Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables

Christopher H. Achen
Department of Political Science
and Institute for Social Research
University of Michigan

Address: 4252 ISR
University of Michigan
PO Box 1248
Ann Arbor, MI 48106-1248

E-mail: achen@umich.edu[1]

November 2, 2001

[1]Prepared for the Annual Meeting of the Political Methodology Section of the American Political Science Association, UCLA, July 20-22, 2000. An preliminary version was prepared for the Annual Meeting of the Midwest Political Science Association, Chicago, April 27-29, 2000. I am very grateful to John Jackson, Larry Bartels, Ido Oren, and Phil Shively for helpful discussions on this topic, and to Jim Granato and Nancy Burns for several conversations about time series analysis. Rob Franzese, Mark Jones, and Chris Lenhardt talked to me about instructive examples from their own recent research of the phenomenon discussed in this paper. The National Science Foundation under grant SBR-9311813 and the Center for Political Studies of the Institute for Social Research, University of Michigan, supplied research support at various stages. Remaining errors are my own.

Abstract

In many time series applications in the social sciences, lagged dependent variables have no obvious causal interpretation, and researchers omit them. When they are left out, the other coefficients take on sensible values. However, when an autoregressive term is put in "as a control," it often acquires a large, statistically significant coefficient and improves the fit dramatically, while many or all of the remaining substantive coefficients collapse to implausibly small and insignificant values. Occasionally, the substantive variables even take on the wrong sign.

This paper explains why this phenomenon occurs and how the resulting confusions have often misled researchers into inaccurate inferences. The standard findings that government budgets are caused primarily by past budgets and that arms races are driven mainly by domestic forces are shown to be likely statistical artifacts. Applications are made to vector autoregressions, error-correction models, and panel studies.

Introduction

Social scientists analyzing time series or panel data often compare two different kinds of specifications. In the first, a dependent variable is regressed on a set of exogenous explanatory factors. The fit may be reasonably successful and the substantive interpretations satisfactory. Thus all seems well. Yet when one or more lagged values of the dependent variable are added "as a control" and the regression is recomputed, in many instances the autoregressive terms are strongly significant and the fit improves sharply, but the original sensible substantive effects of other variables disappear. This pattern frequently occurs even when the lagged variables have no plausible causal interpretation.

For example, Robert Franzese examines welfare spending in 21 OECD countries from 1950-95. (See also Hicks and Swank, 1992.) In some unpublished preliminary studies related to Franzese (2001, chap. 2), his variable to be explained is SWE, defined as social welfare expenditures, including total yearly pensions, benefits including unemployment compensation, and welfare (as defined by the OECD and computed as a percentage of gross domestic product). The explanatory factors are OLD (the percent of population 65 or older) and UNEMP (the percent of the working-age population unemployed according to the internationally comparable definition used by the OECD). The idea is that the more citizens of retirement age, the higher pension payments will be, and that high unemployment will lead to more welfare and unemployment compensation.

If it is assumed that the same linear model applies to all countries, then ordinary least squares (OLS) is consistent. The resulting estimates, based on 683 observations, are:

$$\begin{array}{lllll} \text{SWE} = & -5.36 & +1.34(\text{OLD}) & + .425(\text{UNEMP}) \\ & (.64) & (.054) & (.037) \end{array} \qquad (1)$$

where the standard errors of the coefficients are in parentheses. Here the substantive coefficients make excellent sense, and both have t-ratios greater than 10.[1] The $R^2 = .55$, which is quite respectable.

There is no strong reason to include a lagged value of social welfare benefits here. OECD governments spend a good deal of time carefully planning budgets to meet the needs of the expected number of retirees. After the initial startup years for a social welfare program, last year's expenditures do not matter; this year's needs do. Thus the lagged value of SWE has no obvious causal power. However, if this nonsensical variable, denoted SWE_{-1}, is introduced into the equation, the result is[2]:

$$\begin{array}{lllll} \text{SWE} = & .225 & +.991(\text{SWE}_{-1}) + .016(\text{OLD}) & - .0085(\text{UNEMP}) & (2) \\ & (.130) & (.0074) \qquad\quad (.014) & (.0078) & \end{array}$$

Now only the lagged value seems to matter. It has a coefficient near 1.0 with a t-ratio exceeding 100, which of course is highly significant. Moreover, the R^2 has now improved dramatically to .984, and all the standard errors are smaller. In purely statistical terms, the fit is decisively better. However, the two substantive coefficients have collapsed to about 1% and 2% of their former values, respectively, and neither is statistically significant even at the generous .10 level. UNEMP even has the wrong sign.

In such circumstances, if all else is well, the conventionally respectable interpretation is that the first specification, no matter how sensible, is erroneous, and its implied causal effects are illusory. On the other hand, the strange specification with the lagged variable, no matter how incredible, is preferred. Obeying this logic, applied researchers sometimes talk themselves into accepting outlandish regressions. More often, though, they dislike this

[1] The negative intercept is less plausible and undoubtedly reflects some nonlinearity in the relationship. But it is a forecast of what would happen in a world of no old people and no unemployment—a situation so far from the data as to be nearly irrelevant to the fit.

[2] Two observations are dropped due to missing data for the previous year.

conclusion, claiming that the autoregressive terms somehow "dominate the regression," "swamp the other variables," or otherwise distort the findings. They revert to the original, substantively meaningful specification, though with a sense of guilt about violating sound statistical practice and scientific norms. This leaves them open to the charge of wishful thinking on behalf of their favorite explanatory variables.

The difficulty in the practitioners' argument is that the concept of "dominant variables" has no clear statistical foundation.[3] Regression theory demonstrates that, given proper specification, each explanatory variable is allocated its own proper effect, not that of other variables. No variable can "dominate." Moreover, if all else is well, adding the autoregressive terms to the original specification cannot itself cause bias. If the lagged terms belong there, then they should be included to *avoid* bias. And if they do not belong, then they do no harm: In particular, they cannot systematically alter the other coefficients. The fact that they do dramatically affect the other coefficients and improve the fit seems to argue strongly for their inclusion. Thus the nature of the variables themselves provides no obvious statistical warrant for preferring the weaker specification that excludes the lagged dependent variables.

Then why do specifications with autoregressive terms so often destroy the influence of other variables? In many instances, the *a priori* evidence that a particular exogenous variable has an effect is nearly overwhelming. For example, when the strongly best-fitting regression equation implies that the percentage of old people has a near-zero effect on the level of OECD social welfare benefits, and worse, that the more unemployed people there are, the less it costs to support them, then something is wrong somewhere.[4]

[3] The only textbook treatment I have found occurs in Rao and Miller (1971, pp. 40-43), who discuss the phenomenon in another context, namely the statistical dominance of material inputs when production functions are estimated. They recommend deletion of the dominant variable in some cases.

[4] One might try to save the regression with the dominant lag term by recalling that

This paper uses standard econometric specification analysis to assess the practitioners' argument about "dominant" lagged dependent variables. It turns out that in circumstances often encountered in practice, the applied researchers are right. The lagged dependent variable does bias the substantive coefficients toward negligible values and does artificially inflate the effect of the lagged dependent variable (though not for the reasons practitioners usually give). Some long-standing claims based on apparently strong statistical evidence—that government budgets are caused primarily by last year's budget, for example, or that arms races are driven primarily by domestic forces and not by external threat—are shown to be probable statistical artifacts.

The Case of No True Autoregressive Effect

The nature of the problem is most clearly seen in the simplest case, a bivariate regression in which the lagged dependent variable has no real effect and is therefore excluded:

$$y_t = \beta_0 + \beta_1 x_t + u_t \tag{3}$$

It is assumed that the independent variable is standardized to (stationary) mean zero and is trended over time according to an autoregressive process of order one:

$$x_t = \rho_1 x_{t-1} + e_t \tag{4}$$

Similarly, the disturbance term u_t is composed of unobserved variables trended over time in the same manner:

$$u_t = \rho_2 u_{t-1} + v_t \tag{5}$$

β_1 represents the "impact" or short-term multiplier, while $\beta_1/(1-\gamma)$ is the "total" or long–term multiplier. Thus when γ is near unity, the long-term effect of even a small β_1 could be considerable. But in many applications, this interpretation of the long-term effects is not substantively credible: In the Franzese data, for example, one has to argue that having a lot of old people now does not raise pension payments, but it gradually does so thirty, forty, or fifty years from now (when they are all dead). In addition, this interpretation cannot make sense of coefficients with the wrong sign.

Thus the disturbances are subject to first-order serial correlation. All the other standard regression assumptions are assumed to hold, including stability conditions on the autoregressive processes: $-1 < \rho_1, \rho_2 < 1$. (See the appendix for a list of these assumptions, along with a treatment of the general case of k exogenous variables plus possibly a lagged dependent variable on the right-hand side.)

Now it is well known that, even with the serial correlation, OLS applied to Equation (3) generates unbiased and consistent coefficient estimates. That is, in cases like these, practitioners who omit the meaningless lagged variable from their regressions will get the right coefficients asymptotically. Of course, the nominal OLS standard errors will be wrong, but these are easily corrected, at least in good-sized samples, by one or another version of robust "sandwich" standard errors, which will be asymptotically correct. (See the discussion in Beck and Katz, 1995, 1996.)

Next, suppose that the researcher enters the lagged dependent variable on the right-hand side (with coefficient γ) and re-estimates by OLS. The standard textbook result applies: In the presence of serial correlation, the lagged dependent variable induces bias in all the coefficients. For this case in particular, Equation (18) and footnote (25) of the appendix show that, asymptotically, the estimated coefficient for the exogenous variable is:

$$\text{plim } \hat{\beta}_1 = \left[1 - \rho_1 \rho_2 \, \frac{1 - R^2}{1 - \rho_1^2 R^2}\right] \beta_1 \tag{6}$$

where R^2 is the (asymptotic) squared correlation when OLS is applied to Equation (3).

This equation shows why lagged dependent variables suppress the effects of other variables. For suppose that the exogenous variable is heavily trended, so that $\rho_1 \approx 1$. Then simple substitution shows that, asymptotically, $\hat{\beta}_1$ is reduced in proportion to the fraction ρ_2, the serial correlation parameter. Thus the true effect of the substantive variable is underestimated, and the worse the serial correlation, the worse the bias. When both the exogenous variable and the disturbance term are heavily trended, so that $\rho_1, \rho_2 \approx 1$, the effect of the independent variable disappears almost entirely, as seen so often in applications.

At the same time, Equation (17) of the appendix implies that the lagged dependent variable picks up a non-zero pseudo-effect:

$$\text{plim } \hat{\gamma} = \rho_2 \frac{1 - R^2}{1 - \rho_1^2 R^2} \tag{7}$$

This second equation shows that the bias again gets worse as serial correlation increases. Trending in the exogenous variables is less important for this coefficient, since bias persists even in its absence, but it remains true that heavy trending of the exogenous variable makes matters worse. In that case ($\rho_1 \approx 1$), the effect of the lagged variable, $\hat{\gamma}$, will be asymptotically inflated from its true value zero to nearly ρ_2, the serial correlation coefficient. Then when ρ_2 also approaches unity, the lagged variable will acquire an apparent coefficient of nearly 1.0, again mirroring common practical experience. The lagged variable will seem to be a powerful predictor in spite of its having no real causal impact whatsoever. In this sense, it "dominates" the regression.

The impression of a successful fit will be enhanced in these circumstances by the multiple correlation coefficient, which inevitably rises when the lagged variable is introduced. Indeed, it is easy to show that as the trending coefficients for the exogenous variable and the disturbances both increase toward unity, the R^2 approaches 1.0. For the same reason, the standard errors of the coefficients will decrease dramatically. Thus every sign of good statistical fit will be present except substantive good sense.

What is the culprit here? The answer is a combination of two factors, one familiar and one usually irrelevant, which make a volatile mixture. The first is the well-known bias, typically of modest size, due to using lagged variables when serial correlation is present. The second is heavy trending in the exogenous variable, which makes no difference at all under standard regression assumptions. What seems not to have been noticed previously, however, is that in the presence of a serious version of the first problem, heavy trending in the exogenous variable can generate a catastrophe. An important lagged variable can be created from nothing, and the legitimate impact of everything else destroyed.[5]

Intuitively speaking, the problem is that when a lagged dependent variable is entered into a regression equation with serial correlation, it acts as a proxy, picking up some of the effect of unmeasured variables.[6] However, the autoregressive term does not conduct itself like a decent, well–behaved proxy. Instead, it is a kleptomaniac, picking up the effect, not only of excluded variables, but also of the *included* variables if they are sufficiently trended. As a result, the impact of the included substantive variables is reduced, sometimes to insignificance.

Incremental Budgets or Statistical Error?

In a very influential article, Davis, Dempster and Wildavsky (1966a; see also 1966b) presented extensive statistical evidence that American governmental budgets were set *incrementally*, with the previous year's budget determining most of the subsequent year's budget politics. In particular, last year's budget was an excellent predictor of this year's budget.[7] The detailed and skillfully executed statistical results seemed to reinforce Wildavsky's (1964)

[5]Griliches (1961, pp. 69-70) treats a version of this problem, deriving the bias in $\hat{\gamma}$. But he assumes that the exogenous variable is untrended over time ($\rho_1 = 0$), which minimizes the difficulties. In particular, he does not discuss the bias in $\hat{\beta}_1$, and indeed, there is none under his assumptions, as Equation (6) shows.

[6]Sometimes researchers use lagged dependent variables deliberately for that reason, as in Burkhart and Lewis-Beck (1994, p. 905).

[7]More precisely, they showed that last year's budget closely predicted this year's agency request, and this year's request closely predicted this year's budget.

earlier case studies of the budget process, and they also seemed to confirm the theories of bounded rationality coming from Lindblom (1959), Simon (1955), and others at the time.

Much investigation by other scholars came to similar conclusions. Sharkansky (1968) reviewed the literature on American state budgets, saying that "there is virtually no association between state scores on expenditures per capita and the measures of political characteristics considered here" (p. 63). Furthermore, "across the full range of government programs, the factors likely to correspond most with current levels of activity are previous levels of activity" (p. 152). The implicit vision of boundedly rational, incremental decisionmaking continues to dominate the thinking of most students of the policy process.

In the interim, however, the intellectual basis for incrementalist thinking has come under serious attack. As games of incomplete information became better understood, scholars increasingly saw the incrementalist viewpoint as too limited to explain even the substantial variation in budget success which Davis, Dempster, and Wildavsky had pointed to in their original articles. In one of the early statements of the new viewpoint, McCubbins and Schwartz (1984) argued that fully rational Members of Congress dealing with the bureaucracy would look as if they were behaving incrementally most of the time, even if they were not. More importantly, McCubbins and Schwartz suggested conditions under which Congressional decisions would change non-incrementally, thereby matching better what had been seen in the budget data. Additional research has further exposed the limited explanatory power of "budgetary incrementalism" (see the review in Berry, 1990).

The rational choice perspective has seemed to a growing minority of scholars simply more powerful than incrementalism and better able to account for the full range of evidence. The evidence accounted for includes non-quantitative analyses of budgeting. In fact, rereading Wildavsky's (1964) case studies from the modern perspective is a revelation. The book validates in detail all the main assumptions of the rational choice approach to the topic, and surprisingly, it reads now as a thorough-going critique of the limited-rationality, incrementalist inferences he and most political scientists drew from its contents. Thus, although proper interpretation had to wait on further intellectual developments, the case study methods of the 1960s essentially got the modern story right.

So if lagged budgets are not remarkably influential, why were the statistical studies so wrong? Essentially, the answer lies in the bias results derived earlier: The effects of lagged dependent variables are grossly overestimated when disturbances are heavily trended. Translated to budget data, this means that last year's budget will predict this year's budget very well even if it has little or no real causal impact, so long as other unmeasured influential political factors are strongly trended—much the same from one year to the next.[8] Heavy trending seems likely in budget data most of the time, and if so, the result would be an artificially inflated estimate of the causal power of last year's budget.

Davis, Dempster, and Wildavsky (1966b) provide a partial test for trending problems by computing the Durbin-Watson statistics for their disturbances, and they find that the statistics are not often significant. However, their sample sizes for each government agency are very short—never more than 16 observations, and usually fewer than 10—leaving them with little statistical power to detect non-zero correlations. Even so, many of their test statistics are worrisome, and some achieve statistical significance even with the extremely short time series used, indicating that the underlying correlations are probably very high. Indeed, at one point (1966a, p. 536), they note

[8] An early hint of this result: Looking at arms race data and seeing his lag coefficient drop from .55 to .31 after correction for serial correlation, Hibbs (1974, p. 302) wondered in a brief footnote about the true explanatory power of incrementalism. Raisa Deber made similar brief remarks about her estimates for the same data (Choucri and North, 1975, p. 329.

that when serial correlation was large enough to be statistically significant for a particular agency, their preferred model always had the highest R^2, just as Equations (6) and (7) would predict.

Where budget series are longer, the suspicion that incrementalist findings depend on heavy trending can often be confirmed. In Franzese's budget data, for example, the serial correlation in Equation (1) is .99. This is precisely the condition under which incrementalist effects would be seriously exaggerated and political influences minimized, as the Franzese Equation (2) demonstrates. Subsequent research with longer series, particularly series that include abrupt changes in the political climate to occasionally break the heavy trending, have come to quite different conclusions than the incrementalism literature, just as the bias equations set out earlier would predict. (See, for example, Kiewiet and McCubbins, 1985; see also the dramatic tripling of Irish arts spending in 1973 discussed in Gallagher, Laver, and Mair, 1992).[9]

This argument brings the statistical results into line with the rational choice critique of the qualitative literature. Games of incomplete information show that fully rational players will misleadingly appear to be merely boundedly rational in the usual case when the political environment is stable, but not at other times. Similarly, Equations (6) and (7) show that lagged budgets will falsely appear to be the sole cause of future budgets when the political environment is stable, but not otherwise. Like Wildavsky's often misinterpreted case studies, the Davis, Dempster, and Wildavsky statistical results are completely consistent with contemporary theory. Budget data *are* generally quite stable over time and well predicted by the prior year, but nothing follows about the bounded rationality or incrementalist thinking of

[9] In Franzese's data, the simple step of entering dummy variables for the various countries in the sample reduces the bias, though it does not eliminate it, since within-country trending remains.

the decisionmakers.[10]

When Autoregressive Terms Have Causal Power

In the first part of this paper, the lagged variable was assumed to have no genuine effect on the dependent variable. The example was government budgets: In equilibrium, budgets ought to reflect the felt needs and political pressures of the current year, not last year's. However, budgets may be out of equilibrium for certain periods due to startup friction, unforeseen economic or political contingencies, or just routine year-to-year adjustment costs. Thus even budgets may be subject to greater or lesser degrees of "path dependency," implying that lag terms have at least some causal power. More dramatically, the Bayesian models now in use in public opinion research imply a strong causal dependence on the previous period's lag. Other examples might easily be multiplied.[11]

To examine this situation in a simple case, replace Equation (3) with a specification allowing for a one-period lag:

$$y_t = \beta_0 + \gamma y_{t-1} + \beta_1 x_t + u_t \tag{8}$$

The other assumptions are maintained, notably Equations (4) and (5), which specify first-order autoregressive trending in both the exogenous variable and the disturbance.

With its autoregressive term, Equation (8) is a generalization of the previous Equation (3), of course, and the biases induced in this more general setup by OLS have received greater attention, though the great majority

[10] As noted in footnote (7), Davis, Dempster, and Wildavsky actually use a two-equation system with both serial and contemporaneous correlation, but the same remarks apply.

[11] Sometimes legitimate autoregressive terms appear in statistical models for non-causal reasons, for example as an alternate way to express a model that makes it easier to estimate. The classic instance is the geometric distributed lag (see, for example, Malinvaud, 1970, pp. 124-30; Greene, 1990, pp. 525-538.)

of econometric texts do not discuss the topic in any detail.[12] Griliches (1961) derived the asymptotic bias in the lag coefficient when the exogenous variable is untrended, noting that it would be biased upward when serial correlation was positive. Malinvaud (1970, pp. 558-561), Maddala and Rao (1973, appendix A), and Phillips and Wickens (1978, pp. 376, 412-417) derived in different ways the bias for all coefficients under general conditions on the time path of the exogenous variable, and Malinvaud noted briefly that trending of the exogenous variable adds to the biases.[13] However, none addressed the issues raised by coefficient patterns like those in the Franzese data.

Malinvaud (1970, p. 560-561) also gave some simulation results, most with modest biases, and he showed that adding exogenous variables would reduce the bias further (see footnote 29 of the appendix). The impression left is that the biases are real but not devastating. Subsequent discussions in political science by the best methodologists have generally followed in the same spirit. Hibbs (1974, pp. 289-94) used Griliches's methods to study the same case, and he suggested that "typically" the coefficient on the lagged variable will be biased upward and the coefficient on the exogenous variable will be attenuated, so that the lagged variables "tend to dominate the outcomes of such equations (partially)." However, he also cited Malinvaud in saying that these biases would be less severe when exogenous variables were present, and indeed, the moderate biases in his empirical examples were similar to those Malinvaud obtained in his simulations. Likewise, Rattinger (1976, p. 430) wrote that "Malinvaud (1970) has shown analytically and by Monte Carlo experiments that least squares bias in the estimation of

[12] An exception is Maddala, 1992, pp. 248-249, who mentions the direction of bias in the autoregressive term and in the serial correlation coefficient, but does not discuss the bias in the coefficient of the exogenous variable.

[13] Phillips and Wickens (1978, pp. pp. 376, 417) also derive the OLS bias in this model when the disturbances are moving average of order one, which may turn out in the end to be the more relevant case for government budget data.

autoregressive processes with autocorrelated disturbances is greatly reduced by the presence of exogenous variables."

In fact, however, dramatic biases are possible with arbitrary numbers of exogenous variables, just as they were in the simpler case with the lagged variable excluded. In the present case, the asymptotic values of $\hat{\beta}_1$ and $\hat{\gamma}$ are:

$$\text{plim } \hat{\beta}_1 = \left[1 - \frac{\rho_1 \rho_2 \sigma^2}{(1-\rho_1\gamma)(1-\rho_2\gamma)s^2}\right]\beta_1 \tag{9}$$

and

$$\text{plim } \hat{\gamma} = \gamma + \left[\frac{\rho_2 \sigma^2}{(1-\rho_2\gamma)s^2}\right] \tag{10}$$

where $s^2 = \sigma_y^2 - \rho_1^2 {\beta_1}^2 \text{var}(x_t)/(1-\rho_1\gamma)^2$.

The interpretation of these equations is somewhat less obvious than when the lag term is absent, but the central features are the same. When $\rho_1, \rho_1 > 0$, the appendix shows that asymptotically, $0 < |\hat{\beta}_1| < |\beta_1|$, so that the size of the coefficient on the exogenous variable is always underestimated. Similarly and asymptotically, $\gamma \leq \hat{\gamma} < 1$, so that the estimated lag coefficient is always too high but bounded above by unity. In both cases, the bias increases with the degree of serial correlation. When both the exogenous variable and the disturbance are heavily trended, then the limits of the biases are approached: The lagged variable will have a coefficient of nearly 1.0 and the exogenous variable coefficient will disappear. The appendix shows that this result remains true when additional exogenous variables are included in the regression, so long as they are heavily trended as well.

Thus the lagged variable can artificially dominate the regression whether it has a great deal of explanatory power, a little, or none at all. For applied researchers, the most disturbing aspect of this result is that the bias on the lag variable is not additive. That is, the estimated coefficient does not equal (true value) + (bias), as it does in omitted variables contexts. When the exogenous variable and disturbances are heavily trended, regressions with a powerful lag term will wipe out the other variables, but so will regressions with a weak or impotent lag term. *In the presence of heavy trending in exogenous variables and disturbances, lagged dependent variables will dominate the regression and destroy the effect of other variables whether they have any true causal power or not.* For example, lagged budgets proxy for both their own causal effect and that of unmeasured political factors; OLS lumps both into the coefficient on the lag term. Thus whether or not lagged budgets have some causal power, there is nothing to learn about incrementalism in budgetary decisionmaking from the fact that past budgets predict future budgets very well.

Zinnes's Arms Race Puzzle

Richardson's (1960) arms race model may be written as:

$$y_t = \beta_{10} + \gamma_1 y_{t-1} + \beta_{11} x_{t-1} + u_{1t} \tag{11}$$

$$x_t = \beta_{20} + \gamma_2 x_{t-1} + \beta_{21} y_{t-1} + u_{2t} \tag{12}$$

where x_t and y_t are the armament expenditures of the two sides, typically measured annually. Sometimes measures of hostility or of military effectiveness are substituted for military budgets.

All the coefficients are expected to take on strictly positive values, with the lag coefficients γ_1 and γ_2 interpreted as the "drag" effects of increasing expenditures, while β_{11} and β_{21} are regarded as the "reaction" effects to the other sides' spending.[14] By assuming that the two disturbance terms are serially uncorrelated, either equation may be estimated by ordinary least squares, possibly with additional control variables.

In a widely cited presidential address to the International Studies Association, Dina Zinnes (1980) posed the following puzzle for international

[14] The Richardson equations are often written equivalently with year-to-year changes in armaments on the left-hand side. In that case, the drag coefficients are expected to be negative.

relations researchers. In the empirical literature on arms races, why did the two sides so often appear to ignore each other? That is, why were the estimated values of the drag coefficients so often nearly unity and the reaction coefficients β_{11} and β_{21} so often small and insignificant? Many other reviews have puzzled over the same oddity (Moll and Luebbert, 1980; Russett, 1983; Goldstein and Freeman, 1990, pp. 22-29; Wiberg, 1990).

Some scholars have seen these statistical results, not as an anomaly, but as a finding. In their view, the dominant drag coefficients mean that arms races are driven by a domestic "Eigendynamik" stemming from bureaucratic politics, organizational routines, and/or some form of military-industrial complex. Gleditsch (1990, pp. 1-2) congratulated German social scientists for their intellectual leadership during the Sixties and after in promoting this view that state power has little or no effect on the behavior of one's enemies in the international system. How many r.p.m. this encomium induced in the graves of Ranke and Treitschke no one knows.

More seriously, Rattinger (1976, pp. 430-431), having encountered the same statistical enigma in his own work (Rattinger, 1975), criticized the lack of attention to time series econometrics in the arms race literature. Citing Malinvaud (1970) and Hibbs (1974), among others, he suggested that the combination of lagged dependent variables and serial correlation might cause bias in the estimation of arms race models, though he did not spell out the nature of those biases.

In fact, either one of the arms race equations is just a special case of the lagged regression under study in this paper, namely Equation (8). The lagged variable is the Eigendynamik term, while the "exogenous" variable is the activity of the other state. Moreover, although few authors in this large literature take account of serial correlation, in those few cases where it has been estimated, it appears to be quite substantial. The best early study was done by Raisa Deber in Appendix B of Choucri and North (1975, pp. 326-327). In her models (8) and (9) applied to pre-World War I British data, she uses simultaneous equation techniques to get a consistent estimate of second-order autoregressive disturbances in the presence of lags.[15] Her first-period lag serial coefficient is .6 in one specification and 1.0 in the other. Similarly, in the same specification for the disturbances, Oren (1996, p. 322) gets first-period serial coefficients for the United States and the Soviet Union during the Cold War of 1.1 and .6, respectively.

Moreover, in those cases where serial correlation was corrected, it reduced the lag effect and increased at least some of the other coefficients, just as Equations (9) and (10) predict. In Deber's work, for example, the autoregressive coefficient for the dependent variable was cut in half from the OLS value after the correction for serial correlation, and at least some of the other exogenous coefficients grew in size. Hibbs (1974, pp. 299-303) obtained similar results with what is apparently an earlier version of the same data. Both Deber's and Hibb's corrections might have had even larger effects had they not included the other side's current arms spending, presumably a highly endogenous variable, as one of their instruments.

Thus the bias equations (9) and (10) show that Rattinger's concerns are amply confirmed by econometric theory. When Richardson-type equations are estimated by OLS, Eigendynamik effects are exaggerated and state-to-state reactions are attenuated. Indeed, the potential size of the biases is larger than anyone studying arms races could have imagined at the time.

Zinnes's puzzle appears to be resolved: States *do* interact during arms races, but the statistical techniques conventionally used to study them suppress the interaction and inflate the lagged Eigendynamik effect. Full confidence in this conclusion awaits additional empirical testing, of course, and studies of arms races have been in short supply since the end of the Cold

[15]The Durbin-Watson statistic and the usual generalized least squares correction are useless, of course, due to the presence of the lagged dependent variable. More sophisticated methods must be used to estimate and correct serial correlation in arms race models.

War. However, as the Indian subcontinent and the Taiwan Strait redirect political scientists' attention to the subject, research opportunities should present themselves. It is time Richardson received a fair test.

Can Lags Reverse the Sign of Other Coefficients?

Thus far, nothing in the argument has indicated that sign reversals are possible when autoregressive terms are used in the presence of serial correlation. Yet dubious negative coefficients turn up with disconcerting frequency in studies of government budgets. In the Franzese dataset with which this paper began, for example, the combination of a lagged dependent variable and substantial serial correlation produced an anomalous negative coefficient, though it was not significantly different from zero.

Similarly, Zeigler and Johnson (1972, pp. 60-61) studied 1961 education expenditures per capita by American states (y_{61}) as a function of state and local revenue per capita in the same year (x_{61}), plus other variables. State revenues supply the funds from which education spending is taken, so that a strong positive relationship is expected. Indeed, the bivariate correlation is .87. When the full regression was computed, these expectations were confirmed. Suppressing the intercept and coefficients on other exogenous variables, the result was ($R^2 = .99$):

$$y_{61} = \underset{(.0279)}{.260}\, x_{61} \qquad (13)$$

Thus marginally, about 26% of state and local revenues go to education, which is a very reasonable figure, and the effect is highly statistically significant, with a t-ratio near 10.

Zeigler and Johnson also examined the same relationship in 1965. Again the bivariate correlation between spending and 1961 revenues was quite high (.82), and the same result might be expected. However, this time they also controlled for lagged education expenditures. Again with the intercept and coefficients on a similar set of other exogenous variables suppressed, the result was ($R^2 = .97$):

$$y_{65} = \underset{(.0579)}{-.121}\, x_{61} + \underset{(.123)}{1.17}\, y_{61} \qquad (14)$$

Now the effect of revenue remains statistically significant, but it is bizarrely negative, indicating firmly that the more money states have, the less they want to spend on education. The fact that no revenue data were available for 1965 is a possible culprit, as is the change in the list of exogenous variables, and other forms of specification error. Undoubtedly realizing all these concerns, Zeigler and Johnson, like most researchers encountering the occasional odd negative coefficient when an autoregressive term has been added, simply do not remark on it. Occasional other, less dramatic examples, also unremarked, may be found elsewhere in the same chapter (Zeigler and Johnson, 1972, chap. 3).

As a final example, consider Organski and Kugler's (1980, pp. 190-192, and footnote 33 on pp. 267-268) OLS estimates for a Richardson arms race model. Here y_t is the arms expenditure of the United States and x_t that of the Soviet Union over a 22 year period, 1955-1976 (with 1965 dropped as an outlier):

$$y_t = \underset{(2.12)}{7.83} + \underset{(.092)}{.720}\, y_{t-1} - \underset{(.125)}{.451}\, x_{t-1} \qquad (15)$$

$$x_t = \underset{(1.29)}{1.93} + \underset{(.790)}{.920}\, x_{t-1} - \underset{(.059)}{.038}\, y_{t-1} \qquad (16)$$

The R^2 statistics are .92 and .94, respectively, so that fit is excellent.[16]

[16] The coefficients from the 1952-1976 fit with no outliers dropped are quite similar in character, though the first negative coefficient is then significant only at the .10 level rather than .01.

Strangely, however, both reaction coefficients are negative, with one of them highly statistically significant ($t > 3.5$). The implication is that the more one side built up, the fewer armaments the other side bought. In their interpretation, Organski and Kugler (1980, 191-197) treat their reaction coefficients as zero and accept the argument that the arms race was internally driven, though they warn that some caution is warranted due to the usual statistical gremlins.

In an important article, Oren (1996, pp. 323-327) noticed that a great many arms race researchers had found negative reaction coefficients in their models, just as Organski and Kugler had. In twenty published studies of US-USSR interactions during the Cold War, he found that fourteen (70%) had at least one negative reaction coefficient. Some of the aberrant values were insignificant or only marginally significant, but they occurred with disturbing regularity. Moreover, as Oren notes, these anomalous coefficients came from the *published* regressions; they were probably even more common in the regression runs the researchers threw away. In print, though, everyone just dismissed the abnormalities as accidental, or ignored them entirely.

Why the anomalous coefficients with reversed signs when a lagged dependent variable is added to a regression?[17] Thus far it has been assumed that all variables were stationary, in particular that their means were stable over time. In those circumstances and with enough data, reversed signs are impossible, as the appendix demonstrates.

Even a random variable that becomes stationary in the long run, however, may have a non-constant mean during a start-up period. (An introductory treatment is Enders, 1995, pp. 38-44.) For example, social welfare expenditure as a percent of government spending is presumably asymptotically stable in this sense (Franzese's data), at least when controlled for the relevant variables. The same could be said for arms spending as a percent

[17] To my knowledge, this question is a new one in the econometric and political science literature.

of GNP, controlled for what the other side is doing (Oren, 1996). However, both kinds of data have periods of sharp upward drift, particularly at the beginning, which can pull parameters off their long–term values.

Even worse, the budgets in dollar amounts used in so many studies are not stationary even asymptotically: They nearly always drift upward, simply because government budgets grow. Upward movement will be particularly strong in arms race expenditures; in some sense, that is what defines an arms race. In consequence, either the lag coefficient γ, the autocorrelation coefficient for the exogenous variable ρ_1, or the serial lag coefficient ρ_2 will necessarily exceed 1.0. When researchers run regressions on steadily increasing variables with these characteristics, several of the assumptions made so far break down.

In circumstances like these, when one of the parameters γ, ρ_1, or ρ_2 exceeds unity, reversed signs for the estimated coefficients on exogenous variables become possible. The other two parameters need not even be dramatically large. The most obvious case is $\rho_1 > 1$, meaning upward drift in the exogenous variable. The proof in the appendix at Equation (28), which demonstrates that asymptotically the estimated coefficient on the exogenous variable cannot reverse sign, clearly fails when $\rho_1 > 1$. In that case, it is not hard to find plausible values of the other two parameters that suffice for a sign reversal—see Equation (9). Similar remarks may be made about the cases in which $\gamma > 1$ or $\rho_2 > 1$.[18]

Though it is impossible to be certain without substantial reanalysis of the original data, inspection of the cited empirical examples with a reversed

[18] This argument is only heuristic. One cannot simply plug the values of the parameters into Equation (9) to learn the probability limits, since neither it nor Equation (10) holds when the data are nonstationary. In that case, both equations must be regarded solely as rough-and-ready approximations to the finite sample expectations, which are likely to take on nonintuitive values under nonstationary conditions. A subsequent and more serious investigation is needed to adequately address the probability limits of the regression coefficients in Equations (9) and (10) under nonstationarity.

sign shows that one or another of these key parameters probably exceeded unity in each case.[19] Thus the puzzle of negative substantive coefficients may be resolved in the same manner as the other empirical anomalies found in budget data. Periods of steady upward drift in the variables, whether transitory or permanent, are the guilty parties. In the presence of one parameter exceeding unity, the others need be only of moderate size to create dominant autoregressive terms and small or even wrong-sign coefficients on other variables.[20]

Conclusion and Recommendations

In time series applications such as government budget studies, arms race models, and many others, plausible functional forms with no autoregressive terms often produce theoretically meaningful coefficients in a modestly successful fit. However, when one or more lagged dependent variables are added as explanatory factors, the autoregressive terms take on strongly significant coefficients which improve the fit but squash the effects of the other variables. The conventional inference is that the original variables make no real difference.

This paper has argued that in practice, the anomaly is often due to the combination of high serial correlation and heavy trending in the exogenous variables, which can jointly produce dominating autoregressive terms even when they have little or no real explanatory power.

[19] For arms race dollar budgets, ρ_1 is the autoregressive parameter for the other side's armaments. It nearly always exceeds unity in historical arms races, sometimes substantially. The case $\rho_2 > 1$ in arms race data has already been discussed, and $\hat{\gamma} > 1$ in the education spending example of Equation (14). Finally, note the danger here of looking just at autocorrelations, since they are bounded below unity in absolute value even when the corresponding autoregressive parameter is not.

[20] It should be noted that the discoverer of the aberration (Oren, 1996) believes that the explanation lies elsewhere and, in particular, that the negative coefficients are substantively meaningful. His own negative estimates occur in equations corrected for serial correlation, which may suggest that he is correct, at least for his own data. However, see the next section.

Readers skimming this paper may wish to be reminded that the converse does not follow. That is, the following inference is fallacious: "If I find a dominant autoregressive term, heavy trending in the exogenous variables, and substantial serial correlation when the lag term is omitted, then the autoregressive term is causally unimportant." To the contrary, the autoregressive term may or may not be powerful. The correct inference is instead that the model should be re-estimated with techniques suitable for the combination of lagged dependent variables and serial correlation. When the variables are stationary, various relevant methods for leading cases may be found in editions of standard texts such as Gujarati (1995, pp. 604-612) or Greene (1993, pp. 419-28), and in sophisticated political science applications such as Wood (1992).

In one sense, therefore, nothing in this paper is truly new. The appendix extends only slightly what is already known. Both the potential bias of using lagged endogenous variables in the presence of serial correlation and the methods for coping with it have long been familiar to political scientists (for example, Hibbs, 1974, and Choucri and North, 1975). The new message is simply that everyday biases in this situation may be spectacularly larger and more damaging than the literature suggests and than most researchers have supposed. In particular, budget data, which constitute so many of the datasets in political economy, subnational politics, administrative behavior, and arms control studies, warrant extraordinary concern.

This message may especially deserve repeating as a prolegomenon to the new time series methods entering the profession, where the point is more subtle. A vector autoregression (VAR), for example, regresses the current value of all variables on their own lags and the lags of all other variables in the model. (Freeman, Williams and Lin, 1989, pioneered applications to political science; the Organski-Kugler version of the Richardson model in Equations (15) and (16) is a very simple case in point.) Under suitable conditions, statistical theory tells us that with enough lags (possibly an infinite number) on all variables, serial correlation will be accounted for by the lags and may be ignored. The disturbance will then be white noise, as desired.

With this theoretical background in mind, it is easy to slip into economizing on parameters by using just a few lags on a few variables and supposing that the lagged variables are "predetermined," that is, uncorrelated with the disturbances. In practice, no one worries about serial correlation in VAR models, even though it would make the lagged variables endogenous. The literature warns, however, that omitted variables (of which serial correlation is a special case—see the appendix, for example) can bias VAR coefficients substantially (Harvey 1981, pp. 144-150; Lütkepohl, 1982, pp. 143-59). Thus it pays to include as many relevant variables as possible. Moreover, with the relatively short series available in many applications, the biasing possibility of serial correlation cannot be ignored in applied VAR models. Variables can be determined at a prior time period without being predetermined.

The same biases can afflict error-correction models (see, for example, Banerjee *et al.*, 1993). Define $\Delta y_t = y_t - y_{t-1}$, and similarly for Δx_t. Then a simple error-correction model is:

$$\Delta y_t = \delta \Delta x_t - (1-\gamma)(y_{t-1} - \beta_0 - \beta_1 x_{t-1}) + u_t \tag{17}$$

The idea is, first, that recent deviations in x_t cause changes in y_t, and second, that the long-term equilibrium relationship is $y_{t-1} - \beta_0 - \beta_1 x_{t-1} = 0$. Deviations from it tend to bring y_t back into line.

An equivalent expression for the previous equation is:

$$y_t = \beta_0^* + \gamma y_{t-1} + \beta_1^* x_t + \beta_2^* x_{t-1} + u_t \tag{18}$$

where $\beta_0^* = (1-\gamma)\beta_0$, $\beta_1^* = \delta$, and $\beta_2^* = (1-\gamma)\beta_1 - \delta$. This equation is just a slight extension of Equation (8), now with an additional variable on the right. Thus in the presence of serial correlation due to unmeasured variables or too few lags, bias will enter here as well. In particular, if serial correlation is positive and substantial and if the exogenous variable is heavily trended, γ will be overestimated to be nearly unity, leading to the erroneous conclusion in Equation (17) that the equilibrating force is weak or nonexistent.

The same point applies even more forcefully to panel studies, including time–series of cross–section datasets, where the number of time periods is usually small and may be as few as two or three. Serial correlation often appears in data of this kind, where it causes problems for the estimation of the variance-covariance matrix of the coefficients.[21] To cope with serial correlation, Beck and Katz (1996, p. 5), two of the very best time series specialists in the discipline, make the following suggestion. Since autoregressive terms are always available in panel studies, and since these lagged variables are known theoretically to be substitutes for serially correlated errors, let us use an autoregressive term or two to mop up the serial correlation. Then if that is successful, the model can be estimated by OLS, with the standard errors of the coefficients estimated relatively simply by "panel corrected standard errors," which will have greater efficiency than the standard robust sandwich estimator.

The one blemish on this otherwise attractive proposal stems from the issue discussed throughout this paper: If serial correlation is present and autoregressive terms are entered into the regression equation, then all the coefficients are biased. In particular, when serial correlation is high and the exogenous variables are heavily trended, as will happen frequently in panel data, the lagged variable will falsely dominate the regression and suppress the legitimate effects of the other variables.

Beck and Katz's own empirical example (1996, p. 27) seems to show exactly this syndrome at work. In reanalyzing 1972-1989 data from 131 coun-

[21] Stimson, 1985, lucidly and skillfully surveys the special problems of this topic for political science.

tries reported in Burkhart and Lewis-Beck (1994), Beck and Katz regress a democracy score (y_t) on its lag (y_{-1}), the log of energy consumption per capita (x_1), the interaction of energy with membership in the sub-periphery or periphery of the international system (x_2 and x_3, respectively), and a constant term. They correct for heteroskedasticity using panel-weighted least squares. Their estimates are:

$$\begin{aligned} y_t &= \underset{(.019)}{.004} + \underset{(.002)}{.995}\, y_{t-1} + \underset{(.007)}{.016}\, x_1 + \underset{(.005)}{.006}\, x_2 - \underset{(.007)}{.008}\, x_3 \end{aligned} \tag{19}$$

Here the effect of the lagged variable is dramatically large and the other coefficients nearly vanish, just as the bias equations (9) and (10) would predict if serial correlation were biasing the estimates.

By contrast, using one particular technique for correcting the serial correlation ($\rho_2 = .9$ by their estimates), Burkhart and Lewis-Beck obtained for the same data and the same list of regressors:

$$\begin{aligned} y_t &= \underset{(.06)}{.35} + \underset{(.02)}{.09}\, y_{t-1} + \underset{(.22)}{2.49}\, x_1 - \underset{(.19)}{1.33}\, x_2 - \underset{(.18)}{1.54}\, x_3 \end{aligned} \tag{20}$$

Now the substantive coefficients are far more consequential and the lag effect is modest, indicating that serial correlation bias in the previous equation was probably sizeable. Of course, other estimation techniques might give somewhat different coefficients. The point is simply that autoregressive terms cannot be used to control for serial correlation without taking account of their impact on other coefficients.[22]

Obtaining truly adequate coefficient estimates and standard errors in panel studies or any other time series context with autoregressive terms and serial correlation is no simple matter. There is little reason to think that the

[22] Maddala (1999, p. 61) expresses the same concern about Beck and Katz's (1996) recommendation to use lags to eliminate serial correlation. By contrast, Beck and Katz's deservedly famous (1995) article gives excellent advice on this topic.

first-order lags and first-order serial correlation models that predominate in the best empirical applications are faithful to the data. Yet our datasets are typically too small to teach us the right lag order even if we look for it (Griliches, 1967, pp. 17-18; Rattinger, 1976, pp. 426-428).

In consequence, even some of our best empirical work leaves some questions unanswered. Choucri and North (1975, pp. 210-213) estimated several models for the military spending of Britain, France, Germany, Italy, Russia, and Austria-Hungary in the pre-World War I period. Since lagged dependent variables were used, they tested for serial correlation up to second order. If present, it was corrected for. Almost no other scholars investigating arms races have proceeded so carefully. In spite of the corrections, though, some of the Durbin–Watson statistics fell somewhat short of the magical 2.0, indicating that small amounts of serial correlation remained. As statistical crimes go, this is a parking meter violation.

Even such seemingly routine peccadilloes had dramatic consequences, however. Among the three instances in which the Durbin-Watson fell below 2.0 in the Choucri-North results, every single equation produced a meaningless negative coefficient on the arms spending of the other side, just as in the Organski-Kugler example. Two of these three anomalous coefficients were statistically significant, and the third nearly was. On the other hand, among the six equations with Durbin-Watson statistics exceeding 2.0, not a single one produced a negative reaction coefficient.

These Choucri-North coefficient values fall into the pattern expected from the bias equations above—misleading estimates whose value is a function of the serial correlation coefficient. Choucri and North's work was scrupulously careful; it seems likely that the time series was too short for anyone to get the model exactly right.[23] The only surprise is that it mattered so much. Alas, most of us will find ourselves in the same situation

[23] Similar concerns may apply to Oren's (1996) excellent arms race study, also with negative reaction coefficients.

when we come to our own time-series estimations, facing too many questions with too few observations and knowing that wrong answers are more than usually perilous.

We will need help both from theory and from data. As Griliches (1967, p. 46) puts it,

> Here are a few commandments for virtuous living:...do not expect the data to give a clear-cut answer about the form of the lag. The world is not that benevolent. One should try to get more implications from theory about the correct form of the lag and impose it on the data....Finally, not all is hopeless, but to get better answers to such complicated questions we shall need better data and much larger samples.

Thus we need answers to questions like these: For some government agencies with budget series long enough to permit a firm answer, what is the time series structure of their budgets? Might the log of the ratio of this year's budget to last year's turn out to be a first-order moving average process for most government agencies, for example? From the theoretical side, which time series structure for government budgets is implied by a "fire alarm" model of Congressional oversight? Is it confirmed well enough in our best data that we can safely postulate the same structure in other budget studies with short series? And then finally, how should our models be formulated to learn how politics affects budgets through the time series process that governs them?

These are not small research topics, and they may seem time consuming and narrowly technical even to most quantitative researchers, who just want to get on with the work. Yet good answers would do more to address the serious moral issues of domestic and international politics than the underpowered and dangerously misleading "relevant" research that forever plagues the profession. "To be forward with moral blame before we have diagnosed the nature of the evil is as profitless as the exercise of those who, from piety, used to thrash lunatics in order to drive out the devils of which they were believed to be possessed" (Catlin, 1927, p.293).

Finally, the argument of this paper has focused on the stationary case to make it clear that the biases under study are distinct from those occurring in "nonsense regressions" with nonstationary data (Granger and Newbold, 1986, pp. 205-215). In fact, though, most of the empirical examples set out in this paper use government budget data in dollar form and thus probably contain nonstationary variables. The arguments of the appendix collapse in that case, as do those of Griliches, Malinvaud, and others cited earlier. Standard fixes for serial correlation also fail to work, and VARs and error-correction models have to be treated differently than they have been here. As the vast recent literature on unit root econometrics has demonstrated, nonstationarity is not a minor technical problem. (See recent textbooks such as Enders (1995) and Hamilton (1994), along with *Political Analysis*, vol. 4 (1992) for several articles surveying the topic for political science.)

The new time series methods will reorient how budgets and arms races are studied. Richardson's model, for example, may amount to a claim that the arms stocks of the two sides are *cointegrated* and thus representable in an error-correction model, perhaps in conjunction with total governmental budgets or gross national products. If so, we will have to return to the appropriate datasets with strong formal theory and better time-series tools. Most of the relevant material has traditionally not been taught in standard econometric courses for political scientists. The consequences may be seen in the seriously misleading inferences in several different subfields that this paper has attempted to document.

Appendix

The purpose of this appendix is to study a slightly more general version of Equations (4), (5), and (8) in the text, spelling out the assumptions, deriving the OLS bias, giving bounds on the coefficients, and showing that the probability limits of the OLS coefficients are monotonic in the serial correlation coefficient.

To that end, suppose that $y = [y_1, y_2, \ldots, y_T]'$ is a weakly stationary[24] T-dimensional vector of observations on a dependent variable, and define y_{-1} to be the T-dimensional vector of one-period lagged values of y. This notation is meant to encompass two different sampling designs. The first is the univariate time series case, in which all observations refer to a single sampling unit and y_t denotes the observation on that unit at time t. In that case, $y_{-1} = [y_0, y_1, \ldots, y_{T-1}]'$. The second design is the panel study with several different units of observation and two or more time periods. Then y is the vector whose elements contain observations on all units at every time period after the initial period, while y_{-1} is the corresponding vector lagged one time period.

Similarly, X denotes a $T \times k$ matrix of observations on k explanatory variables, and X_{-1} denotes the corresponding matrix of one-period lagged values. Both y and X are taken to be mean-deviated, so that X contains no constant.

The matrix X is partitioned as $X = [X_1 \; X_2]$, where X_1 and X_2 have k_1 and $k - k_1$ columns, respectively. X_1 is the matrix of all causal variables that are observed, while X_2 contains those that are not observed and thus enter the disturbance. The corresponding lagged values are $X_{1,-1}$ and $X_{2,-1}$.

[24] A scalar random variable x_t observed at time periods indexed by t is weakly stationary (or covariance-stationary) if at each time period it has the same mean and variance and its autocovariances depend only on the lag length. The definition extends in an obvious way to the case in which x_t is a vector. See, for example, Hamilton (1994, pp. 45-46).

It is assumed that:

$$\begin{aligned} y &= \gamma y_{-1} + X_1\beta_1 + X_2\beta_2 \\ &= \gamma y_{-1} + X_1\beta_1 + u \end{aligned} \tag{1}$$

where γ, β_1, and β_2 are parameter vectors of dimension 1, k_1, and $k - k_1$, respectively, with $-1 < \gamma < 1$, and $u = X_2\beta_2$ is a disturbance term taken to be uncorrelated in probability with the observed explanatory variables:

$$\text{plim } X_1{}'X_2/T = 0 \tag{2}$$

Both the observed and the unobserved exogenous variables evolve according an autoregressive process of order one:

$$X_1 = \rho_1 X_{1,-1} + E_1 \tag{3}$$

and

$$X_2 = \rho_2 X_{2,-1} + E_2 \tag{4}$$

where ρ_1, ρ_2 are scalar parameters, with $-1 < \rho_1, \rho_2 < 1$, and E_1 and E_2 are disturbance matrices with mean zero. It is further assumed that:

$$\text{plim } X'_{-1}E_1/T = \text{plim } X'_{-1}E_2/T = 0 \tag{5}$$

and

$$\text{plim } E_1'E_2/T = 0 \tag{6}$$

It follows that the disturbance u is subject to first-order serial correlation with parameter ρ_2.

In Equation (1), denote the matrix of right-hand side variables by $Z = [y_{-1}\ X]$. Then their variance-covariance matrix is assumed to converge to a fixed matrix:

$$\text{plim } Z'Z/T = G \tag{7}$$

where G is positive definite.[25] The two positive definite submatrices corresponding to the observed and unobserved exogenous variables are denoted by $G_1 = \text{plim } X_1{}'X_1/T$ and $G_2 = \text{plim } X_2{}'X_2/T$. The variance of the disturbance term u is $\sigma^2 = \beta_2{}'G_2\beta_2$.

Next, some preliminary results.

$$\begin{aligned}\text{plim } X_1{}'y/T &= \text{plim } [X_1{}'(\gamma y_{-1} + X_1\beta_1 + X_2\beta_2)/T] \\ &= \text{plim } [\gamma(\rho_1 X_{1,-1} + E)'y_{-1}/T] + G_1\beta_1 \\ &= \rho_1\gamma \text{ plim } X'_{1,-1}y_{-1}/T + G_1\beta_1 \end{aligned} \tag{8}$$

But since $\text{plim } X'_{1,-1}y_{-1}/T = \text{plim } X_1{}'y/T$ under the assumptions above, then

$$\text{plim } X_1{}'y/T = \frac{1}{1-\rho_1\gamma}G_1\beta_1 \tag{9}$$

By similar logic, after solving for y_{-1} from Equation (1):

$$\text{plim } X_1{}'y_{-1}/T = \frac{\rho_1}{1-\rho_1\gamma}G_1\beta_1 \tag{10}$$

and

$$\text{plim } X_2{}'y_{-1}/T = \frac{\rho_2}{1-\rho_2\gamma}G_2\beta_2 \tag{11}$$

Finally, using these results and the fact that under stationarity, plim $y_{-1}{}'y_{-1}/T = \text{plim } y'y/T = \text{var}(y) = \sigma_y^2$:

$$\text{plim } y_{-1}{}'y/T = \gamma\sigma_y^2 + \frac{\rho_1}{1-\rho_1\gamma}\beta_1{}'G_1\beta_1 + \frac{\rho_2}{1-\rho_2\gamma}\sigma^2 \tag{12}$$

and

$$\sigma_y^2 = \frac{1}{(1-\gamma^2)}\left[\frac{1+\rho_1\gamma}{(1-\rho_1\gamma)}\beta_1{}'G_1\beta_1 + \frac{1+\rho_2\gamma}{(1-\rho_2\gamma)}\sigma^2\right] \tag{13}$$

[25] In the univariate time–series case, it is sufficient to require that plim $X'X/T$ converge to a fixed positive definite matrix, since the convergence of plim $Z'Z/T$ then follows from stationarity. Indeed, under normality assumptions, stationarity alone would imply both. In the panel design, by contrast, the full assumption is needed even under stationarity and normality, since in that case the number of units goes to infinity, creating the need for conditions on that sequence.

Now suppose that ordinary least squares (OLS) is applied to Equation (1). By definition, the resulting estimates are:

$$\begin{bmatrix} \hat{\gamma} \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} y_{-1}'y_{-1}/T & y_{-1}'X_1/T \\ X_1'y_{-1}/T & X_1'X_1/T \end{bmatrix}^{-1} \begin{bmatrix} y_{-1}'y/T \\ X_1'y/T \end{bmatrix} \tag{14}$$

Applying the results above gives:

$$\begin{aligned} \text{plim} \begin{bmatrix} \hat{\gamma} \\ \hat{\beta}_1 \end{bmatrix} &= \begin{bmatrix} \sigma_y^2 & \rho_1\beta_1'G_1/(1-\rho_1\gamma) \\ \rho_1 G_1\beta_1/(1-\rho_1\gamma) & G_1 \end{bmatrix}^{-1} \\ &\quad \times \begin{bmatrix} \gamma\sigma_y^2 + \rho_1\beta_1'G_1\beta_1/(1-\rho_1\gamma) + \rho_2\sigma^2/(1-\rho_2\gamma) \\ G_1\beta_1/(1-\rho_1\gamma) \end{bmatrix} \end{aligned} \tag{15}$$

Using the standard formula for the inverse of a partitioned matrix and setting $s^2 = \sigma_y^2 - \rho_1^2\beta_1'G_1\beta_1/(1-\rho_1\gamma)^2$ yields:

$$\begin{aligned} \text{plim} \begin{bmatrix} \hat{\gamma} \\ \hat{\beta}_1 \end{bmatrix} &= (1/s^2) \begin{bmatrix} 1 & -\rho_1\beta_1'/(1-\rho_1\gamma) \\ -\rho_1\beta_1/(1-\rho_1\gamma) & s^2G_1^{-1} + \rho_1^2\beta_1\beta_1'/(1-\rho_1\gamma)^2 \end{bmatrix} \\ &\quad \times \begin{bmatrix} \gamma\sigma_y^2 + \rho_1\beta_1'G_1\beta_1/(1-\rho_1\gamma) + \rho_2\sigma^2/(1-\rho_2\gamma) \\ G_1\beta_1/(1-\rho_1\gamma) \end{bmatrix} \end{aligned} \tag{16}$$

After multiplying out and rearranging[26]:

$$\text{plim } \hat{\gamma} = \gamma + \left[\frac{\rho_2\sigma^2}{(1-\rho_2\gamma)s^2}\right] \tag{17}$$

Similarly, from Equation (16) and after some simplification:

$$\text{plim } \hat{\beta}_1 = \left[1 - \frac{\rho_1\rho_2\sigma^2}{(1-\rho_1\gamma)(1-\rho_2\gamma)s^2}\right]\beta_1 \tag{18}$$

Note that if the asymptotic bias in $\hat{\gamma}$ is defined as $g = \text{plim}(\hat{\gamma}) - \gamma$, the previous equation may be written as:

$$\text{plim } \hat{\beta}_1 = \left[1 - \frac{\rho_1 g}{(1-\rho_1\gamma)}\right]\beta_1 \tag{19}$$

[26] In the special case when $\gamma = 0$ (no included lagged dependent variable), $\beta_1'G_1\beta_1$ is the explained variance. Hence $\sigma_y^2 = \beta_1'G_1\beta_1 + \sigma^2$, and the associated squared correlation coefficient is $R^2 = \beta_1'G_1\beta_1/\sigma_y^2$. Thus in this special case, $s^2 = \sigma_y^2(1-\rho_1^2R^2)$, which in turn allows Equation (18) here to be written in the form of Equation (6) in the text. Similar remarks apply to Equation (17) here and the corresponding Equation (7) in the text.

These probability limits may now be used to prove the main result of this paper: In the case for which $0 \leq \gamma < 1$ and $0 < \rho_1, \rho_2 < 1$, then the coefficient on the lagged variable, $\hat{\gamma}$, is asymptotically biased upward toward unity, while the coefficient vector for the exogenous variables, $\hat{\beta}_1$, is asymptotically biased toward zero.[27] Both biases grow with increases in ρ_2, the degree of serial correlation.[28]

Formally, denote by $|\beta_1|$ the vector whose elements are the absolute values of the elements of β_1, and similarly for $|\text{plim } \hat{\beta}_1|$. Then:

Proposition. Under the assumptions listed above and when $0 \leq \gamma < 1$ and $0 < \rho_1,\ \rho_2 < 1$:

$$\gamma \leq \text{plim } \hat{\gamma} < 1 \tag{20}$$

$$0 < |\text{plim } \hat{\beta}_1| \leq |\beta_1| \tag{21}$$

$$\partial \text{plim}(\hat{\gamma})/\partial \rho_2 > 0 \tag{22}$$

and

$$\partial |\text{plim}(\hat{\beta}_1)|/\partial \rho_2 < 0 \tag{23}$$

Proof. The four assertions of the Proposition are taken up in order. First, in Equation (17), replace s^2 by its definition and replace σ_y^2 in turn from Equation (13). This gives, after some rearrangement:

$$\text{plim } \hat{\gamma} = \gamma + \rho_2 \left[\frac{(1-\gamma^2)(1-\rho_1\gamma)^2\sigma^2}{(1-\rho_1^2)(1-\rho_2\gamma)\beta_1{}'G_1\beta_1 + (1-\rho_1\gamma)^2(1+\rho_2\gamma)\sigma^2} \right] \tag{24}$$

The quantity in square brackets is clearly strictly positive, so that the lower bound for plim $\hat{\gamma}$ follows immediately by setting $\rho_1 = 0$.[29]

[27] The probability limits also show that in the absence of serial correlation, no coefficients are biased, as is well known. A bit less obviously, the probability limits demonstrate that in the absence of trending in the observed exogenous variables, serial correlation causes bias in the lag coefficient but not in the other regression coefficients.

[28] The relation of ρ_1, the autocorrelation parameter for the observed exogenous variables, to the biases in $\hat{\gamma}$ and $\hat{\beta}_1$ is slightly more complex, but both biases increase monotonically with ρ_1 when $\rho_1 > \gamma$, which is perhaps the most common case in applications.

[29] Note that additional exogenous variables would raise $\beta_1{}'G_1\beta_1$. (This is most easily

To show that the upper bound holds, note that, for fixed values of ρ_1, γ, and σ^2, maximizing expression (24) clearly requires $\beta_1{}'G_1\beta_1 = 0$, the case when the observed exogenous variables have zero coefficients and thus no explained variance. Making that substitution yields:

$$\begin{aligned} \max \text{ plim } \hat{\gamma} &= \gamma + \frac{\rho_2(1-\gamma^2)}{(1+\rho_2\gamma)} \\ &= \frac{\rho_2+\gamma}{1+\rho_2\gamma} \end{aligned} \tag{25}$$

It must be shown that the latter expression is less than unity. However, under the conditions of the Proposition:

$$(1-\rho_2)(1-\gamma) > 0 \tag{26}$$

Multiplying out and rearranging gives:

$$\frac{\rho_2+\gamma}{1+\rho_2\gamma} < 1 \tag{27}$$

which was to be proved.

For the second claim of the Proposition, the upper bound on $|\text{plim } \hat{\beta}_1|$ is obvious. (Set $\rho_1 = 0$ in Equation (19), noting from Equation (22) that $g = \text{plim}(\hat{\gamma}) - \gamma \geq 0$.) For the lower bound, it must be shown in Equation (19) that $\rho_1 g/(1-\rho_1\gamma) < 1$. But from the first part of the proof, $g < 1-\gamma$, so that:

$$\begin{aligned} \max \left[\frac{\rho_1 g}{1-\rho_1\gamma}\right] &< \frac{\rho_1(1-\gamma)}{1-\rho_1\gamma} \\ &= \frac{\rho_1-\rho_1\gamma}{1-\rho_1\gamma} < 1 \end{aligned} \tag{28}$$

Third, to show that plim $\hat{\gamma}$ is positively monotonic in ρ_2, take the partial derivative of Equation (24) with respect to ρ_2.[30] Only the second term on

seen when $\gamma = 0$, in which case $\beta_1{}'G_1\beta_1$ is the explained variance.) Hence Equation (24) implies that adding exogenous variables will reduce the asymptotic bias, all else equal, a point made by Malinvaud (1970, p. 559).

[30]Note that under the assumptions, σ_y^2 is a function of ρ_2, so that one cannot treat it as fixed and then work with the simpler Equation (17).

the right-hand side enters the derivative, and this term is just g. Setting $g = a/b$ implies that the derivative is $(ba' - ab')/b^2$, where the primes denote derivatives with respect to ρ_2. Hence we must show only that under the assumptions, the numerator of this expression is strictly positive for all values of the parameters.

Routine differentiation establishes that the numerator equals $(c + d)$, where

$$c = (1 - \gamma^2)(1 - \rho_1^2)(1 - \rho_1\gamma)^2\sigma^2\beta_1{}'G_1\beta_1 \tag{29}$$

and

$$d = (1 - \gamma^2)\ (1 - \rho_1\gamma)^4\sigma^4 \tag{30}$$

both of which are clearly strictly positive under the conditions of the Proposition. Hence:

$$\partial \text{plim}(\hat{\gamma})/\partial\rho_2 = \partial g/\partial\rho_2 > 0 \tag{31}$$

which establishes the (positive) monotonicity of plim $\hat{\gamma}$ with respect to ρ_2.

Fourth and finally, this result for the derivative of g may be applied to Equation (19), where the (negative) monotonicity of $|\text{plim } \hat{\beta}_1|$ with respect to ρ_2 follows after a simple application of the product rule for differentiation. This ends the proof.

Note that the bounds on the estimated coefficients are sharp. For example, Equations (19) and (22) imply that plim $\hat{\gamma} \to 1$ and plim $\hat{\beta}_1 \to 0$ as $\rho_1,\ \rho_2 \to 1$.

References

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

[2] Beck, Nathaniel, and Jonathan Katz. 1995. What To Do (and Not To Do) with Time–Series Cross-Section Data. *American Political Science Review* 89, 3 (September): 634-647.

[3] Beck, Nathaniel, and Jonathan Katz. 1996. Nuisance vs. Substance: Specifying and Estimating Time–Series–Cross–Section Models. *Political Analysis* 6: 1-36.

[4] Berry, William D. 1990. The Confusing Case of Budgetary Incrementalism: Too Many Meanings for a Single Concept. *Journal of Politics* 52, 1 (February): 167-96.

[5] Burkhart, Ross E., and Michael S. Lewis-Beck. 1994. Comparative Democracy: The Economic Development Thesis. *American Political Science Review* 88, 4 (December): 903-910.

[6] Catlin, G. E. G. 1927. *The Science and Method of Politics.* New York: Knopf.

[7] Choucri, Nazli, and Robert C. North. 1975. *Nations in Conflict.* San Francisco: W. H. Freeman.

[8] Davis, Otto A., M. A. H. Dempster, and Aaron Wildavsky. 1966a. A Theory of the Budgetary Process. *American Political Science Review* 60, 3 (September): 529-47.

[9] Davis, Otto A., M. A. H. Dempster, and Aaron Wildavsky. 1966b. On the Process of Budgeting. In Gordon Tullock, ed. *Papers on Non-Market*

Decision Making. Charlottsville, Virginia: Jefferson Center for Political Economy, University of Virginia. Pp. 63-132.

[10] Enders, Walter. 1995. *Applied Econometric Time Series.* New York: Wiley.

[11] Franzese, Robert. 2001. *The Political Economy of Macroeconomic Policy in Developed Democracies.* Cambridge: Cambridge University Press.

[12] Freeman, John R., John T. Williams, and Tse-min Lin. 1989. Vector Autoregression and the Study of Politics. *American Journal of Political Science* 33: 842-877.

[13] Gallagher, Michael, Michael Laver, and Peter Mair. 1992. *Representative Government in Western Europe.* New York: McGraw Hill.

[14] Gleditsch, Nils Petter. 1990. Research on Arms Races. In Nils Petter Gleditsch and Olav Njolstad. *Arms Races.* London: Sage: 1-14.

[15] Goldstein, Joshua S., and John R. Freeman. 1990. *Three-Way Street: Strategic Reciprocity in World Politics.* Chicago: University of Chicago Press.

[16] Granger, C. W. J., and Paul Newbold. 1986. *Forecasting Economic Time Series*, 2nd ed. Orlando, Florida: Academic Press.

[17] Greene, William H. 1993. *Econometric Analysis,* 2nd ed. Englewood Cliffs, New Jersey: Prentice-Hall.

[18] Griliches, Zvi. 1961. A Note on Serial Correlation Bias in Estimates of Distributed Lags. *Econometrica* 29, 1 (January): 65-73.

[19] Griliches, Zvi. 1967. Distributed Lags: A Survey. *Econometrica* 35, 1 (January): 16-49.

[20] Gujarati, Damodar N. 1995. *Basic Econometrics,* 3rd ed. New York: McGraw-Hill.

[21] Hamilton, James D. 1994. *Time Series Analysis.* Princeton: Princeton University Press.

[22] Harvey, A. C. 1981. *The Econometric Analysis of Time Series.* Oxford: Philip Allan. (See also 2nd ed., Cambridge: MIT Press, 1993.)

[23] Hibbs, Douglas A., Jr. 1974. Problems of Statistical Estimation and Causal Inference in Time–Series Regression Models. In Herbert L. Costner, ed. *Sociological Methodology 1973-1974.* San Francisco: Jossey-Bass. Pp. 252-308.

[24] Hicks, Alexander, and Duane Swank. 1992. Politics, Institutions, and Welfare Spending in Industrial Democracies, 1960-1982. *American Political Science Review* 86: 658-74.

[25] Kiewiet, D. Roderick, and Matthew D. McCubbins. 1985. Congressional Appropriations and the Electoral Connection. *Journal of Politics* 47, 1 (February): 59-82.

[26] Lindblom, Charles E. 1959. The Science of "Muddling Through." *Public Administration Review* 19 (Spring): 79-88.

[27] Lütkepol, Helmut. 1982. The Impact of Omitted Variables on the Structure of Multiple Time Series: Quenouille's Data Revisited. In O.D. Anderson, ed., *Time Series Analysis: Theory and Practice 2.* Amsterdam: North Holland. Pp. 143-59.

[28] Maddala, G. S. 1992. *Introduction to Econometrics*, 2nd ed. Englewood Cliffs, New Jersey: Prentice Hall.

[29] Maddala, G. S. 1999. Recent Developments in Dynamic Econometric Modelling: A Personal Viewpoint. *Political Analysis* 7 (1998): 59-87.

[30] Maddala, G. S., and A. S. Rao. 1973. Tests for Serial Correlation in Regression Models with Lagged Dependent Variables and Serially Correlated Errors. *Econometrica* 47, 4 (July): 761-774.

[31] Malinvaud, E. *Statistical Methods of Econometrics,* 2nd revised edition. 1970. Amsterdam: North-Holland.

[32] McCubbins, Mathew D., and Thomas Schwartz. 1984. Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science* 28, 1 (February): 165-179.

[33] Moll, Kendall D., and Gregory M. Luebbert. 1980. Arms Race and Expenditure Models. *Journal of Conflict Resolution* 24, 1 ((March): 153-185.

[34] Newbold, Paul, and C. W. J. Granger. 1974. Experience with Forecasting Univariate Time Series and the Combination of Forecasts. *Journal of the Royal Statistical Society A* 137: 131-46.

[35] Oren, Ido. 1996. Capability and Hostile Behavior in Arms Race Models. *International Interactions* 21, 4: 309-334.

[36] Organski, A. F. K., and Jacek Kugler. 1980. *The War Ledger.* Chicago: University of Chicago Press.

[37] Phillips, P. C. B., and M. R. Wickens. 1978. *Exercises in Econometrics*, vol. 2. Oxford: Philip Allan.

[38] Rao, Potluri, and Roger LeRoy Miller. 1971. *Applied Econometrics.* Belmont, California: Wadsworth.

[39] Rattinger, Hans. 1975. Armaments, Detente, and Bureaucracy. *Journal of Conflict Resolution* 19, 4 (December): 571-595.

[40] Rattinger, Hans. 1976. Econometrics and Arms Races: A Critical Review and Some Extensions. *European Journal of Political Research* 4: 421-439.

[41] Richardson, Lewis F. 1960. *Arms and Insecurity.* Pittsburgh: Boxwood Press.

[42] Russett, Bruce. 1983. International Interactions and Processes: The Internal vs. External Debate Revisited. In Ada Finifter, ed., *Political Science: The State of the Discipline.* Washington: The American Political Science Association. Pp. 541-568.

[43] Sharkansky, Ira. 1968. *Spending in the American States.* Chicago: Rand McNally.

[44] Simon, Herbert A. 1955. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics* 69: 99-118.

[45] Stimson, James A. Regression in Space and Time: A Statistical Essay. *American Journal of Political Science* 29, 4 (November): 914-947.

[46] Wiberg, Hakan. 1990. Arms Races, Formal Models, and Quantitative Tests. In Nils Petter Gleditsch and Olav Njolstad. *Arms Races.* London: Sage. Pp. 31-57.

[47] Wildavsky, Aaron. 1964. *The Politics of the Budgetary Process.* Boston: Little, Brown.

[48] Wood, B. Dan. 1992. Modeling Federal Implementation as a System: The Clean Air Case. *American Journal of Political Science* 36, 1 (February): 40-67.

[49] Zeigler, Harmon, and Karl F. Johnson. 1972. *The Politics of Education in the States.* Indianapolis: Bobbs-Merrill.