Let the Researcher Beware: The Use of the Richardson Equations to Estimate the Parameters of a Dyadic Arms Acquisition Process

Author(s): Richard J. Stoll


Published by: Midwest Political Science Association


Accessed: 25/10/2013 17:08
Let the Researcher Beware: The Use of the Richardson Equations to Estimate the Parameters of a Dyadic Arms Acquisition Process*

Richard J. Stoll, Rice University

Most attempts to untangle the factors accounting for arms acquisition by nation states rely implicitly or explicitly on the work of Lewis Richardson. However, very few findings from these investigations are durable; one study finds a particular nation state dyad to be involved in an arms race, while the next concludes that this same pair of nations is driven by factors internal to each. The question asked in this paper is: how likely is the true model of a dyadic arms acquisition process to emerge when the Richardson equations are used? This question is investigated by constructing two data sets representing the military capability of a pair of nations—one when an arms race is present, and the other when each nation’s acquisition of capability is driven by internal factors. These data are then used to estimate the Richardson equations. The conclusion is that in many situations, estimation of the Richardson equations may not uncover the true parameters of the process and may well lead to different conclusions about the same dyadic arms acquisition process.

Lewis Richardson’s “linear theory of two nations” was one of the first (and certainly the most well known) formal models in international relations. Briefly, Richardson (1960; original 1919) proposed to model the “defenses” of a pair of nations by means of the two equations displayed below. Richardson’s original form specified the relationships as differential equations; here they are restated in the more usual discrete form:

\[
\Delta X_i = b_{xy} Y_{i-1} + b_{x} X_{i-1} + c \\
\Delta Y_i = b_{yx} Y_{i-1} + b_{y} X_{i-1} + d
\]

where:
- \(X_i, Y_i\) = defense expenditures of nations \(X\) and \(Y\) at time \(i\).
- \(\Delta X_i, \Delta Y_i\) = change in defense expenditures of nations \(X\) and \(Y\) from time \(i-1\) to time \(i\).
- \(b_{xy}, b_{yx}\) = “defense coefficients” indicating the effects of one nation’s previous defense expenditure level on current changes in the other nation’s defense expenditure.

* My thanks to Bruce Bueno de Mesquita, who originally raised the issues discussed in this paper, and to David John Gow for a careful reading of the manuscript. Responsibility for any errors is mine.

American Journal of Political Science Vol. 26, No. 1, February 1982
©1982 by the University of Texas Press 0092-5853/82/010077-13$01.40
\[ b_{i1}, b_{i2} = \text{"fatigue coefficients" indicating the effect of each nation's previous level of defense expenditure on its own current change in defense expenditure.} \]

\[ c, d = \text{"grievance coefficients" indicating the net hostility (or friendship) felt by each nation toward the other, independent of all levels of expenditure.} \]

This formulation has had a powerful effect on the study of arms acquisition processes. There are several reasons for this. First, Richardson's equations appear compelling as a general description of the possible influences on arms acquisition. Further, the indicators suggested by Richardson's work, arms expenditures, are generally available to researchers and usually as "hard figures" (currency units). Finally, the form of the equations lends itself quite readily to parameter estimation by a variety of econometric techniques; these methods are the staples of quantitative empirical work in the social sciences.

However, despite numerous studies using Richardson's equations in the "pure" form shown above or some more exotic variant, despite more extensive data bases, and despite the use of increasingly sophisticated methodology, little progress has been made in identifying either the causal mechanisms of a particular arms race, or classes of arms races. A recent review (Moll and Luebbert, 1980, p. 166) of the arms race literature concludes: "The common failure of statistical tests to discriminate among various models indicates that there are fundamental problems with either the explanatory power of the models or the discriminating ability of the tests. Which is the real difficulty remains a subject for future research."

As an example of the problems scholars have had in untangling the underlying processes of dyadic arms acquisition processes, consider two recent studies of the U.S.-Soviet dyad in the post--World War II era. On the one hand, Majeski and Jones (1980, p. 19) conclude that "no causal relationship was found between [U.S. and Soviet] military expenditures." They attribute the apparent "interaction" between the superpowers to internal factors operating independently within each nation. On the other hand, Wallace (1980, p. 271) is equally emphatic that his findings "offer no support whatever to those who are convinced that the arms race results from the ineluctable nature of the political or economic system on

---

1 This paper will not deal, in more than a superficial manner, with the use of military expenditure figures as measures of defense effort or capability. Suffice it to say that these data are more elastic than is desirable. For example, with regard to the U.S.-Soviet comparisons, note the wide variance between the work of Lee (1977), whose method of estimating Soviet expenditures is accepted as better than most, and Holzman (1980), who demonstrates that by making a number of not unreasonable assumptions, very different figures result.
either side. Rather, it would appear that both superpowers are reacting to external [emphasis in original] rather than internal stimuli.”

There are numerous differences between these studies, but nevertheless, the fact that opposite conclusions were reached with strong empirical evidence surely gives one pause. How is this possible? I believe a great deal of the cross-study variance in findings can be laid at the door of the basic conception of the arms acquisition process handed down from Richardson’s pioneering research.

The Internal-Internal Case

To illustrate these problems I will consider two hypothetical dyadic arms acquisition processes. Using simple algebra, correlations, multiple regression, and Monte Carlo experiments, I will investigate whether the Richardson equations can be expected to uncover these underlying processes.

First, consider a dyad in which the arms acquisition of both nations is a result of internal forces, such as bureaucratic momentum. That is,

2 The statistical technique used by Majeski and Jones is statistical causality analysis; Wallace used seemingly unrelated regressions. For predictor variables, Majeski and Jones use the military expenditures of both superpowers; Wallace uses expenditures, serious disputes, and, for the United States, a war mobilization index. The data used by Majeski and Jones come from Stockholm International Peace Research Institute (SIPRI), while Wallace’s data are from the Correlates of War Project. Finally, the years of Majeski and Jones’ study are 1946–1976. Wallace does not indicate what years are used in his data base.

3 A number of authors have noted and/or tried to resolve the difficulties that have appeared in the empirical estimation of the Richardson equations. In fact, the two contradictory findings noted above (Majeski and Jones, 1980; Wallace, 1980) are attempts to do just that. Wagner, Perkins, and Taagepera (1975), in an investigation of the pre-World War I arms race, plot curves for both internal-internal and external-external models; there is little difference between them. Lutcher (1975) and Rattinger (1976) comment on the difficulties of empirical estimation of arms race parameters; both advocate the use of more complicated estimation techniques to solve the problem. Another demonstration of the difficulties in evaluating the drives behind an arms acquisition process is given by Ostrom (1977). He compares the fit of three models—an external, an internal, and a null—in forecasting U.S. military expenditures in the post-World War II era. He finds that the fits are so similar that a choice among them is impossible; further, he notes a “similarity in the dynamic structure” between the internal and external models (Ostrom, 1977, p. 260). The published work that most closely follows this effort is that of Schrodt (1978). He uses Monte Carlo techniques to examine the effects of random error on the $R^2$ and coefficient values when estimating an internally driven model. However, the outcome variable in his study is the level of expenditures. Further, his experiments, concentrating on the effects of applying alternative estimation algorithms, do not consider what happens when both internal and external variables are entered into the equation, with various quantities of random error. Thus, most previous work fails to deal directly with the question addressed here: does the fault lie with the conceptualization of the arms acquisition process, rather than with any particular data set or estimation technique?
Richard J. Stoll

\[ \Delta X_i = b_{xx} X_{i-1} \quad \text{(Eq. 1)} \]
\[ \Delta Y_i = b_{yy} Y_{i-1} \quad \text{(Eq. 2)} \]

If the Richardson equations are used, then we will be estimating each nation’s change in expenditure (i.e., \( \Delta X \) and \( \Delta Y \)) from both nations’ previous levels of expenditure, as illustrated in the first display. Will we be able to uncover the true relationship?

A simple internally driven process (i.e., one with a constant rate of growth) is a compound interest problem. In other words, we can represent the level of arms acquisition for nation \( X \) at time \( t \) as follows:

\[ X_t = (1 + g_t)X_0 \quad \text{(Eq. 3)} \]

In Equation 3, \( g_t \) is the growth rate of \( X \)'s expenditure (i.e., .05 if expenditures are growing at 5 percent per annum), and \( X_0 \) is the level of expenditure of \( X \) when the “arms race” begins.

Through substitution, the change in \( X \)'s expenditures can be expressed as

\[ \Delta X_t = X_0(1 + g_t)^{t-1}g_t \quad \text{(Eq. 4)} \]

The change in expenditures for nation \( Y \) can be expressed similarly. We can find constants \( k \) and \( m \), such that:

\[ X_0 = kY_0 \quad \text{and} \quad g_t = m + g, \]

Substituting these values into the expression for \( \Delta X_t \) (Eq. 4) yields:

\[ \Delta X_t = kY_0 (1 + m + g)^{t-1} (m + g) \quad \text{(Eq. 5)} \]

Thus we can derive an expression for the change in nation \( X \)'s expenditure in terms of constants, the initial value for \( Y \)'s expenditures, and \( Y \)'s growth rate, even though the true model for change relies only on \( X \)'s own previous level of expenditures.

All this is fairly simple mathematics; the reader may feel it merely illustrates what Hubert Blalock (1961, ch. 2) said long ago—that algebra is sometimes too flexible for model building in the social sciences. But it does merit some thought. Suppose a pair of nations has only a small difference in their arms acquisition growth rates. We should expect that the two (causally independent) series of expenditure levels will look very similar, and therefore, look to be highly intercorrelated.

This offers the potential for much confusion if we try to estimate the parameters of the Richardson equations. Multicollinearity will exist between the two nations’ levels of expenditures, and it will not be a “data problem” in the sense of a characteristic of a particular sample of points but rather an inevitable consequence of the underlying processes used to generate the data. Just how high the intercorrelation can be is shown in Table 1. Series
Use of Richardson Equations

Table 1
Correlation between Two Hypothetical Expenditure Series with Constant Growth Rates
(N = 30)

<table>
<thead>
<tr>
<th>Growth Rate, Y's Expenditures</th>
<th>Growth Rate, X's Expenditures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>.989</td>
</tr>
<tr>
<td>5%</td>
<td>.953</td>
</tr>
<tr>
<td>10%</td>
<td>.907</td>
</tr>
<tr>
<td>15%</td>
<td>.859</td>
</tr>
<tr>
<td>20%</td>
<td></td>
</tr>
</tbody>
</table>

Note: Average r = .953.

of 30 observations were generated (30 is a generous upper figure for the number of years in an arms race study), with a variety of growth rates (1, 5, 10, 15, and 20 percent per annum). Even a great disparity in growth rates (1 percent versus 20 percent) produces an excessive and dangerous degree of multicollinearity between the two "independent" variables in the equation.

How many of us would feel comfortable making inferences about the significance (statistical and otherwise) of the two Richardson equation coefficients given this level of multicollinearity between the independent variables in the equation? Even a model that predicts changes in expenditure for one nation, from the previous level of expenditure of the other nation (i.e., ΔX, from Y, or ΔY, from X), will produce a good (perhaps outstanding) fit.

If there is no fluctuation from the true relationship between current changes in expenditure of a nation and its own previous level (i.e., "perfect data"), the parameters estimated from the Richardson equations will be correct; the internal coefficient (b_y or b_x) will reflect the growth rate, and the external coefficient (b_y or b_x) will be zero.

What happens when the data contain some fluctuations, and the parameters of the Richardson equations are estimated? Tables 2a and 2b display a summary of the results from 20 Monte Carlo runs designed to answer this question.

Each nation's level of expenditure was given error rates of ±3, ±5, and ±10 percent of the mean value in its error-free series, and the Richardson equations for all combinations of these error rates were estimated. The exact procedure is as follows: first, add the (possibly negative) normally

4 Since the correlation expresses the linear relationship between the two series in standardized form, the starting point for each series (X_0 and Y_0) has no effect on its value.
distributed random component to each of the predictor variables (i.e., the level of expenditures of X and Y); second, compute the change in X's expenditure from the previous year; third, use both X's and Y's level of expenditure variables to predict to X's change in expenditures.5

The Monte Carlo results allow for two different tests of the Richardson equations. First, we can check to see how much bias occurs by testing if the estimates are significantly different from the true (i.e., error-free) model.6 Second, we can check to see if the coefficient estimates are significantly different from the standard error estimates. This aids us in determining what a researcher would conclude about the arms acquisition process given these results.

A t-statistic is used for both tests. For the first test, both regression coefficients and the standard error of the internal coefficient (i.e., the standard error of \( b_{xy} \)) are examined. The observed average values of these estimates are compared to their "true" values, and t-tests are run to determine if these three parameter estimates are significantly different. The true values assumed for each estimate are indicated in Table 2a. The true values for the regression coefficients are simply the values that would occur in the errorless data. A two-tailed t-test is used, since I am interested in whether the Monte Carlo estimates diverge in either direction from the true values. The value for the standard error of the internal coefficient is taken to be one-half of the true value of its associated coefficient. If the estimated value for the standard error was this size or smaller, and the regression coefficient was correctly estimated, then we would be likely to conclude that the coefficient is significantly different from 0. This line of reasoning leads to a one-tail t-test for the standard error: the key question is whether the estimated standard error is too large. If it is smaller than the test value, so much the better. Since the assumed true value for the external regression coefficient (i.e., \( b_{xy} \)) is 0, there is no convenient test value for its associated standard error; no t-test is conducted for this estimate.

The other test of the Monte Carlo results is conducted by computing the t-ratios for all of the regression coefficients. The mean value of each

5 The same analysis was done with Y's change in expenditures as the dependent variable. The results were the same as those described in the text with X's change as the dependent variable. Tables with results from all Monte Carlo runs are available from the author.

6 Introducing random error in the series is, admittedly, an arbitrary device to break down the relationship of the true model. It also produces some attenuation bias in the coefficients, since error is being introduced to the independent variables (Duncan, 1975, p. 117). However, it is the simplest way to observe the effects of estimation with an imperfect model. Further, the question that interests me the most is what happens to the coefficients of the variables that should have no effect on the change in expenditures. The introduction of random error allows us to assess these effects.
coefficient is divided by the mean value of its standard error. If the resulting ratio is 2 or greater (the usual rule of thumb), the coefficient is interpreted as being significant. That is, if the researcher obtained such a t-ratio, then s/he would be likely to conclude that the variable had a significant effect in predicting the rate of change of the dependent variable. Summaries of these two tests are displayed in Tables 2a and 2b.

Turning first to Table 2a, we see that most of the Monte Carlo estimates are significantly different from the true values (19 of the 27 estimates, including 11 of the 18 coefficient estimates). From inspection of the results, it appears that only when the error in Y is 10 percent are the coefficients for \( b_{xx} \) close to the true values. And when the error in X is 10 percent, the \( b_{xx} \) coefficients are negative.

The regression coefficient t-ratios displayed in Table 2b tell an equally depressing story. Only three are equal to or greater than 2.00. Of these three, two are for the \( b_{xy} \) coefficient, which should be 0. The single significant \( b_{xx} \) coefficient is negatively signed.

**TABLE 2a**

Monte Carlo Results, Internal Model

<table>
<thead>
<tr>
<th>Estimate</th>
<th>True Value</th>
<th>Number of Estimates Significantly Different from True Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b_{xx} )</td>
<td>.10</td>
<td>5</td>
</tr>
<tr>
<td>( SE(b_{xx}) )</td>
<td>( \leq .05 )</td>
<td>8</td>
</tr>
<tr>
<td>( b_{xy} )</td>
<td>0.00</td>
<td>6</td>
</tr>
</tbody>
</table>

*Note: Every estimate is an average from 20 runs of each combination of \( \pm 3\% \), \( \pm 5\% \), and \( \pm 10\% \) error in X and Y.*

**TABLE 2b**

Monte Carlo Results, Internal Model

t-ratios of 9 Estimated Coefficients over Estimated Standard Errors

<table>
<thead>
<tr>
<th>Coefficient Estimate</th>
<th>Number of t-ratios ( &gt; 2.00 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b_{xx} )</td>
<td>1</td>
</tr>
<tr>
<td>( b_{xy} )</td>
<td>2</td>
</tr>
</tbody>
</table>

*Note: Every estimate is an average from 20 runs of each combination of \( \pm 3\% \), \( \pm 5\% \), and \( \pm 10\% \) error in X and Y.*
Overall, the conclusions I draw from this investigation are that, with a simple model and moderate amounts of error in the data, an internally driven dyadic arms acquisition process will not be successfully uncovered by the estimation of the Richardson equations. Estimated coefficients are likely to diverge greatly from their true values, and on those occasions when a coefficient is significant, it may well contain an incorrect sign.

The External-External Case

I now turn to a situation in which each nation's change in military expenditures is driven by the previous level of the other nation. This is the pure arms race case. Once again, I begin with some algebraic manipulation.

In this situation the equation for X's expenditures is

\[ \Delta X_i = b_{xy} Y_{i-1} \]  

(Eq. 6)

Noting that: (a) the level of Y's expenditure at any point in time is the sum of its pre-arms-race spending, \( Y_0 \), and all the successive increments to this base; and (b) each year's increment to Y's expenditure is itself determined by last year's level of X's expenditure, we can express the change in X's expenditure as follows:

\[ \Delta X_{m+1} = b_{xy} \sum_{j=0}^{m-1} X_j + b_{yx} Y_0 \]  

(Eq. 7)

| TABLE 3 |
Correlation between Two Hypothetical Series with Ratio Goals  
(\( N = 30 \))

<table>
<thead>
<tr>
<th>X's Ratio Goal</th>
<th>( Y_i = 1.01 X_{i-1} )</th>
<th>( Y_i = 1.05 X_{i-1} )</th>
<th>( Y_i = 1.10 X_{i-1} )</th>
<th>( Y_i = 1.15 X_{i-1} )</th>
<th>( Y_i = 1.20 X_{i-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y_i = 1.01 X_{i-1} )</td>
<td>-.690</td>
<td>.246</td>
<td>.677</td>
<td>.817</td>
<td>.878</td>
</tr>
<tr>
<td>( Y_i = 1.05 X_{i-1} )</td>
<td>.156</td>
<td>.573</td>
<td>.765</td>
<td>.844</td>
<td>.886</td>
</tr>
<tr>
<td>( Y_i = 1.10 X_{i-1} )</td>
<td>.549</td>
<td>.714</td>
<td>.810</td>
<td>.859</td>
<td>.889</td>
</tr>
<tr>
<td>( Y_i = 1.15 X_{i-1} )</td>
<td>.681</td>
<td>.768</td>
<td>.828</td>
<td>.864</td>
<td>.888</td>
</tr>
<tr>
<td>( Y_i = 1.20 X_{i-1} )</td>
<td>.737</td>
<td>.792</td>
<td>.836</td>
<td>.864</td>
<td>.885</td>
</tr>
</tbody>
</table>

Note: Starting value for X's expenditures was 150. Starting value for Y's expenditures was 100. Average \( |r| = .740 \).
Setting $kX_0 = Y_0$, the algebra of the equations can be used to show that in the pure arms race situation, the change in a nation's expenditures can be related to the lagged sum of its own previous expenditures.

Repeating the investigations of the internal-internal situation, I pick a simple model to explore. I assume that $X$ spends so as to have a percentage advantage over $Y$'s previous year's expenditures; i.e., if $X$ wants a 15 percent advantage over $Y$, then $X = 1.15Y_{-1}$. Table 3 displays the intercorrelations between various pairs of such percentage goals.

Once again, a series of 30 observations is used. Although two of the correlations are low (when either nation seeks to have 105 percent of the other's previous level), the pattern is of moderate to high relationships between the two series. As with the internal-internal case, using both series as predictor variables in estimation of a Richardson equation will entail a high degree of multicollinearity.

### TABLE 4a

Monte Carlo Results, External Model

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Estimate</th>
<th>True Value</th>
<th>Number of Estimates Significantly Different from True Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{xx}$</td>
<td>-1.0</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>$SE (b_{xx})$</td>
<td>$\leq .5$</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>$b_{xy}$</td>
<td>1.1</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>$SE (b_{xy})$</td>
<td>$\leq .55$</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

*Note: Every estimate is an average from 20 runs of each combination of $\pm 3\%, \pm 5\%$, and $\pm 10\%$ error in $X$ and $Y$."

### TABLE 4b

Monte Carlo Results, External Model

$t$-ratios of 9 Estimated Coefficients over Estimated Standard Errors

<table>
<thead>
<tr>
<th>Coefficient Estimate</th>
<th>Number of $t$-ratios $&gt; 2.00$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{xx}$</td>
<td>9</td>
</tr>
<tr>
<td>$b_{xy}$</td>
<td>9</td>
</tr>
</tbody>
</table>

*Note: Every estimate is an average from 20 runs of each combination of $\pm 3\%, \pm 5\%$, and $\pm 10\%$ error in $X$ and $Y$."

This content downloaded from 128.42.228.89 on Fri, 25 Oct 2013 17:08:10 PM
All use subject to JSTOR Terms and Conditions
If the Richardson equations with perfect data are estimated, the true model emerges. But the coefficient for the internal past level (i.e., $b_{xx}$) will be $-1.0$. This occurs because the true model is $X_t = 1.1Y_{t-1}$. But we are estimating the change in expenditures of $X$. This is expressed by subtracting $X_{t-1}$ from the left side of the equation representing the true model. However, this same quantity must be subtracted from the right side of the equation to balance out. Thus, if the arms race relationship involves a simple percentage goal, we should expect to find that the internal coefficient has a negative, not a zero value, although it has no causal impact. I will return to this point later.

The Monte Carlo results for the external-external dyad are shown in Tables 4a and 4b. The format of presentation and the type of analysis are the same as in Tables 2a and 2b. The only difference is that, with the external-external model, the true parameters have nonzero values for both regression coefficients, so that Table 4 contains $t$-tests for both coefficients and their standard errors (once again, I assume that the true value for the standard error should be one-half the size of the regression coefficient).

As can be seen, the results of these Monte Carlo runs are a lot more satisfactory than for the internal-internal case. Most of the Monte Carlo estimates are not significantly different from the true values (only 5 of 36). Furthermore, as can be seen from Table 4b, every estimated coefficient is much larger than its associated standard error.

Although the findings of the Monte Carlo study indicate that parameter estimates for an externally driven arms acquisition process will be accurate, a problem still remains. What are we to conclude if the internal coefficient has a statistically significant negative value? Given the model used for the Monte Carlo runs, the correct interpretation is that the coefficient value is a mathematical artifact of the expression of the relationship in terms of change over time. However, there are other theoretical frameworks that would lead to the expectation of a negative internal coefficient. Richardson himself believed that large previous levels of military expenditure by a nation would lead to a fatigue effect on current changes of expenditure, as the economic and psychic costs associated with military preparations made themselves felt. The existence of two different theoretical rationales leading to the same parameter prediction is, to say the least, unsatisfactory.

**Implications, Suggestions, and Conclusions**

This paper has examined the vulnerability of the Richardson equations to misleading inferences in the study of dyadic arms acquisition processes. The results are very sobering. The probability of reaching a misleading or ambiguous conclusion appears to be high. If both nations’ arms acquisition processes are driven by internal forces, empirical estimation is unlikely to
produce accurate results. When a true arms race is present, the results will be unbiased, but could be accounted for by other theoretical processes.

To be quite fair, this paper has not proven that these conclusions are valid for all possible internal and external arms acquisition processes. Two simple, "pure" cases were selected for examination. However, if there are problems with these processes, should we assume that they will be diminished or disappear if more complicated processes drive the arms acquisition of nations? I think not.\(^7\)

At the very least, this paper has shown that high multicollinearity in arms race research may be endemic to the phenomena being studied and not be a problem with a particular data series. Considering the debilitating effects of this problem on inferences about the relative importance of variables, and the large amount of effort that has gone into partitioning variance between internal and external factors, this is a serious problem.\(^8\)

In most empirical applications, the researcher is on the horns of a very nasty dilemma. Estimating a model that contains only internal or only external components is likely to give a reasonable fit, even if the true process is driven totally by the excluded component. Estimation of a model with both internal and external components and attempting to partition the variance between the two is likely to be difficult because of the multicollinearity problem discussed above.\(^9\)

\(^7\) I have assumed that our investigations usually contain strong exploratory (as opposed to confirmatory) elements. My reading of the arms race literature leads me to conclude that this is correct. Very rarely, for example, do researchers hypothesize the magnitude, or in many cases, even the direction of a coefficient. Most commonly, the question asked is whether the variables in the equations exercise any effect (i.e., whether they are nonzero).

\(^8\) In many studies the intercorrelations between predictor variables are not reported. I would hope that this paper would lead to a heightened skepticism of the validity of inferences about arms acquisition processes when this information is left out of a research report.

\(^9\) The researcher may have a priori information that would allow him or her to order the impacts of internal and external factors on changes in expenditures. For example, Ostrom (1978) assumes that Soviet expenditures affect budget requests by the U.S. military, and the service requests then pass through the budgetary cycle—clearly, the external factor is placed first in his scheme. However, given the multicollinearity problems outlined above, one should always be aware that reversing the temporal sequences is also likely to result in a good fit. The two contradictory findings noted in the beginning of this paper provide an illustration of this. Majeski and Jones use Auto-Regressive-Integrated-Moving-Average (ARIMA) techniques separately on U.S. and Soviet expenditures and correlate the residuals from this analysis. Since the correlations are low, they conclude that external factors play little part in determining superpower arms levels. However, by placing the internal factors first, all the shared variance between internal and external factors is allocated to previous expenditure levels within both nations. In a test, I took my "true" arms race (i.e., the data from my pure external-external process) and used ARIMA on both series. Not surprisingly, the fit of the ARIMA models was very good and the residuals very small. Wallace (1980) devotes some attention to the presence of multicollinearity in his data and
As a general observation, I believe the study of arms acquisition processes would profit by moving from general tests of general models to specific tests of specific models.

Perhaps the best alternative is to accept Richardson not as a specific model of arms acquisition but as a general conception (more akin to a paradigm) and to craft a more explicit representation of the particular arms acquisition process under investigation. Such a representation may be built on information from historical or contemporary accounts that reveal (or purport to reveal) values or ranges of values for various parameters. Despite difficulties, testing a more explicit set of parameter values will result in a more satisfactory model of an arms acquisition process, although at the cost of a lesser degree of generality.10

If there is one lesson to be learned from this paper, it is that a loosely expressed series of ideas will cause great difficulty when an attempt is made at rigorous estimation. Even if the relationship between variables is a simple one, the results may fail to indicate the true parameters of a process. The wide variance in findings about various dyadic arms acquisition processes serves as a clear example of this statement. All arms race researchers (indeed, all international relations researchers) owe a great debt to Lewis Richardson and his pioneering efforts, but if we are to produce durable findings about arms acquisition processes, we must make our models and our tests stronger and more explicit.

Manuscript submitted 6 November 1980
Final manuscript received 27 March 1981

its deleterious effects on resolving the internal-external debate. However, his solution consists of leaving out of his equations the internal expenditure component. Thus, the shared variance between internal and external expenditure levels is allocated entirely to the external variable. The fit of his models for the Soviet Union and the United States are very good. I took the “true” internally driven arms acquisition that I used earlier and predicted the change in expenditure of each nation from the previous level of expenditure of the other nation, using the same econometric technique as Wallace (seemingly unrelated regressions). As anticipated, the fits were quite good. Thus, in both cases, the incorrect model will fit the data very well, due to the large amount of shared variance between the internal and external factors.

10 Along with resorting to a more limited model, researchers might also be advised to move away from total military (or even total service) expenditures as an indicator of capability. After all, wars are fought by soldiers with weapons, not by accountants with dollars. Although, in principle, valuing capability by currency appears to produce a single recognizable standard of comparison, specific program decisions may represent a more accurate tapping of the capability that drives further arms acquisition by both sides. It goes without saying that constructing these alternative measures of military capability will be a long, arduous task fraught with difficulties.
REFERENCES