Measuring Military Allocations

A COMPARISON OF DIFFERENT APPROACHES

GARY GOERTZ
Department of Political Science
University of Michigan

PAUL F. DIEHL
Department of Political Science
University of Georgia

This study investigates the validity of different indicators of resources allocated to the military that are commonly used in conflict research. The various indicators are first described and then evaluated on four criteria for face validity. The convergent, discriminant, and predictive validity are then assessed according to the results of various statistical tests. A common hypothesis that conflict involving "overallocating" states is more likely to escalate to war is examined using all indicators under consideration. Implications and suggestions for conflict research are discussed in the conclusion.

In principle (and surely in propaganda), military establishments exist in order to protect a nation-state from its enemies. All nations, with the possible exception of Costa Rica and Iceland, have some level of military preparedness. In order to maintain that preparedness, economic resources, be they human or capital, must be allocated to the military. Accompanying this allocation to the military are certain opportunity costs and possibly harmful economic side-effects. Our goal is not to make a contribution to the existing literature on the domestic consequences of military spending, but rather to analyze and evaluate various operational measures of that allocation.

AUTHORS' NOTE: An earlier version of this article was presented at the Annual Meeting of the International Studies Association, Washington, D.C., March 5-9, 1985. We would like to thank Robert Axelrod, J. David Singer, and two anonymous reviewers for their suggestions.

JOURNAL OF CONFLICT RESOLUTION, Vol. 30 No. 3, September 1986 553-581

Measurement approaches in the social sciences are inherently theoretically based and vary according to the purposes and preconceptions of the individual researcher. Rather than assess all possible measurement approaches for military allocations, we chose to concentrate on those that relate to the study of conflict. By making such a choice, we bypass those indicators of military allocations designed to assess economic costs (that is, those that consider domestic economic effects; for example, see Russett, 1982). We focus instead on indicators that have actually been used to identify the magnitude of military allocations and to study their impact on conflict escalation (for example, military expenditures divided by GNP).

This is not simply a methodological exercise. The implications for the study of conflict are potentially great. Military allocations are thought to play a prominent role in national decisions for war. First, high military allocations may be an early warning indicator of conflict escalation. Nations may be reluctant to fight unless they have adequate military preparations. High military allocations are indicative of this preparedness and perhaps also of the willingness to use military force.

Second, military allocations might also provide a link in the causal chain for war. High allocations could indicate the influence that military officials have in government decision making. In addition, high allocations could foster the growth of militarized movements or strengthen their public appeal (Noel-Baker, 1958). For these reasons, the use of force might be a more likely policy choice during a serious confrontation. The probability of war may also be affected if one nation seeks to bring its opponent's economy "to its knees" through protracted arms competition. Some argue that a strategy of increasing military allocations to unacceptable levels could lead one side to back down and avoid conflict. Another school of thought states that the overallocated protagonist could launch a preventive war before it falls behind its less encumbered foe in the future.

Whatever their effects, military allocations have been seen as an important factor in accounting for the outbreak of war by a variety of scholars. Prior empirical research on military allocations and conflict, using a variety of different indicators, has yielded varied and often inconsistent results (Weede, 1981). Newcombe and Wert (1973) discovered a positive relationship between high allocations and conflict involvement. Rummel's (1972) work shows similar findings, but the relationship is much weaker. In contrast, Kegley et al. (1978) find high allocations positively associated with external conflict only in the relative absence of domestic conflict. Choucri and North (1975) report

wide variation in the military allocation ratios of the major powers prior to World War I.

Although some of this inconsistency in results undoubtedly can be traced to differences in the spatial-temporal domain, the variation in indicators could also be responsible. Before any further theoretical models are constructed or empirical analyses using military allocations performed, it might be useful to step back and consider the different indicators of military allocations. We can perhaps account for some of the differences among previous studies by identifying the effects of using different indicators. More important, the identification of the strengths and weaknesses of each indicator will enable a more informed choice of indicators in the future, hence greater validity and a better understanding of how military allocations affect the likelihood of war.

In considering different indicators of military allocations, we wish to appraise the face, convergent, discriminant, and predictive validity of each measure. In the first section, we specify a number of face validity criteria for the construction of a "good" indicator of military allocations. Next, we describe each measurement approach and evaluate it along those criteria, detailing its advantages and shortcomings. We then specify a set of empirical criteria and proceed to conduct various statistical tests in order to assess the other forms of validity for each indicator. We end with a study of militarized confrontations and the role of military allocations. The latter includes separate analyses for each indicator in order to permit a comparison of the results. In the conclusion, we hope to be able to evaluate each military allocation indicator and provide a guide to those contemplating research on military allocations and conflict. A note of caution to the reader is appropriate at the outset. A "good" indicator of any concept is more than a function of its ability to pass a series of statistical tests. The choice of indicator should also be dependent on the theoretical framework used and should be appropriate for the model tested. Our analysis is meant only to investigate the issues of comparability and empirical validity and not those related to the appropriate theoretical approach, which can vary greatly.

CRITERIA FOR INDICATOR CONSTRUCTION

Before conducting extensive empirical analyses, we believe a good indicator of military allocations must pass at least four tests:

(1) Is the indicator valid over space and time? A good military allocation indicator should be applicable to a wide range of countries. When dealing with a large spatial domain, the indicator must be valid for both large and small countries, making allowances if necessary to insure comparability. Because economic data are frequently used in indicator construction, the approach must guarantee that the indicator will permit valid comparisons between countries across the spectrum from capitalist to socialist economies and from developed to underdeveloped countries.

Equally important, the indicator must be useful throughout the long temporal domain that many conflict studies analyze. Standardization of measures across different historical epochs, however, is often problematic. There must be a balance between choosing a valid indicator for a particular historical period and insuring comparability across periods.

(2) Does the measurement approach produce a baseline by which abnormal allocations to the military can be detected? A major emphasis in conflict studies is the hypothesized relationship between high military allocations and nation-state behavior. It seems to us, then, that a good military allocation indicator should provide the scholar with a means of determining what is a "normal" or "average" military allocation and by implication what is an "overallocation."[1] This becomes particularly important in light of the third criterion below. Most desirable would be a method that can identify "overallocating nations," and "underallocating nations," while still permitting subsequent interval level data analysis.

(3) Can the measurement approach adjust for changes in the baseline over time? More than validity over a long period of time, the indicator should be able to adjust for changes over time in the baseline for normal allocations. Russett (1970) noted a "ratchet" effect in military personnel allocations for the United States following participation in a war. In fact, normal military allocations for all major powers have risen dramatically since the Congress of Vienna (Diehl and Goertz, 1985). What may be a normal military allocation in one period could be an

1. It could be argued that an explicit method for determining "overallocation" is not necessary for a good indicator of military allocations. Nevertheless, if such controls are not introduced, the problem of parameter instability arises, as noted in the third criterion. This problem has plagued a great deal of conflict research. For example, some of the Correlates of War research shows that estimated parameters are often different for the nineteenth and the twentieth centuries. For elaboration on this and related issues, see Goertz (1984).

overallocation and a burden in an earlier epoch. Thus, a good indicator must be sensitive to changes in the norm for military allocations, lest it label all nations in the most recent time period as "overallocators" and all those in earlier periods as "underallocators."

(4) Are the data needed for the indicator available? Expediency in data collection is a poor justification for choosing one measurement approach over another, but even an indicator meeting all the above criteria is useless if the data required are unavailable. The lack of data may force a scholar to narrow the spatial-temporal domain of a proposed study, thereby limiting the level of generalization possible.

Given these criteria, we now turn to a description and evaluation of the various measurement approaches.

MEASUREMENT APPROACHES

A comparison of some portion of a nation's resource base with the resources devoted to the military is most often used to measure military allocations in conflict research. Most studies use military expenditures to represent resources devoted to the military. The variation in measurement approaches tends to center on the choice of indicator for a nation's resource base and the technique of comparison.

With respect to techniques currently used, approaches to measuring military allocations can be roughly divided into two categories: those based on a simple ratio and those that are regression-based. The former use a measure of military appropriation and divide it by an indicator of a nation's resource base. The latter use regression analyses to obtain a predicted or normal military allocation (as a function of the resource base) and then compare the actual allocation with the predicted allocation.

The first measurement approach to be considered has variations that encompass both of these basic techniques.

MILITARY EXPENDITURES AND GROSS NATIONAL PRODUCT

Perhaps the most common indicator of military allocations is a simple ratio of a nation's military expenditures to its Gross National Product, ME/GNP (Russett, 1964; 1970; Weede, 1977; Kegley et al.,

1978; Reisinger, 1983). The rationale for this approach is that GNP is the best indicator of a nation's total available resources. Consistent with most other measures of military allocations, military expenditures are believed to be the best available measure of resources appropriated to the military. Military expenditure figures generally include a wide range of items (from research and development to military hardware) and therefore are more accurate in defining what is actually allocated to a nation's military establishment.

A more sophisticated, regression-based measurement approach (Newcombe, 1969; Newcombe et al., 1973, 1974) uses GNP and military expenditures in a different manner. The size of a nation's military establishment is postulated to be, to a large extent, a function of the size of its economic base. Therefore, military expenditures are regressed on GNP for a large number of nations over a three-year period. After the regression equation is derived, it is applied to yearly GNP figures for each country to obtain a predicted value for that state's military expenditures. The observed (actual) expenditures are then divided by the predicted values and multiplied by 100, the product being called a "tensiometer." Scores that deviate significantly from 100 indicate over- or underallocation to the military relative to other countries in the period studied.

The Newcombe approach has the advantage of establishing a baseline by which to compare different nations' military allocations. The simple military expenditures-GNP ratio, however, provides no systematic method of defining which nations are over- or under-allocating. In addition, the foreign policies of minor powers have a narrower scope, and those nations may spend proportionately less on the military than major powers (Weede, 1977). Some minor powers (for example, Japan) are protected by alliances and may feel little need to maintain a large military establishment, whereas others receive a great deal of external military aid. By including a wide range of countries in the regression baseline (most of which are small countries), the tensiometer might identify major powers to be overallocating, whereas their minor power counterparts would appear to be below the average. Thus, the tensiometer does not distinguish between the potentially different norms of military allocations for major and minor powers.

There are a number of problems with using GNP in either the ratio or regression-based technique. First, GNP is a relatively new concept, dating only to the inception of Keynesian economics. There are serious data problems in a longitudinal study extending before World War I.

Estimates of GNP are available (often only at 10-year intervals) for the nineteenth century, but they must always be used with a great deal of caution.

More than just data availability problems complicate the utility of GNP; its validity over space and time must be questioned. A high GNP in recent times may be, in large part, the result of a rapidly expanding service sector instead of industrial strength, giving a false indication of the resources that can be converted to military purposes. Thus, a nation may have a growing GNP, but still experience increasing strain from its military allocations because of declining industrial production, increased military spending, or both. GNP is also a statistic that is heavily biased toward capitalist economies. Gross National Product figures are hard to calculate for centrally planned economies; estimates are often made, but these can be misleading, if not wholly inaccurate (for instance, note the widely varying estimates of Soviet GNP). In addition, production not exchanged in markets is not included or crude estimates are made (for example, China).

Finally, neither of the GNP approaches describes a distinct method for ascertaining possible changes over time. Comparing military allocation ratios of nations in the nineteenth century with those in the twentieth century could be deceiving; the normal level of allocations may have risen considerably over time (see below).

Overall, the GNP approach is simple enough and relies on military expenditure data that are both available and comparable across a broad spatial-temporal domain. Nevertheless, GNP itself has validity and data availability problems that can be quite severe, and even the Newcombe variation, which does provide a baseline for determining overallocation, does not attempt to detect changes in that baseline.[2]

MILITARY EXPENDITURES AND NATIONAL INCOME

Another approach to measuring military allocations substitutes national income (NI) for GNP as an indicator of a nation's resource base, ME/NI (Nincic, 1983). The advantage of national income relative to GNP is that NI does not include indirect business taxes and allows for

2. Problems with Newcombe's approach would be particularly evident when comparing pre- and post-World War II epochs. The tremendous increase in the number of sovereign states during the postwar period could have dramatic effects on the baseline.

a capital consumption allowance. Thus, NI is able to present a more accurate picture of the actual amount of national resources available.

The national income indicator suffers from the same drawbacks as the GNP-based indicator. Service sector distortions are still present, and NI data for centrally planned economies are suspect. GNP and NI are highly correlated, as they share the same components, except for the two noted above. Consequently, we will not undertake separate analyses for the national income indicator in the remainder of the study.

MILITARY EXPENDITURES AND GOVERNMENT BUDGETS

Another variation of measuring military allocations consists of taking military expenditures as a percentage of government budget (Rummel, 1972; Haas, 1974; Choucri and North, 1975; Cusack and Ward, 1981). Here, the assumption is that the only resources actually available for military purposes are those available to a national government.

Government budgets have the advantage of being a readily available data item for an extended period of time (the data, however, are sporadic for certain closed societies, such as the People's Republic of China). Nevertheless, there are important limitations to the military expenditures/government budget indicator. Government budgets are not sufficiently comparable, given the great range of items that is found in different types of economies. Socialist economies tend to channel a greater share of their resources through government budgets than do market economies, thereby making military allocations of socialist nations appear smaller than those in which the governmental role is more modest. This difficulty will also arise when comparing mid- to late twentieth century governments, in which the role of government in society is relatively great, with their eighteenth and nineteenth century predecessors, which maintained a smaller presence. In the absence of any method to control for historical changes in the scope of government budgets and considering that change occurs at different rates in different countries, the risk of distortion and misconception is great; the lack of a specified baseline for determining normal allocations compounds this problem.

In summary, the government budget approach offers the minor advantage of data availability over the previously cited measurement

approaches, while presenting similar or more difficult problems and disadvantages.

MILITARY PERSONNEL AND TOTAL POPULATION

Rather than focus on capital resources allocated to the military, another approach to measuring military allocations concentrates on human resources devoted to the military. Andreski (1968) developed the idea of a "military participation ratio" to signify the proportion of individuals in the population used by the military. In conflict research, this has been operationally defined as the number of military personnel of a nation divided by its population (Stoll and Champion, 1977; Rummel, 1972; Russett, 1964). The reasoning behind this indicator is that those individuals in military service are unable to contribute to a nation's production of goods and services, and therefore are a burden on the economy. Benoit and Lubbell (1967) carry this one step further operationally: they multiply the number of men in the armed forces by the average civilian wage or salary, labeling the result "lost production" from military service.

Data availability problems with this indicator are all but nonexistent, but its validity can be challenged. The significance of manpower for military preparedness has declined greatly over the last century probably as a result of a shift in emphasis from labor-intensive to capital-intensive warfare. Success in combat has become more dependent on weaponry than sheer manpower. Accordingly, new resources are being channeled into weapons development and procurement, rather than into increases in troop strength. Thus, the number of military personnel in a nation probably no longer adequately reflects the economic investment in its military. The most technologically advanced nations might have the lowest scores on this measure, even though they may allocate the greatest proportion of financial and industrial resources to the military. The validity of this approach over time is, therefore, in doubt.

In addition, the approach provides no baseline for comparing normal and abnormal allocations. Exacerbating this difficulty is a notable change in the pattern of human resource allocations to the military. Following World War II, virtually all of the major powers maintain large standing armies, regardless of the size of their population. Unfortunately, the present approach cannot adjust for this change.

Overall, the military personnel-based measure may be an adequate indicator of military allocations for the preindustrial age, but its ignorance of the changing structure of appropriations makes it less valid in the modern era.

COMPARATIVE SHARES OF INDUSTRIAL CAPABILITY AND MILITARY EXPENDITURE USING SYSTEM PERCENTAGES

A recent approach (Wayman et al., 1983) relies, in part, on industrial indicators of resource capability, rather than aggregate GNP or national income figures. Iron/steel production and energy consumption are used to represent a nation's industrial strength and, in turn, its capacity for military production and supply.

Wayman et al. begin by calculating the major power subsystem totals for military expenditures and the two industrial indicators. Then, they determine the percentage of the total system capabilities that each major power has on each of the three dimensions.[3] The underlying assumption of their next set of transformations is that a nation should have approximately the same percentage of the system's military capabilities as it does for industrial capabilities. For example, if Nation A has 10% of the industrial capabilities in the subsystem, it should also be expected to have 10% of the military capabilities. This assumption provides an *a priori* norm for military allocations and as such is not empirically confirmed. This approach calls for dividing the percentage figures (military expenditures by each industrial indicator), with values over one signifying overallocation to the military.

This approach has some attractive features. It establishes a baseline by which the researcher is able to identify military allocations in excess of the norm, thus permitting analyses of the effects of overallocation. The baseline is defined a priori, and there is, therefore, no problem with it changing over time, as potentially exists with empirical baselines. The indicator permits comparisons across many different kinds of national economies throughout a broad time frame. By using percentage shares instead of raw data, the approach is not subject to problems from inflation. Finally, the data required are readily available, and the

3. The rationale and procedure for calculating system capabilities is laid out in Singer et al. (1972).

authors were able to conduct a set of analyses extending back to 1816.[4]

One potential problem is that the indicator is relative and based on single-year capability figures. Thus, a finding that all major powers are overallocating simultaneously is impossible. A nation may be severely strained by its military spending, but this approach will show it to be underallocating if its peers are allocating more heavily to their militaries. In addition, large concurrent increases in military expenditure for all major powers are not adequately reflected by this approach. Each nation could triple its military spending without any increase in industrial capability and yet retain the same military allocation score.

This method has the potential to be applied to minor as well as major powers, although some allowance must be made for the large number of small countries that do not produce any steel. This approach has shown that it fulfills that criteria better than previous attempts, yet it still has some drawbacks for use in conflict research.

PREDICTED MILITARY EXPENDITURES AND ACTUAL MILITARY EXPENDITURES STRATIFIED BY EPOCHS

This approach assumes that military personnel are more significant than other aspects of the military in determining military effectiveness for the early and mid-nineteenth century. Thus, the focus is on the relationship between military personnel and total population until 1860. Citing the problems with GNP, national income, and government budgets, the analysis centers on military expenditures and the two industrial indicators of Wayman and his colleagues for the post-1860 period (Diehl, 1985).

It is hypothesized that the size of a nation's army and its military budget are a function of the human and capital resources, respectively, available to that nation, and an empirical analysis confirmed this point. To establish a baseline, a variation of the Newcombe and Wert (1973) technique is employed. For the period 1816-1860, military personnel are regressed on total population. For the period after 1860, military expenditures are the outcome variable and the two industrial indicators (energy consumption and iron/steel production) serve as separate predictor variables. The post-1860 regressions are stratified by historical

4. One data limitation of their analysis was the unavailability of energy consumption figures prior to 1860.

epochs (1861-1913, 1919-1938, 1946-1980) corresponding to observed changes in the baseline.

The observed yearly population and industrial figures are applied to the equations obtained from the four regressions and provide expected values for military personnel and military expenditures. The actual personnel and expenditure values are then divided by the predicted values and the quotient is multiplied by 100. The scores of the industrial indicators are averaged to form a composite score. A score of 100 indicates the average or "normal" number of troops or military expenditure for a nation of its particular size.

This approach has the advantages of adjusting for parameter changes over time; thus, different historical epochs can be easily compared with each other. Furthermore, the measurement approach seems valid over space and time, using different indicators for the pre- and post-1860 period as well as employing comparable measures (military personnel, energy consumption, and so on) throughout. Data availability is not a significant problem and different baselines for each epoch, empirically derived, are important by-products of the approach.

As to the limitations of this approach, the relationship between military spending and resource capability may be strong over a 165-year period, but it is noticeably weaker if confined to the 1919-1938 epoch. Because the "fit" of the regression line is rather imperfect for this period, the discrepancy between predicted and actual expenditures is quite large, leading to greater variation in the military allocation scores for this period than other epochs. The outlying scores (particularly the highest ones) exaggerate the actual allocations of countries and distort the results when used in analyses on conflict. (It should be noted that the 1919-1938 period is extremely confict ridden and includes a number of cases in which conflicts escalated to war.)

This approach has been attempted thus far only on major powers, and its applicability to minor powers is uncertain. It appears that some adjustments relating to the technological development of minor powers and military aid would be necessary before the measure can be useful in this domain. Finally, there is a serious problem when attempting to chart military allocations over time by analyzing the slopes and intercepts of the regression lines. Because the dependent variable is monetary (for the post-1860 period), inflation will exaggerate the actual "ratchet" effect or changes occurring across epochs; this problem can be solved if suitable deflators are found.

TABLE 1
Summary of Military Allocation Indicators and Their Attributes

Indicator	*Military Resources*	*Resource Base*	*Technique of Construction*	*"Norm" or Baseline*	*System Change Sensitivity*	*Data Availability*	*Applicability over Space and Time*	*Sample Citation*
1. System %	mil. exp.	energy iron/steel	ratio of system % shares	yes (a priori)	yes	good	good	Wayman, Singer, Goertz (1983)
2. Stratified epochs	mil. exp.	energy iron/steel	regression	yes (empirical)	yes	good	good	Diehl (1985)
3. Budget	mil. exp.	gov't budget	ratio	no	no	fair	poor	Rummel (1972)
4. GNP	mil. exp.	GNP	ratio	no	no	fair	poor	Weede (1977)
5. Personnel	mil. personnel	total population	ratio	no	no	good	poor	Stoll and Champion (1977)

These are the main approaches to measuring military allocations in the conflict literature; by no means is this an exhaustive list of the *possible* measurement approaches, although this is almost an exhaustive list of *actual* approaches. Table 1 gives a summary of the different approaches along with an evaluation of their validity on the criteria given above. Scholars may wish to combine different techniques with various indicators of resource capability. Techniques and indicators beyond those mentioned here are also options for future study.

DEVELOPING SOME EMPIRICAL CRITERIA

Having established a number of criteria for face validity and evaluated each measurement approach according to those criteria, we can now consider different indicators' validity based on empirical tests. We investigate three kinds of validity as they relate to each measurement approach: convergent, discriminant, and predictive validity (Campbell and Fiske, 1959; Ghiselli et al., 1981).

(1) *Convergent validity* requires agreement between scores obtained with two or more indicators presumably measuring the same construct. To investigate this, we will run simple intercorrelation analyses, with special attention to measures that yield the highest coefficients. The greater the correlation of a measure to all others, the more validity that measure has in terms of capturing the same concept.

The reader and the investigators must be careful in assessing the intercorrelations. Similarities index construction or data may account for some portion of a high correlation coefficient, and thereby give a false indication of the convergent validity of the measures in question.

(2) *Discriminant validity* requires disagreement between indicators measuring different constructs. This insures that the indicator is actually measuring what it purports to, rather than some other phenomena. Because we are interested in indicators used in conflict research, we have chosen to run correlations between each military allocation indicator and an indicator of the intensity of a nation's arms buildup. This also provides for a comparison of a static measure (military allocations) and a dynamic one (a military buildup). There should be some positive correlation between the allocation measures and the arms buildup indicator; in fact, a nation's allocation ratio often increases when that nation rapidly increases its weapons acquisition. Nevertheless, the

relationship is far from a perfect one. An increase in arms is not always translated into an increase in the allocation ratio because a nation's resource base may concurrently increase, leaving the allocation ratio approximately the same, even after a significant increase in weapons stocks. A high allocation ratio may also persist long after the termination of an arms buildup, reflecting the impact of prior increases in arms acquisition.

Therefore, we might expect some positive correlation between the rate of increase in arms acquisition and the indicators of military allocations, but that correlation should not be very high. A negative or a high, positive correlation for any of our measures would call into question the validity of that indicator.

(3) *Predictive validity* is an evaluation of an indicator's utility in predicting or postdicting a given outcome, according to some theoretical model. In assessing predictive validity, there is the inherent risk of choosing an incorrect model on which the analysis is based. As a consequence, one must be cautious when interpreting results. The failure to confirm a hypothesis may result from a faulty model rather than from an invalid indicator.

In this study, each measure will be assessed to determine its ability to predict escalation to war, according to the hypothesis that high military allocations increase the risk of a conflict escalating to full-scale war. We will analyze a set of militarized disputes and their outcomes, using each allocation measure as the independent variable. Those measures that yield the strongest results can be said to have the greatest validity on this dimension. These tests should also reveal the extent to which results differ according to the measurement approach used. This could be an important exercise in reconciling divergent findings in the literature, as well as a valuable piece of information for scholars contemplating work in this area.

In empirically assessing the various measures of military allocations, we focus on major power nations in the years 1861-1980.[5] The Correlates of War Project provided data for military expenditures,

5. The major powers for this study are those identified by Small and Singer (1982): United States (1899-1980), the United Kingdom (1861-1980), France (1861-1940, 1945-1980), Austria-Hungary (1861-1918), Germany (1861-1918, 1925-1945); Russia (1861-1917), Soviet Union (1922-1980), China (1950-1980), Italy (1861-1943), and Japan (1895-1945). Inasmuch as the participants' military allocations during major power wars are all but impossible to measure when economies are fully devoted to the war effort, we bypass analysis for the World War I and World II years (1914-1918, 1939-1945).

military personnel, energy consumption, iron and steel production, and total population. Government expenditure data are taken primarily from Banks (1971), supplemented by Mitchell (1981) and the Arms Control and Disarmament Agency (1983). Gross National Product figures were primarily taken from Bairoch (1974), who also graciously provided some unpublished estimates. Because the GNP data are given at 10-year intervals, a polynomial regression was run for each country (all r^2 values are over .99) in order to obtain yearly estimates. It must be pointed out that estimates represent basic trends and do not reflect fluctuations due to the business cycle; thus, there is some error associated with them. Additional GNP data were taken, when necessary, from the World Bank (1983), U.S. Bureau of the Census (1975), Arms Control and Disarmanent Agency (1983), and the International Monetary Fund (1982).[6] "Missing data" are a problem, particularly for government expenditures in the case of China and, to a lesser extent, Germany and the Soviet Union.[7] All monetary data are converted to a common currency,[8] according to Correlates of War Project conversion rates, in order to permit comparability.

For the analysis of discriminant validity, we use a measure of military buildup intensity that is an exponentially weighted average of military expenditure increases over a 5-year period (Diehl, 1985).

In the analysis of predictive validity, we concentrate on "militarized disputes" and their outcomes. Militarized disputes are "a set of interactions between or among states involving threats to use military force, displays of military force, or actual use of military force . . . these acts must be explicit, overt, nonaccidental, and government sanctioned" (Gochman and Maoz, 1984: 587). In determining the outcome of these disputes, we code all disputes that resulted in 1,000 or more battle-

6. Banks (1971) supplied government expenditure data for all years except: 1966-1970 (Mitchell, 1981), 1971-1980 (Arms Control and Disarmament Agency, 1983), and 1929-1933 for Germany and the Soviet Union (Mitchell, 1981). Bairoch (1974) provided all data on GNP except: 1929-1979 for the United States (U.S. Bureau of the Census, 1975), 1950-1980 for China (World Bank, 1983), and 1971-1980 (Arms Control and Disarmament Agency, 1983). Other scattered missing values were filled in by data from the International Monetary Fund (1982). In all cases where data from different sources were combined, we undertook a careful effort to make sure that the figures formed a consistent pattern. Where the data exhibited no consistent pattern or were contradictory, we coded those years as "missing data" as noted below.

7. "Missing data" for government expenditures included 1861-1970 and 1934-1938 for Germany and 1934-1938 for the Soviet Union. In addition, there are "missing data" for Chinese government expenditures and GNP for many years between 1950 and 1980.

8. Monetary data are measured in U.S. dollars.

TABLE 2
Correlations of Military Allocation Indicators

	1. System %	2. Stratified epochs	3. Budget	4. GNP	5. Personnel
1. System %	1.00	−.09	−.19	−.15	−.05
2. Stratified epochs	−.09	1.00	.02	.05	.13
3. Budget	−.19	.02	1.00	.65	.27
4. GNP	−.15	.05	.65	1.00	.41
5. Personnel	−.05	.13	.27	.41	1.00

NOTE: All coefficients are Pearson's r coefficients.

related fatalities as a "war" (Small and Singer, 1982). All those not satisfying this criterion are labeled as "no war." A list of major power militarized disputes and their outcomes is taken from the files of the Correlates of War Project.

ANALYSIS OF CONVERGENT VALIDITY

The simplest and most straightforward way to compare our measures is by correlating them; the correlations are presented in Table 2.

For indicators purporting to measure the same thing, one is struck by the low correlations; only two coefficients are greater than .4. This is surprising in that four of the five indicators use military expenditures and differ only in the measure of economic resources and the technique of construction. The importance of the technique of construction can be seen in the fact that the system percentage method (Wayman et al., 1983) is virtually uncorrelated with the stratified epochs one (Diehl, 1985), despite using exactly the same data (military expenditures, iron and steel production, and energy consumption). The popular GNP method (Russett, 1970 and others) is not highly correlated with either of those two indicators, but is highly associated with indicators using the government budget (Rummel, 1982 and others) and military personnel (Stoll and Champion, 1977) methods. Indeed, the GNP method yields the highest average correlation with the other indicators, but the coefficient is still quite low (.24).[9]

9. The average correlations for each indicator with all others are as follows: system percentages (−.12), stratified epochs (.03), budget (.19), GNP (.24), and personnel (.19).

One possible explanation for the low correlations with the stratified epochs approach is that for each epoch (1861-1913; 1919-1938; 1946-1980), a different regression line is used to calculate the indicator. Using this hint, correlations were calculated for each epoch separately, and the results are given in Table 3.

A quick glance will reveal that the correlations are much higher for the epochs individually than for the whole 120-year period. A broad overview of the results for the three periods seems to indicate that World War II is a watershed in that the correlations are significantly lower (except for military personnel) in the pre-World War II period. Once again, the GNP-based indicator yields the highest average correlations with the other indicators; it had the highest average correlation in the first two epochs and finished second best to the system percentage method in the third epoch.[10]

If one examines the changes over time in the various coefficients, it becomes quite clear why the correlations in Table 2 are so low. The evolution of the indicators over time with respect to one another is nonlinear. For example, the correlations between GNP and government budget indicators go from .23 in 1861-1913, to .66 in 1919-1938, to .18 1946-1980; the correlations between indicators based on the system percentage and stratified epochs approaches go from -.22, to .00, to .76 over the three time periods. Thus, the relations between indicators are changing over time in various nonlinear ways.

Another possible explanation for why the overall correlations are so low is that certain indicators are dependent on the composition of the major power subsystem (the system percentage and stratified epochs methods), whereas others (personnel, GNP, and government budget methods) are unaffected by the composition of that subsystem and its changes. In the former, the baseline used to calculate the indicator will change when countries enter and depart the major power subsystem; the latter are constructed with data for each country only and hence are not influenced by such changes. The cutoff points that we have chosen are the World War I and II years when some dramatic changes occurred in the composition of the major power subsystem.

One way to examine if the values of nonsystem dependent indicators are changing fundamentally over time is to see if the average values of

10. The average correlations for each indicator with all others, stratified by the epochs 1861-1914, 1919-1939, and 1945-1980 are: system percentages (-.06; .29; .56), stratified epochs (.15; .34; .44), budget (.13; .35; .3), GNP (.39; .56; .52), and personnel (.31; .20; .37).

TABLE 3
Correlations of Military Allocation Indicators, Stratified by Epochs

1861-1913 Indicator					
1. System %	1.00	−.22	−.14	.13	−.03
2. Stratified epochs	−.22	1.00	.27	.36	.21
3. Budget	−.14	.27	1.00	.23	.16
4. GNP	.13	.36	.23	1.00	.87
5. Personnel	−.03	.21	.16	.87	1.00
	1. System %	2. Stratified epochs	3. Budget	4. GNP	5. Personnel

1919-1938 Indicator					
1. System %	1.00	.00	.45	.39	.33
2. Stratified epochs	.00	1.00	.44	.78	.14
3. Budget	.45	.44	1.00	.66	−.04
4. GNP	.39	.78	.66	1.00	.40
5. Personnel	.33	.14	−.04	.40	1.00
	1. System %	2. Stratified epochs	3. Budget	4. GNP	5. Personnel

1946-1980 Indicator					
1. System %	1.00	.76	.31	.75	.44
2. Stratified epochs	.76	1.00	.37	.52	.09
3. Budget	.31	.37	1.00	.17	.32
4. GNP	.75	.52	.17	1.00	.65
5. Personnel	.44	.09	.32	.65	1.00
	1. System %	2. Stratified epochs	3. Budget	4. GNP	5. Personnel

NOTE: All coefficients are Pearson's r coefficients.

these ratios increase or decrease dramatically over the three epochs. To investigate this possibility, we use an ANOVA (although technically not all of its assumptions are met here), which tests to see if the means of the three epochs are equal; these results are given in Table 4.

The average value for each indicator has increased over time (the F-statistic is significant at .0001 in each case, except the military personnel indicator, and the follow-up tests between individual means are significant at .01), although the rates of increase are different in each case. This, incidentally, supports a finding reported elsewhere (Diehl and Goertz, 1985) that the major subsystem is becoming increasingly militarized over time.

A less stringent comparison of the five indicators is a rank-order correlation that gives the degree to which the indicators rank the countries concerned in the same order. Given the variability in the sources of data, measures of economic resources, and techniques of construction, if (despite relatively low Pearson's r coefficients) they were to agree on the same ordering, it would still be an indication that they are tapping related concepts.

As Table 5 indicates, the convergence among the indicators signifies that on a grosser level there is more agreement, but still not the high correlations one would expect from indicators of the same phenomenon (although there appears to be quite close convergence on the ranking of major powers following World War II). The GNP method is again the one with the highest average correlation among the indicators.

TABLE 4
Mean Values of Nonsystem Dependent Indicators, Stratified by Epochs

	Epochs		
Indicator	*1861-1913*	*1919-1938*	*1946-1980*
3. Budget	.07	.20	.31
	(.05)	(.16)	(.16)
4. GNP	.01	.04	.11
	(.003)	(.05)	(.08)
5. Personnel	.009	.007	.01
	(.006)	(.005)	(.006)

NOTE: Standard errors are in parentheses.

TABLE 5
Rank-Order Correlations of Military Allocation Indicators

Indicator	1. System %	2. Stratified epochs	3. Budget	4. GNP	5. Personnel
1. System %	1.00	.26 (.22)	−.12 (.09)	.57 (.44)	.31 (.24)
2. Stratified epochs	.26 (.22)	1.00	.44 (.33)	.59 (.49)	.50 (.40)
3. Budget	−.12 (.09)	.44 (.33)	1.00	.23 (.18)	.25 (.20)
4. GNP	.57 (.44)	.59 (.49)	.23 (.18)	1.00	.62 (.50)
5. Personnel	.31 (.24)	.50 (.40)	.25 (.20)	.62 (.50)	1.00

NOTE: Correlations given are Spearman's rho coefficients; those in parentheses are the more conservative coefficients of tau-beta.

When attempting to determine if various indicators are measuring the same concept, one is usually concerned with central tendencies, and hence the use of correlational techniques is quite normal. The situation under investigation here is somewhat different in that researchers have been particularly concerned with countries that are spending abnormally large amounts on the military; in statistical terms, they have been interested in outliers. The previous analyses focused on the degree to which these indicators have the same central tendencies. We now move to a discussion of the degree to which they agree on which states are "over-" or "underspenders."

In this analysis, each state's military allocation, according to each indicator, was classified as either "high," "average," or "low." As we noted above, some indicators inherently define normal levels of allocations and some do not. For those that do not, the mean value for the epoch was classified as "normal" (as we have determined above, an overall mean would be invalid). A definition of "low" or "high" allocations was based on being more than ½ to 1 standard deviations away from the mean (this varied according to the epoch and the indicator, because in some cases outlying values distorted the variance). Each pair of indicators was then compared in a contingency table analysis in order to test the similarity of their classifications of military

allocations; the tau-beta values in Table 6 give a summary of the test results.

The agreement between indicators on the outlying values is clearly higher than their agreement on central tendency (compare with Table 2). This is particularly the case between indicators using military expenditure data. Again, the GNP method yielded the best results, although even under the simplified conditions of only three categories, none of the relationships is particularly strong. As in previous analyses, the tests were repeated for the three historical epochs separately; in contrast to the results reported above, we did not find a regular increase over time in the level of agreement between the different measures.

In summary, our analysis of convergent validity yielded somewhat disappointing results. All forms of analysis revealed only weak congruence between the five indicators, although there was greater agreement when rank ordering countries and identifying outliers. Overall, the GNP method was slightly better than the others on most tests, yet on an absolute level, its performance was less than impressive. We also discovered that increases in the norm or average allocation occurred over time.

ANALYSIS OF DISCRIMINANT VALIDITY

Another test of an indicator is its ability to differentiate itself from indicators of other phenomena. Here, we test the ability of each of the five measures of the military allocations indicator to differentiate itself from an indicator of arms buildup intensity. Earlier, we stated our belief that the correlations between the buildup indicator and the allocation indicators should be positive, but not dramatically high. Table 7 gives the results of our correlational analysis.[11]

The stratified epochs approach and the GNP method perform in the manner desired for discriminant validity. Both indicators exhibit a positive correlation (just above .20), but not so high as to suggest that the same phenomenon is being measured. The other measurement approaches yield very weak correlations, suggesting that different concepts are being measured (this is desirable), but the indicators appear

11. In this analysis, outlying scores (> 20) on the arms buildup intensity indicator are eliminated from consideration so as not to distort the results. Including those cases in the analysis yields correlations of less than .10 for all indicators, none of them significant at .01.

TABLE 6
Cross-Classification Agreement Among Military Allocation Indicators

Indicator					
1. System %	1.00	.23	−.04	.40	.20
2. Stratified epochs	.23	1.00	.32	.46	.10
3. Budget	−.04	.32	1.00	.33	.02*
4. GNP	.40	.46	.33	1.00	.38
5. Personnel	.20	.10	.02*	.38	1.00
	1. System %	2. Stratified epochs	3. Budget	4. GNP	5. Personnel

NOTE: Correlations are tau-beta coefficients. Chi-square values are not given, but were calculated for each pair and were significant in all cases at .001, except where noted by asterisk.
*Significant at .05.

TABLE 7
Correlation Between Indicators of Military Allocation and Arms Buildup Intensity

Military Allocation Indicator	
1. System %	.004
2. Stratified epochs	.22
3. Budget	.11
4. GNP	.21
5. Personnel	.05
	Arms buildup intensity

NOTE: All correlations are Pearson's r coefficients, outliers (> 20.0) on the arms buildup variable are eliminated from consideration.

unable to reflect situations in which military buildups increase the military allocations of a nation.

We now move to our final set of analysis, concerning each indicator's ability to predict conflict escalation.

ANALYSIS OF PREDICTIVE VALIDITY

In looking through the literature that uses these indicators, the majority of it is concerned with the problem of war and the question of

whether "high-allocating" nations are more prone to war than those allocating at more normal levels. Although the indicators considered here do not have high intercorrelations, it remains to be seen whether they produce the same results in the kind of analysis in which they are commonly used.

In order to conduct these exploratory tests, each state in each year was categorized as an "average" (including "low" allocations) or "high" allocator, in the same manner as described in our earlier analysis. Two-by-two contingency tables were constructed with the allocation level as the predictor variable,[12] and the war/no war outcome of militarized disputes (described above) as the outcome variable. The analyses were performed for: (1) the initiator of the dispute (that nation committing the first act involving military force) and (2) the target of the dispute (the victim of the first act). In addition, a third analysis used the ratio of the allocation levels of the initiator and the target as the predictor variable. In that test, the two categories were "equality" ($.67A < B < 1.5A$) and "inequality" ($B < .67A$ or $B > 1.5A$). A summary of all the results is given in Table 8.

In considering the results, the stratified epochs approach is the only one that strongly and consistently associates high allocations with disputes that escalate to war; the chi-square and tau-beta values are significant at .01 under all three testing conditions. The tau-betas for the GNP method are significant in two instances, but at quite low levels.

The remaining three indicators turn in a poor performance in predicting the onset of war. One reason why the system percentage method fared so poorly was its inability to reflect large concurrent increases in military allocations. The system percentage indicator remains unchanged if all or most nations in the major power subsystem increase their allocations prior to a war; in fact, the indicator is capable only of reflecting differences in the growth rates of allocations and not the growth itself. One can apply the same line of reasoning to the government budget indicator. As military expenditures increase, so usually do government budgets, and although the indicator will reflect some the military increase, it still dampens the full impact.

That the personnel indicator is not strongly associated with war is not surprising, as this indicator showed the least variation of all those

12. Military allocation values in this analysis are taken from the year *before* the militarized dispute, in order to insure that the resources allocated are actually available at the time of the confrontation and to eliminate any distortions from allocations that were reactive to the dispute.

TABLE 8
Comparative Performance of Military Allocation Indicators in Predicting Escalation to War

	Initiator		*Target*		*Ratio*	
Indicator	χ^2	*Tb*	χ^2	*Tb*	χ^2	*Tb*
1. System %	.00	.01	.01	.01	3.63	−.20**
2. Stratified epochs	8.33**	.31**	9.78**	.33**	9.00**	.32**
3. Budget	.10	−.04	.08	.03	.43	.09
4. GNP	1.29	.13**	4.44*	.23**	.03	.02
5. Personnel	.28	.06	.33	.06	4.33*	−.22**

*Significant at .05; **significant at .01.

considered, and one might suspect that drastic changes in troop levels occur after, rather than before, the war has begun (Mullins, 1975). In addition, by the Franco-Prussian War of 1870, industrial power was becoming a more important factor in warfare than the number of men that could be put on the field of battle.

Considering our analysis here, it is not surprising that past research has produced a diversity of findings on the effects of high military allocations on conflict. A large portion of this disparity apparently can be traced to the different indicators employed by various scholars. Only the stratified epochs approach performs well in this analysis of predictive validity. The GNP method shows some promise that might be fulfilled if the *ad hoc* procedure used here to define "high allocation" were refined to provide a more sensitive measure. Nevertheless, we must once again remind the reader against interpreting these findings too rigidly. High military allocations may have little relationship with conflict escalation and consequently the indicators with approximately zero correlations could be the most valid. Although we believe otherwise, a greater knowledge of allocation and conflict is necessary before a more definitive interpretation can be made.

CONCLUSIONS

At the outset of this study, we identified the major approaches in conflict research for measuring military allocations. Our goal was to test the validity of those indicators and make a series of generalizations regarding their use.

One conclusion is that one should proceed cautiously in choosing a military allocation indicator; that choice is likely to strongly affect the strength of the relationships discovered. Researchers might do well to test their hypotheses using more than one indicator of miliary allocations in order to determine whether the results are mere artifacts of the indicator construction. Indeed, it might be desirable to use multiple indicators to form a composite score. Different indicators are not likely to be affected by the same measurement error, and in a composite score, the errors could cancel each other out (Manheim and Rich, 1981). Furthermore, scholars must be sensitive to changes over time in the average allocation. What is high allocation in one period may be well below the norm in a later period as nations grow accustomed to increasing levels of military allocation.

Finally, one must be sensitive to the possibility that his or her allocation indicator is inappropriate for a particular historical epoch or set of countries. Even within our limited domain (major powers since 1860), we noted problems with using certain indicators for some time periods and different economic systems. We would expect these difficulties to be at least as prevalent when the spatial-temporal domain is expanded.

In the way of specific conclusions, we cannot give our unqualified endorsement to any of the indicators. Despite using the same spatial-temporal domain and, in many cases, the same data in indicator construction, the convergence among the indicators was rather poor. Nevertheless, two indicators appear to be superior to the others. The stratified epochs approach has strong face validity, scoring well on the criteria of data availability, ability to detect system change, and applicability over a wide spatial-temporal domain. It also yielded positive results when tested for discriminant and predictive validity, although the relationships in the latter are not terribly strong. Its drawbacks, however, were that it was virtually unrelated to its peers as indicated by the low correlations in the analysis of convergent validity. We believe that this approach might be best used in longitudinal studies of allocations by countries that collectively exhibit a variety of economic systems.

The GNP method consistently performed better than the other indicators in the convergent validity analyses and was better than average with respect to discriminant and predictive validity. Although on a relative basis this indicator did well, on an absolute scale, the case for its validity was not overly compelling. The GNP method appeared to

have severe problems with the face validity criteria. Data availability and accuracy seem to be serious problems in the years before World War II, and at any point in time with respect to noncapitalist states. It appears the GNP method might be useful in a study that has a narrow spatial-temporal domain, encompassing only capitalist countries in the last 40 years (for example, a study on burden sharing in NATO).

The remaining three indicators generally scored low on almost all the dimensions of validity. The system percentage method intuitively has some attractive attributes, but its empirical performance usually rated it poorly compared to the other indicators. The government budget indicator suffers from the same spatial-temporal limits as the GNP indicator and did not yield strong empirical results in any phase of the investigation. Finally, the military personnel method seems inappropriate for modern society. Its ignorance of the technological features of conflict is born out by its poor performance in the empirical analyses. It may be that this indicator has some value for the early nineteenth century and before, or for underdeveloped countries, but because of the limited domain of our venture we can only suggest this possibility.

Overall, we suggest that scholars be cautious in their choice of military allocation indicators, cognizant of the scope and theoretical approach of the study in that choice. We would, however, hope that researchers will not be content to utilize only the indicators outlined here. The need for a better measure of military allocations seems to outweigh the convenience associated with acceptance of current methods.

REFERENCES

Arms Control and Disarmament Agency (1983) World Military Expenditures and Arms Transfers 1971-1980. Washington, DC: ACDA.

ANDRESKI, S. (1968) Military Organization and Society. Berkeley: Univ. of California Press.

BAIROCH, P. (1974) "Europe's GNP, 1800-1970." J. of European Econ. History 3, 3: 557-608.

BANKS, A. S. [ed.] (1971) Cross Polity Times Series Data. Cambridge: MIT Press.

BENOIT, E. and H. LUBELL (1967) "The world burden of national defense," pp. 29-59 in E. Benoit (ed.) Disarmament and World Economic Interdependence. New York: Columbia Univ. Press. 29-59.

CAMPBELL, D. and D. FISKE (1959) "Convergent and discriminant validity by the multitrait-multimethod matrix." Psychological Bulletin 56: 81-105.

CHOUCRI, N. and R. NORTH (1975) Nations in Conflict. San Francisco: W. H. Freeman.

CUSACK, T. and M. WARD (1981) "Military spending in the United States, Soviet Union, and the People's Republic of China." J. of Conflict Resolution 25, 3: 429-469.

DIEHL, P. (1985) "Arms races to war: testing some empirical linkages." Soc. Q. 26, 3: 331-349.

——— and G. GOERTZ (1985) "Trends in military allocations since 1816: what goes up does not always come down." Armed Forces and Society 12, 1: 134-144.

GHISELLI, E., J. CAMPBELL, and S. ZEDECK (1981) Measurement Theory for the Behavioral Sciences. San Francisco: W. H. Freeman.

GOCHMAN, C., and Z. MAOZ (1984) "Militarized interstate disputes: procedures, patterns and insights." J. of Conflict Resolution 28, 4: 585-616.

GOERTZ, G. (1984) "Changing parameters across time and space." Presented at European Consortium for Political Research Meetings, Salzburg, Austria.

HAAS, M. (1974) International Conflict. Indianapolis: Bobbs-Merrill.

International Monetary Fund (1982) International Financial Statistics. Washington, DC: IMF.

KEGLEY, C., N. RICHARDSON, and G. RICHTER (1978) "Conflict at home and abroad: an empirical extension." J. of Politics 40: 742-752.

MANHEIM, J. and R. RICH (1981) Empirical Political Analysis. Englewood Cliffs, NJ: Prentice-Hall.

MAOZ, A. (1982) Paths to Conflict. Boulder, CO: Westview.

MITCHELL, B. R. (1981) European Historical Statistics, 1750-1975. New York: Facts on File.

MULLINS, A. F. Jr. (1975) "Manpower Data as a Measure of Arms Race Phenomena." Ann Arbor, MI. (mimeo)

NEWCOMBE, A. (1969) "Towards the development of an inter-nation tensiometer." Peace Research Society Int. Papers 13: 11-27.

——— N. NEWCOMBE, and G. LANDRUS (1974) "The development of an inter-nation tensiometer." Int. Interactions 1, 1: 3-18.

——— and J. WERT (1973) "The use of an inter-nation tensiometer for the prediction of war." Peace Sci. Society Int. Papers 21: 73-83.

NINCIC, M. (1983) "Fluctuations in Soviet defense spending: a research note." J. Conflict Resolution, 27, 4: 648-660.

NOEL-BAKER, P. (1958) The Arms Race: A Programme for World Disarmament. New York: Oceana.

REISINGER, W. M. (1983) "East European military expenditures in the 1970s: collective good or bargaining offer?" Int. Organization 37, 1: 143-155.

RUMMEL, R. (1972) The Dimensions of Nations. Beverly Hills, CA: Sage.

RUSSETT, B. (1964) "Measures of military effort." Amer. Behavioral Scientist 7, 6: 26-29.

———(1970) What Price Vigilance?: The Burdens of National Defense. New Haven, CT: Yale Univ. Press.

———(1982) "Defense expenditures and national well-being." Amer. Pol. Sci. Rev. 76, 4: 767-777.

SINGER, J. D., S. BREMER, and J. STUCKEY (1972) "Capability distribution, uncertainty, and major power war 1820-1965," pp. 19-48 B. Russett (ed.) Peace, War, and Numbers. Beverly Hills, CA: Sage.

SMALL, M. and J. D. SINGER (1982) Resort to Arms: International and Civil Wars, 1816-1980. Beverly Hills, CA: Sage.

STOLL, R. and M. CHAMPION (1977) "Predicting the escalation of serious disputes to war: some preliminary findings." Presented at the North American Peace Science Conference, Philadelphia.

U.S. Bureau of the Census (1975) Historical Statistics of the USA from Colonial Times to 1970. Washington, DC: U.S. Bureau of the Census.

WAYMAN, F., J. D. SINGER, and G. GOERTZ (1983) "Capabilities, allocations, and success in militarized disputes and wars, 1816-1976." Int. Studies Q. 27, 4: 497-515.

WEEDE, E. (1977) "National position in world politics and military allocation ratios in the 1950s and 1960s." Jerusalem J. of Int. Relations 2, 3: 63-80.

———(1981) "Preventing war by nuclear deterrence or by detente." Conflict Management and Peace Science 6, 1: 1-78.

World Bank (1983) China: Socialist Economic Development, vol. I: The Economy. Washington, DC: World Bank.