

MEASURING NATIONAL POWER

KELLY M. KADERA

*Department of Political Science
University of Iowa
Iowa City, IA 52242-1409*

GERALD L. SOROKIN

*Hillel Foundation
Iowa City, IA 52245-1730*

Power's central role in international relations theory is unsurpassed, yet considerable debate persists over the quality of its most commonly used indicator, the Correlates of War project's Composite Indicator of National Capabilities (*CINC*). At issue is whether *CINC*'s main feature, its ability to measure a nation's power relative to other nations' power levels, inadvertently creates errors when membership in the comparison group fluctuates. Using mathematical proofs and an empirical investigation of the major power system, we show that Organski and Kugler (1980) and Gleditsch and Ward (1999) are correct: changes in the comparison group *do* create errors in *CINC*. In particular, *CINC* inadvertently mismeasures dyadic power distributions. Using power transition theory as a context within which to evaluate *CINC*, we find that it creates artificial power transitions, masks actual transitions, changes the timing of transitions, alters the magnitude by which one state overtakes another, and produces specious relationships between transitions and conflict. We also offer a viable alternative measure, called the Geometric Indicator of National Capabilities (*GINC*), and demonstrate how its use of the geometric mean retains *CINC*'s notion of systemically-based *relative* power and immunizes it from the problems afflicting *CINC*. *GINC* is strongly recommended for dyadic analyses, especially when membership in the comparison group fluctuates frequently.

Received for publication August 2003.

*Previous versions of this paper were presented at the annual meeting of the International Studies Association, March 18–22, 1997, Toronto, ON, and at the annual meeting of the ISA-Midwest, October 18–20, 1996, St. Louis, MO. We thank Daniel Morey and Ronald Schwenn for their research assistance and Stuart Bremer, Mark Crescenzi, Paul Diehl, Douglas Dion, Paul Hensel, James Lindsay, Doug Lemke, Karen Nelson, Megan Shannon, and anonymous reviewers for their comments on earlier drafts. All errors remain ours.

Address correspondence to Kelly Kadera, Department of Political Science, 341 Schaeffer Hall, University of Iowa, Iowa City, IA 52242-1409, USA. E-mail: kelly-kadera@uiowa.edu

Keywords: power, *CINC*, *GINC*, measurement error, power transition theory

Because power's central role in international relations theory is unsurpassed, confidence in its most commonly used measure, namely the Correlates of War (COW) project's composite indicator of national power (*CINC*), is crucial to theoretical advancement. Considerable debate over *CINC*'s performance persists (cf. Organski and Kugler, 1980; Gleditsch and Ward, 1999; Kim, 1989; and Houweling and Siccama, 1996). The purpose of this paper is to settle this debate by closely examining *CINC*. We are particularly concerned with whether *CINC*'s main feature, its ability to measure a nation's power relative to other nations' power levels, inadvertently creates errors when the number and identity of states in the comparison group fluctuates. Our findings support Organski and Kugler's and Gleditsch and Ward's conclusion that *CINC* exhibits validity problems, and we demonstrate the ramifications within the context of a popular theoretical framework. A viable alternative measure, called *GINC*, is proposed.

The paper is organized as follows. First, we review power's role in international relations theory and the debate over whether *CINC* is prone to measurement errors. Second, we assess the frequency with which the suspected sources of error occur. Third, we resolve the debate over *CINC*'s susceptibility to mismeasurement by mathematically analyzing its functional form. Here, we use power transition theory as an applied context within which to evaluate *CINC*'s validity. We show that *CINC* inadvertently mismeasures dyadic power distributions: it produces artificial transitions, masks actual transitions, changes the timing of transitions, and alters the magnitude by which one state overtakes another. We also offer a remedial measure and demonstrate its immunity to these problems. Fourth, we show how *CINC*'s errors affect substantive conclusions, again using power transition theory as an example. Last, we summarize our findings and make recommendations about when scholars should use *GINC*.

RELEVANT LITERATURE

Numerous theories of international relations heavily rely on the concept of power. Realism, the dominant paradigm, focuses so strongly on power that it is often called *power politics*. In addition, the fierce debate between balance of power theorists and power transition scholars has long centered on the conflictual effects of two rival nations' relative power levels (DiCicco and Levy, 1999). Power is also instrumental to expected utility models of war, which rely on it to calculate the likelihood of victory (Bueno de Mesquita, 1981; Bueno de Mesquita and Lalman, 1992; Powell, 1996). Polarity theories (Singer et al., 1972; Mansfield, 1992), hegemonic stability theory (Gilpin, 1981 and 1988), long cycle theories (Modelski, 1978, Modelski and Thompson, 1989; Rasler and Thompson, 1994), and relative power cycle theory (Doran, 1989; Doran and Parsons, 1980) also give power a "determining role" (DiCicco and Levy, 1999, p. 679). Power has also been used as a control variable for empirical investigations of many other substantive issues such as the normative and structural determinants of democratic peace (Maoz and Russett, 1993), regime change and alliance shifts (Siverson and Starr, 1994), and the role of regime type in the

settlement of disputes (Dixon, 1994).

Despite the centrality of power in the study of world politics, scholars disagree on the merits of its most common indicator, *CINC*. At the crux of this debate is the question of whether *CINC*'s use of a baseline comparison group makes a nation's score sensitive to variation in the membership of that group. *CINC* is based on two indicators each of military, industrial, and demographic dimensions. The six resulting indicators are each expressed as a proportion of all group members' capabilities on that dimension. The six proportions are then averaged to create a mean share of capabilities, representing how well a state performs relative to its peers. Are these average shares sensitive to variation in which states are included in the ratios' denominators and when they are included?

Intuitive ideas about the effects of changes in the comparison group on the *CINC* measure vary widely. Organski and Kugler criticize *CINC*'s sensitivity to fluctuations in group members' capabilities (1980, p. 36). When a state enters or exits the relevant system, the group's overall capabilities fluctuate more dramatically than they would from a simple shift in one's capabilities. Therefore, entries and exits make this sensitivity even more acute. Gleditsch and Ward make a similar claim, arguing that changes in membership mean that "proportional system characteristics, such as the *CINC* measure...differ over time *by construction*, particularly after the First World War" (1999, p. 397). In contrast, Kim, citing "simple arithmetic," proposes that a state's *CINC* score is immune to changes in the group of states included in the calculation (1989, p. 262, n. 8). Houweling and Siccama concur with Kim, reasoning that as long as a composite measure is used for dyadic purposes, changes in system membership have no effect (1996, p. 109).

These two assessments of the effect of the entry and exit of states are at odds. One tells researchers that entries and exits make *CINC*'s use of a comparison group problematic. The second, apparently based on sound mathematical reasoning, assures us there is nothing to worry about because two nations' scores relative to one another will remain constant regardless of membership in the comparison group. Confident, noted scholars offer both assessments. Who is right? We will mathematically demonstrate that Organski and Kugler's and Gleditsch and Ward's intuition is correct: a nation's *CINC* score does inadvertently reflect the composition of the comparison group at a particular point in time. Before doing so, we address the question of whether or not group membership changes are sufficiently common to present serious problems for *CINC*.

THE REALITY OF GROUP MEMBERSHIP CHANGES

How often does group membership actually change? In any comparison group, shifts can occur: 1) when states are born or when they die, 2) when data become available or unavailable, 3) as states qualify or disqualify as a member of the selected group, and 4) when scholars alter membership rules. Examples follow.

Let us look at states' movement in and out of the entire international system, a common comparison group (e.g., Kim, 1989; Bennett, 1997; Bremer, 1992 and 1993; Mitchell et al., 1999). Changes primarily result from the birth and death of actors and whether or not those actors qualify as states according to the heavily used COW

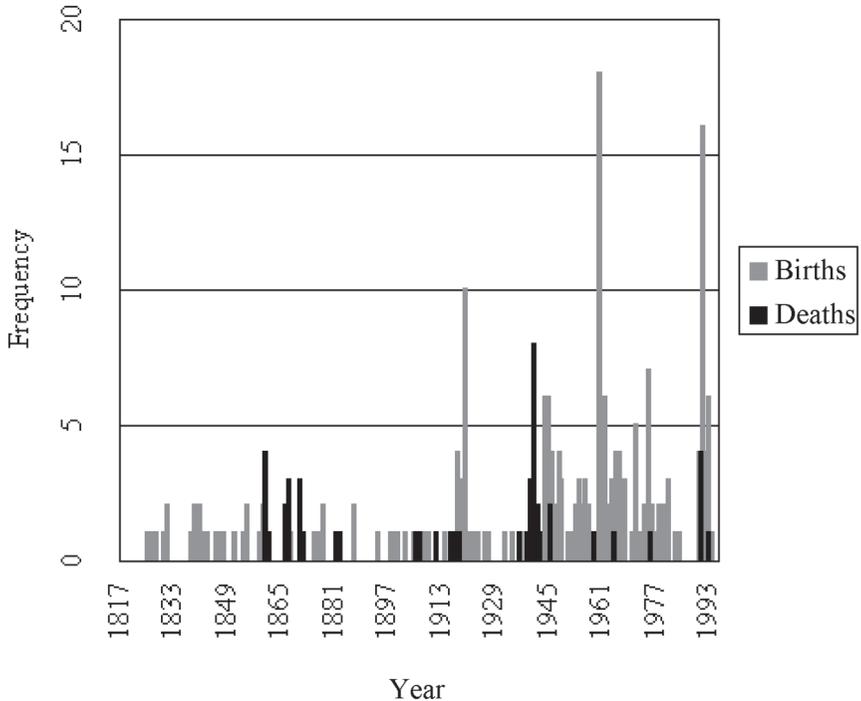


Figure 1. Entry to and Exit from the International System.

Interstate Membership list (Small and Singer, 1982). Figure 1 shows the frequency of state births (e.g., the establishment of Israel) and deaths (e.g., the disintegration of Yugoslavia and the USSR and the subsuming of states as in the Balkans). At least 1 birth occurs in 47.8 percent of the years since 1816, and at least 1 death occurs in 14.3 percent of the years. The aggregate number of states in the system shifts (up or down) from one year to the next over 54 percent of the time. The notable spike in membership at 1920 is largely due to the COW alteration in membership rules after 1919.¹ Sudden upward shifts are also seen during the post World War decolonization in Asia, Africa, and the Middle East and in the early 1990s following the dissolution of the Soviet Union. Generally, membership changes are very common; at least 1 birth or death occurs in 56.6 percent of the years.²

If entries and exits inadvertently affect states' power levels, then modifications of the coding rules for system membership also have serious implications. Several studies propose nontrivial amendments to the COW criteria for statehood (Bennett and Zitomersky, 1982; Gleditsch and Ward, 1999; Fazal, 2000). For example, switching from the COW guidelines to Fazal's adds 19 states to the system, revises the entry date for 24, and increases the proportion that experience death from 22.86 to about 30 percent (2000).

Appraisal of a specialized subset of states as a referent group is also instructive. The most commonly analyzed subgroup consists of the "major powers" (e.g., Goertz

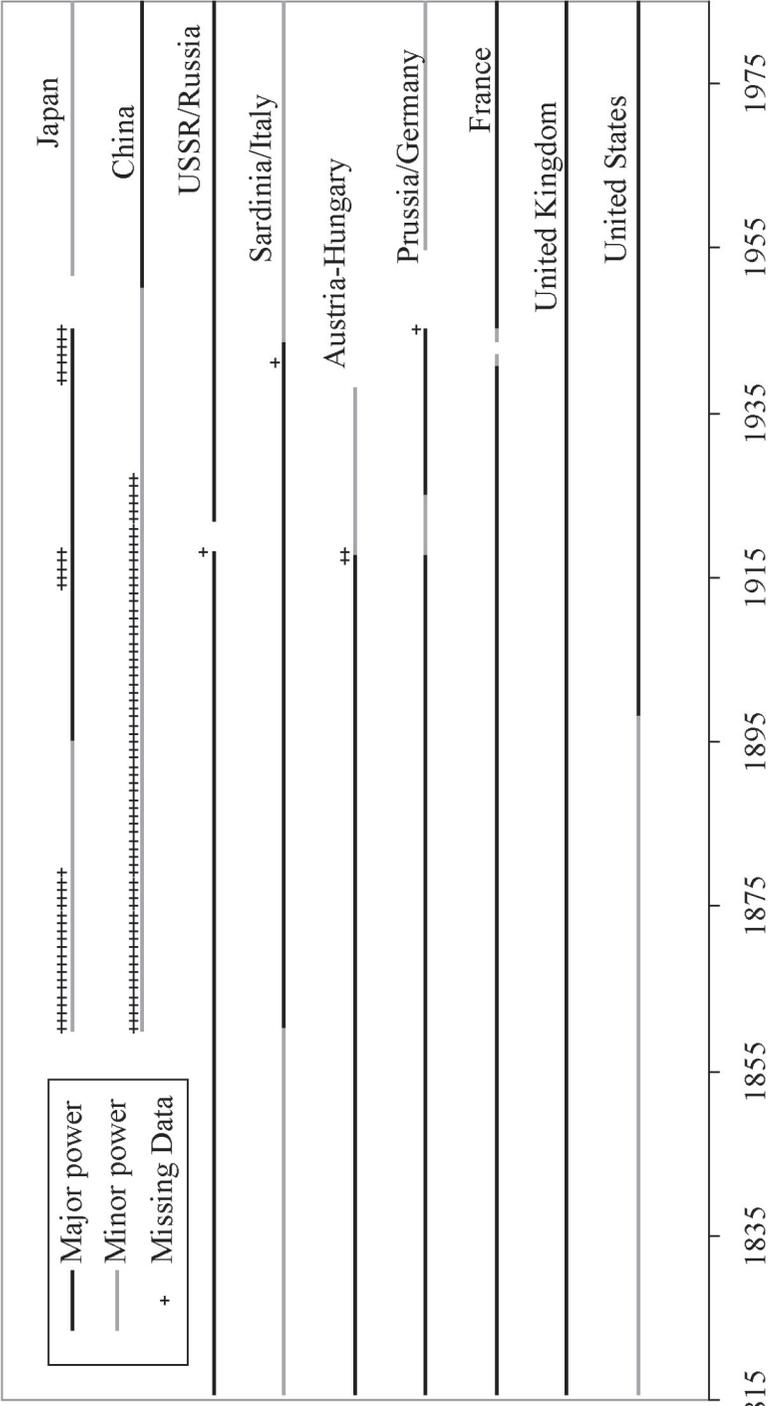


Figure 2. Major Power System Membership and Data Availability.

and Diehl, 1995; Huth et al., 1993; Kadera, 2001; Mansfield, 1992 and 1994), states that are “involved in the bulk of the system’s diplomatic and martial activities” and that “set the tone for other nations’ interactions” (Small and Singer, 1979, p. 65).³ Figure 2 illustrates the frequency of entries and exits in this elite comparison group, according to the COW rules. The composition of the major power system changes 12 times in 170 years, and the arrangement often lasts for only a handful of years.⁴ Sustained membership is prevalent from 1816 to 1859 and from 1950 on. Four states—Japan, Italy, France, and Austria-Hungary/Austria—disappear completely for one or more years, and missing data for Japan and China are especially problematic.

Therefore, exits and entries occur frequently in common analysis sets. Seemingly innocuous issues such as which set of coding rules is used, the birth and death of states, the qualification and disqualification of states as members of a stylized group, and data availability are all culprits. In the following section, we demonstrate how these entries and exits affect *CINC*’s measure of national power and dyadic power distributions.

PROBLEMS AND SOLUTIONS

This section focuses on formal reasoning based on the functional form of the *CINC* score. After introducing some mathematical notation, we discuss *CINC*’s measurement of power transitions and the errors that are introduced by its averaging procedure. Then we propose an alternative composite measure and demonstrate how it remedies *CINC*’s shortcomings.

*Measuring Power with a Composite Indicator*⁵

We establish the following notation for *CINC*’s six national capabilities indicators:

ME_i = nation I’s military expenditures,
 MP_i = nation I’s military personnel,
 IS_i = nation I’s iron and steel production,
 NRG_i = nation I’s energy consumption,
 $UPOP_i$ = nation I’s urban population,
 $TPOP_i$ = nation I’s total population.

The first step in building a *CINC* score is to determine the group of nations comprising the relevant system of analysis. Next, a single nation’s system share of a particular capability is computed. For example, in a 3-nation system, X’s share of total military expenditures, $\%ME_x$, is given by:

$$\%ME_x = \frac{ME_x}{ME_x + ME_y + ME_z} \quad [1]$$

A similar calculation determines nation X’s share of each of the five remaining aspects of national power, producing values for $\%MP_x$, $\%IS_x$, $\%NRG_x$, $\%UP_x$, and

$\%TP_x$. Each proportion ranges from 0 to 1. Averaging combines the proportions into a single indicator:

$$CINC_x = \frac{\%ME_x + \%MP_x + \%IS_x + \%NRG_x + \%UP_x + \%TP_x}{6} \quad [2]$$

This procedure is followed for each nation in the system for each year. The entry and exit of additional states into the referent group is the suspected source of errors, so let:

$CINC_i$ = nation I's power level before nation Z enters the system, and

$CINC_i^*$ = nation I's power level after nation Z enters the system.

Unintended Consequences of CINC

Numerous scholars use *CINC*, or similarly constructed composite measures, to measure dyadic power relationships. In fact, a JSTOR search revealed over 40 such studies since 1985 (see, e.g., Bueno de Mesquita et al., 1997; Geller, 1993; Sample, 1998). We caution researchers about the possibility of erroneous findings stemming from *CINC*'s mismeasurement. To demonstrate *CINC*'s errors, we consider their effects within a specific theoretical context, namely power transition theory (Organski, 1958; Organski and Kugler, 1980). Many tests of this theory also use *CINC* (e.g., Kim, 1989 and 2002; Lemke and Werner, 1996). Our choice was based on the prominence of the theory (DiCicco and Levy, 1999, p. 675). A handful of power-centric theories do not focus on power distributions (e.g., Powell, 1996 and 1999; Werner, 1999); but the many that do typically predict peaks in conflict near power transitions, although there is some disagreement about whether this happens before, during, or after a transition (cf., Bueno de Mesquita and Lalman, 1992; Geller, 1992; Kadera, 2001; Kim and Morrow, 1992; Organski, 1958; Organski and Kugler, 1980). Even those that consider other power-based factors such as the speed of change and the capabilities contributions of allies also include the power transition itself as a key predictor of heightened hostilities (e.g., Kim, 1989 and 2002; Lemke and Werner, 1996). More importantly, we emphasize that the focus on one particular theory is merely for purposes of demonstration. Our finding, that the entry and exit of states inadvertently affects *CINC*, applies to all endeavors to measure dyadic power relations, which has broad-based implications.⁶

Power transition scholars are interested in: 1) whether two opponents' power levels are equal or unequal, 2) whether one opponent is overtaking the other, and 3) when such takeovers, or power transitions, occur in relationship to conflict (Organski and Kugler, 1980, p. 36).⁷ These questions guide our investigation of *CINC*'s performance. Hence, consider the relationship that signals an empirical power transition between X and Y:

$$CINC_x = CINC_y \quad [3]$$

Equivalently:

$$CINC_x - CINC_y = 0 \tag{4}$$

which can be rewritten in terms of the six indicators:

$$\frac{ME_x - ME_y}{ME_x + ME_y} + \frac{MP_x - MP_y}{MP_x + MP_y} + \dots + \frac{TP_x - TP_y}{TP_x + TP_y} = 0 \tag{5}$$

Some of the fractions on the left-hand side of Equation 5 will be positive, indicating that nation X holds the advantage over nation Y for those particular dimensions. Other fractions will be negative, indicating that nation Y holds the advantage in those domains. When a transition is underway, the magnitude of the sum of the positive fractions is exactly equal to the magnitude of the sum of the negative fractions:

$$\Sigma \text{ nation X's advantages} = \Sigma \text{ nation Y's advantages} \tag{6}$$

In the hypothetical example given in the first two columns of Table 1, nation X holds the advantage in iron and steel production and energy consumption while nation Y holds the advantage in military expenditures, military personnel, urban population, and total population. Nation Y enjoys the overall advantage because:

$$\Sigma \text{ nation X's advantages} < \Sigma \text{ nation Y's advantages} \tag{7}$$

That is,

$$\frac{IRST_x - IRST_y}{IRST_x + IRST_y} + \frac{NRG_x - NRG_y}{NRG_x + NRG_y} < \frac{ME_y - ME_x}{ME_x + ME_y} + \frac{MP_y - MP_x}{MP_x + MP_y} + \frac{UPOP_y - UPOP_x}{UPOP_x + UPOP_y} + \frac{TPOP_y - TPOP_x}{TPOP_x + TPOP_y} \tag{8}$$

Given the inequality in Equation 8, the hypothetical example does not represent a transition. Could the entry of nation Z cause an artificial transition to appear? The

Table 1
A Hypothetical Dataset for Calculating Relative Power

Aspect of Power	Nation			System Total		Revised Data for Z
	X	Y	Z	X&Y	X, Y&Z	
Military Expenditures	50	60	40	110	150	120
Military Personnel	17	25	10	42	52	130
Iron & Steel Production	20	15	5	35	40	5
Energy Consumption	65	55	30	120	150	10
Urban Population	15	20	15	35	50	70
Total Population	40	45	35	85	120	90

hypothetical dataset yields the following *CINC* scores for X and Y with and without Z as a group member:

$$\begin{aligned} CINC_x &= .48 & CINC_x^* &= .37 \\ CINC_y &= .52 & CINC_y^* &= .40 \end{aligned} \quad [9]$$

Nations X and Y experience a transition neither with nor without Z's inclusion in the calculation. While the difference between their scores decreases from .04 to .03, it does not become 0. Similarly, the ratio of the two scores increases from .923 to .925, thereby moving closer to, but not becoming 1. Nor does X take the lead. Is this universally true for all cases? In other words, if the difference between two rivals' *CINC* scores is not 0, can it speciously become 0 simply due to an additional actor?

We answer this question by investigating the impact of a third actor on both sides of the inequality (Eq. 8). Nation Z's capabilities, ME_z , MP_z , IS_z , NRG_z , UP_z , and TP_z , are added to the denominators, and each fraction becomes slightly smaller. Both nation X's advantages and nation Y's advantages diminish.

$$\begin{aligned} &\frac{IRST_x - IRST_y}{IRST_x + IRST_y + IRST_z} + \frac{NRG_x - NRG_y}{NRG_x + NRG_y + NRG_z} < \frac{ME_y - ME_x}{ME_x + ME_y + ME_z} + \\ &\frac{MP_y - MP_x}{MP_x + MP_y + MP_z} + \frac{UPOP_y - UPOP_x}{UPOP_x + UPOP_y + UPOP_z} + \\ &\frac{TPOP_y - TPOP_x}{TPOP_x + TPOP_y + TPOP_z} \end{aligned} \quad [10]$$

Does the inequality hold? In the hypothetical example, it does. Nation Y maintains its advantage, albeit a smaller one. The general answer can be found in the sizes of Z's six indicators relative to those of the other two nations. In order for X to gain the advantage after Z's entry, Z's iron and steel production and energy consumption capabilities must be very small compared to X's and Y's; while its military expenditures, military personnel, urban population, and total population must be very large. This would make the fractions on the left-hand side of inequality [8] diminish by much less than the fractions on the right-hand side, so that:

$$\begin{aligned} &\frac{IRST_x - IRST_y}{IRST_x + IRST_y + IRST_z} + \frac{NRG_x - NRG_y}{NRG_x + NRG_y + NRG_z} > \frac{ME_y - ME_x}{ME_x + ME_y + ME_z} + \\ &\frac{MP_y - MP_x}{MP_x + MP_y + MP_z} + \frac{UPOP_y - UPOP_x}{UPOP_x + UPOP_y + UPOP_z} + \\ &\frac{TPOP_y - TPOP_x}{TPOP_x + TPOP_y + TPOP_z} \end{aligned} \quad [11]$$

Nation Z's entry brings about a transition if its capabilities are bifurcated in a particular way. Values for one subset of the six indicators must be especially high, and this subset must match the group for which the previously stronger state (Y in the example) held the advantage. Values in the second subset must be low, and this subset should match the group for which the challenger (X in the example) held the advantage. In other words, the new system member must compete almost exclusively with the stronger state for shares of the group's capabilities. For example, let us examine what would happen if nation Z's capabilities are changed to the values in the last column of Table 1. The scores for X and Y become $CINC_x^* = .281$ and $CINC_y^* = .275$. This new calculation reveals an *artificial power transition*, since X surpasses Y solely because Z enters the picture.⁸

Not only do scholars need to know when power transitions are brought about by the entry of a third nation, they also need to know when the exit of a nation might lead to the avoidance of a transition. Under what conditions might this happen? Suppose that while nation Z is a member, Y surpasses X. Accordingly, we would observe:

$$\text{condition 1: } \Sigma \text{ nation X's advantages} < \Sigma \text{ nation Y's advantages} \quad [12]$$

What happens instead if nation Z exits just before we take stock of the two states' relative advantages? If nation Z's exit causes X and Y to avoid a transition they otherwise would have experienced, then we would observe:

$$\text{condition 2: } \Sigma \text{ nation X's advantages} > \Sigma \text{ nation Y's advantages} \quad [13]$$

For condition 2 to occur when Z leaves, X's power level must grow by much more than Y's does. This means that the capabilities removed from the denominators of the fractions where X had the advantage should be much larger than the capabilities removed from the denominators where Y had the advantage. That is, the departing nation must have been competing for resources most vigorously with the potential power transition loser (X), and ineffectively with the potential transition winner (Y).

Until now, this discussion on the inadvertent effects of *CINC* on power transitions has focused on those measured by the difference between X and Y's scores. It can be demonstrated that transitions measured by the ratio of their scores are similarly

affected. Suppose $\frac{CINC_x}{CINC_y} < 1$. What would artificially cause the inequality to reverse

when Z enters the system? Nation X would gain the advantage if Z were most competitive on dimensions where Y had enjoyed an advantage and least competitive on those where X was previously strong. And what would artificially cause an in-process transition to be masked? Just as was the case for the difference-measured transition, it would be nation Z's strengths vying primarily with the erstwhile loser's strengths.⁹

So, nation Z's entry will artificially cause a transition and its exit will artificially

avert a transition only if its scores across the six indicators are dichotomized such that the values for one subset of indicators are unusually high while the values for the remaining subset are unusually low.¹⁰ A bifurcation in Z's capabilities can also manifest two nuanced versions of artificial or masked transitions: 1) changing their *timing* by hastening or delaying them and 2) changing their *magnitude* by augmenting or ameliorating the amount by which the weaker overtakes the stronger. One example from major power data demonstrates both. Without the United States' entry into the major power system, Russia overtakes Germany in 1905. It does so with a *CINC* score that is 119.65 percent of Germany's. In 1907, Germany regains the lead. The scenario is quite different if the United States enters the major power system in 1898, as prescribed by the COW system membership rules. With the United States in the system, Russia still overtakes Germany in 1905, but this time with a much larger ratio over Germany's *CINC* score, namely 151.43 percent. In addition, Germany reestablishes its lead much later, not until 1916, or 11 years later than it would have had the United States not entered the system.

One caveat concerning the dichotomous capabilities culprit is in order. Such dichotomies among the major powers, a common focus group, should be relatively mild. Major powers are precisely those actors who are strong across many dimensions. If a nation has an uncommonly low score on one or more dimensions, it would be less likely to be included as a major power. For regional powers, however, this may not be unusual. Egypt, with high values for population and military personnel but low industrial scores, is a good example. The entry or exit of such a state would have a pronounced effect on the *CINC* values of Middle Eastern and North African states. Although we later demonstrate that entries and exits among the major powers do indeed produce transition measurement errors, such errors are likely to be more severe for regional groups of states.

An Alternative Composite Indicator: GINC

Unlike the arithmetic mean, the geometric mean does not affect dyadic power relationships when membership in the baseline group changes.¹¹ Therefore, we design *GINC*, a measure of nation I's power based on the geometric mean:

$$GINC_i = \sqrt[6]{\frac{ME_i}{ME_{total}} * \frac{MP_i}{MP_{total}} * \dots * \frac{TP_i}{TP_{total}}} \tag{14}$$

Using $GINC_i^*$ to represent I's score when Z is in the referent group, we consider how Z's entry affects the power distribution between X and Y:

$$\frac{GINC_x^*}{GINC_y^*} = \sqrt[6]{\frac{\frac{ME_x}{ME_x + ME_y + ME_z} * \frac{MP_x}{MP_x + MP_y + MP_z} * \dots * \frac{TP_x}{TP_x + TP_y + TP_z}}{\frac{ME_y}{ME_x + ME_y + ME_z} * \frac{MP_y}{MP_x + MP_y + MP_z} * \dots * \frac{TP_y}{TP_x + TP_y + TP_z}}} \tag{15}$$

This reduces to:

$$\sqrt[6]{\frac{ME_x * MP_x * \dots * TP_x}{ME_y * MP_y * \dots * TP_y}} \quad [16]$$

which is equivalent to the ratio of X and Y's *GINC* scores before Z enters the system. It can generally be shown that:

$$\frac{GINC_x^*}{GINC_y^*} = \frac{GINC_x}{GINC_y} \quad [17]$$

Because dyadic power relationships are unaffected by entries and exits when the geometric mean is used; artificial, masked, magnified, minimized, and mistimed transitions are impossible with *GINC*.

Although membership changes alter a state's *GINC* score ($GINC_i^* \neq GINC_i$), they do so in a predictable manner that is identical across all states. That is:

$$\frac{GINC_x}{GINC_x^*} = \frac{GINC_y}{GINC_y^*} = \frac{GINC_i}{GINC_i^*} \quad [18]$$

GINC maintains several of *CINC*'s advantages. First, it conveniently uses the COW capabilities data as inputs. Second, it offers a *relative* measure of national power. An additional advantage of the geometric mean is that it is conceptually appealing for measuring percentage changes (University of Toronto Mathematics Network, 1997), a useful feature for scholars interested in how power *shifts* affect conflict. Third, *GINC* reflects the performance of an individual state or a pair of states within the context of a relevant system. Each state's score conveniently translates as its share of group capabilities. For example, Britain's *GINC* score can be calculated as a share of global, regional, or major power capabilities. In settings where states can have "multiple enemies" (Bennett, 1996), it is conceptually appealing to make power share comparisons for each possible or politically relevant dyad. This approach, which assesses numerous dyads within a broader, sometimes global, structure, is taken by a variety of scholars investigating the connection between power distributions and hostility (e.g., Bennett, 1996; James, 2002; Kim, 2002; Organski and Kugler, 1980).

AN APPLIED COMPARISON OF MEASURES

Does the technique for measuring national power affect substantive results? In this section, we consider practical consequences for measuring dyadic power relationships and their relationship to militarized conflicts. For reasons outlined above, our focus is on the questions derived from power transition theory, but the procedure can be applied to other theoretical investigations as well.¹²

Data

We compare *GINC* and *CINC* in two universes of cases: the major powers and a special subset of them—the three or four most powerful—known as the “contenders” (Organski and Kugler, 1980, pp. 44–45).¹³ As explained above, major powers are the group *least* likely to manifest artificial or masked transitions produced by the traditional *CINC* indicator. Therefore, we have biased the analyses *against* confirming our mathematical demonstration of *CINC*’s problems. Within the two universes, we gather all cases of dyadic power transitions, which occur whenever one state’s power score becomes equal with another’s and then overtakes it by at least 10 percent. In other words, the stronger to weaker power ratio must be at least 1.1.

For war cases, we use level 5 militarized disputes from the dyadic version of the militarized interstate dispute (MID) dataset (Jones et al., 1996). Because some empirical investigations of power transition theory consider its application to lower levels of conflict as well (Kadera, 2001; Geller, 1993; Bueno de Mesquita et al., 1997), we also consider the occurrence of all levels of MIDs. Conflicts are considered to be associated with power transitions if they occur within five years of the transition. Occasionally, the resulting 10-year window encompasses another power transition between the same pair of major powers. In such cases, we average the date of the transitions, count them as a single transition, and look for associated conflicts in the span running from five years before the earliest transition to five years after the latest.

Power Transitions and the Occurrence of War

Because a central prediction of power transition theory is that power transitions are associated with wars, we compare how often each power measure results in transitions associated with war,¹⁴ as reported in Table 2. The difference between *GINC* and *CINC* is statistically significant for neither major power nor contender dyads.¹⁵

In the process of aggregating information to construct Table 2, we regularly noticed distinctions between a dyad’s record according to one measure versus that according to another. These differences, when aggregated, appear to wash out. In order to examine the differences at the dyadic level, we look at the magnitude of the difference between war rates for each dyad, and then calculate the average dyadic difference between the two indicators. For contenders, the difference (21.67 percent) is not statistically significant because it involves such a low number of cases ($n = 8$). However, the difference for major powers, 17.54 percent, is statistically significant

Table 2
Power Transitions and War

	Percent of power transitions associated with ≥ 1 war	
	<i>GINC</i>	<i>CINC</i>
Major powers	28.57%	22.58%
Contenders	20.0%	28.57%

Table 3
Power Transitions and Militarized Disputes

	Percent of power transitions associated with ≥ 1 MID		Rate of MIDs per power transition	
	<i>GINC</i>	<i>CINC</i>	<i>GINC</i>	<i>CINC</i>
Major powers*	60.61%	58.06%	1.69	1.52
Contenders	55.0%	64.29%	2.05	2.15

*Delaying Japan and Germany's exit from the major power set from 1939 until 1945 changes the proportions in this row only slightly and does not affect the conclusions.

at the .05 level.¹⁶ Thus, a sizable discrepancy in war rates results from the two methods of measuring national power and detecting dyadic power transitions. Switching from *CINC* to *GINC* can raise or lower a transitioning, major power dyad's war rate by an average of 17.54 percent. Because this value is significant, we conclude that changing from *CINC* to *GINC* results in drastically new estimates for war risks associated with dyadic power transitions. The pronounced difference at the dyadic level should lead researchers to be concerned about *CINC*'s effect on empirical findings when the unit of analysis is the dyad.

Power Transitions and the Occurrence of MIDs

Turning to an analysis of all MIDs associated with power transitions, we report the results for the probability that a transition will have at least 1 associated MID and the rate of MIDs per transition in Table 3. As in the systemic study of wars, none of the differences between values in Table 3 is statistically significant.

Strong dyadic differences, however, are found in Table 4. Discrepancies in the propensities for at least 1 MID are large and statistically significant, consistent with our casual observation of the dyadic data. We notice, for instance, that among the

Table 4
***CINC* vs. *GINC*: Average Dyadic Differences for MIDs**

	Propensity for ≥ 1 MID associated with power transitions	Rates of MIDs per power transition
	(<i>t</i> values)	(<i>t</i> values)
Major powers	32.32%*** (3.49) df = 22	0.89* (2.18) df = 19
Contenders	37.08%* (2.90) df = 7	1.25** (3.13) df = 7

* $p < .05$, ** $p < .02$, *** $p < .01$

Table 5
Timing of Conflict

	Average number of years from power transition*			
	All MIDs		War	
	<i>GINC</i>	<i>CINC</i>	<i>GINC</i>	<i>CINC</i>
Major powers**	.173	1.02	-.353	1.0
Contenders	.876	.706	-.16	.75

*When the 10 year window was adjusted to account for multiple power transitions near one another, the resulting window and associated conflicts were re-scaled to a 10 year base. Thus, the maximum distance from a power transition is ± 5 years.

**The ending year for Japan and Germany (1939 or 1945) makes no difference.

contenders, only 2 of the 8 dyads have the same probability of at least 1 associated MID for both methods of power transition detection. For both major powers and contenders, using *CINC* instead of *GINC* changes a dyadic power transition's propensity for at least 1 MID by over 30 percent. Dyadic sensitivity to power measures is even greater for all MIDs than it is for wars alone. Dyadic differences in the number of MIDs per power transition are also large and significant. Switching power indicators results in the average gain or loss of almost 1 MID per transitioning major power dyad. For the contenders, it produces an average shift of more than 1 MID per transitioning dyad. Confidence intervals for both estimates are well within acceptable limits. Clearly, the level of analysis has a great impact on the assessment of MID risks associated with transitions measured by the two indicators.

The Timing of Wars and MIDs vis-à-vis Power Transitions

Power transition theory specifies not only that power transitions make wars more likely, but also that these wars will occur preceding the transition (Organski, 1958, p. 333; Organski and Kugler, 1980, p. 60). How do *GINC* and *CINC* compare in terms of the timing of wars? Table 5 conveys information about the distribution of all MIDs and of wars within the five years before and the five years following transitions. Negative values indicate that conflicts typically precede the power transition, and positive values indicate that they tend to follow it. Differences between timing averages are not statistically significant for wars or MIDs. Even so, an important substantive difference between *CINC* and *GINC* results emerges from the war cases. *GINC*'s negative war timing results are consistent with power transition theory's expectation of conflict preceding the transition. In contrast, *CINC*-detected transitions are always associated with conflict afterward.

A dyad-by-dyad assessment of conflict timing is problematic because so few transition dyads experience conflict within the ± 5 year window for both power measures. Lack of militarized conflict makes it impossible to compare timing according to the two measures.¹⁷ For example, France and Italy have several MIDs associated with transitions detected using *CINC*, but no MIDs associated with transitions detected using *GINC*. Therefore, we cannot compare the timing of MIDs using *CINC*

with the timing of MIDs using *GINC*. In principle, there is actually an enormous difference in timing because *CINC* produces Franco-Italian transitions that experience MIDs 3.67 years afterward, whereas *GINC* produces transitions that *never* experience associated MIDs. We simply have no meaningful way of mathematically expressing this difference so that it can be compiled with other timing differences. At most, there are only a handful of existing comparisons that can be made between timing results, certainly not enough for statistical analysis.

FINAL RECOMMENDATIONS

Power plays a central role in international relations theory, and *CINC* is ubiquitously used to measure it. Mathematical reasoning and empirical application both demonstrate that shifts in group membership, common and seemingly innocuous phenomena, inadvertently affect *CINC*'s measurement of national power and dyadic power relationships. Specifically, errors in *CINC* can result in artificially revealed or masked power transitions, alterations in the magnitude of transitions, and shifts in their timing.

A single indicator, such as GDP, might seem like a useful alternative to *CINC*. Data limitations, such as the availability of annual GDP values only after World War II, often prevent scholars from using a single indicator. Or, a single indicator may not always be useful. For example, Saudi Arabian and Venezuelan oil give those states large GDPs without much correlation to their respective regional shares of power using *CINC* (Kadera and Sorokin, 1997; Taber, 1989). Theoretical focus might also require a composite indicator. A measure of power based only on military expenditures, for example, might be an excellent referent for the probability of winning a head-to-head combat, but a poor predictor of a nation's overall influence in world affairs. For these reasons, several scholars recommend *CINC* over single indicators (de Soysa et al., 1997; Sample, 1998; Senese, 1997).

The *GINC* indicator, which is based on the geometric rather than arithmetic mean, is a superior alternative. It allows researchers to retain the advantages of a composite indicator while avoiding the pitfalls inherent in *CINC*. *GINC* is immune to entry and exit effects on dyadic power relationships, maintains a systemically-based idea of *relative* power, and uses the readily available COW data. For major power dyads, replacing *CINC* with *GINC* changes a transitioning dyad's rate of war by over 17 percent; modifies a dyad's propensity for at least 1 MID by over 30 percent; and shifts the average transition time to *after* wars, which yields additional support for power transition theory. We therefore recommend *GINC*'s use, especially when dyadic power relationships are important to the study at hand and especially for analytical groups with frequent membership changes.

Embedded in this paper are many suggestions for further study. If *CINC*'s use matters for power transition theory, it might also matter for other power-centric theories and for studies that use power as a control variable. *CINC* might produce errors in additional power-based variables such as *CON*, a common indicator of the concentration of power in the international system (Singer et al., 1972); *security* (Bennett, 1997), which accounts for the relationship between a state's power level, those of its allies, and those of its enemies; or the *velocity* of capability change, which measures

the speed of power ratio shifts (Schampel, 1993). And if *CINC* errors make a difference for the major power system, the manifestations will be more pronounced for less stable groups such as regional sub-systems. Last, if *CINC*'s problems result from its functional form, then other composite indicators based on the arithmetic mean will exhibit similar pitfalls.

In sum, when *CINC* mismeasures power, the inferences we draw are, in part, artifacts of indicator construction. *GINC* eliminates this problem by providing a more reliable measure and less specious statistical tests of power-centric theories. Using *GINC* will advance our understanding of power's role in international relations.

NOTES

1. Tanisha Fazal kindly pointed this out to us.
2. The various types of state birth and death and their causes (e.g., revolution, annexation, conquest, unification, dissolution, regime collapse, and economic failure) are interesting subjects, but they are beyond the scope of this paper. Readers interested in these matters should consult the numerous sources that deal directly with them (e.g., Small and Singer, 1982; Gleditsch and Ward, 1999; Fazal, 2000; Bennett and Zitomersky, 1982).
3. A regional comparison group is another type of specialized subset. See, for example, Lemke's (1995) and Lemke and Werner's (1996) study of South American dyads or Kim's (2002) study of East Asian dyads.
4. Short-lived major power system configurations are found in the following periods: 1895–1898, 1918, 1919–1921, 1922–1924, 1941–1943, 1944, 1945, and 1946–1949.
5. It can be easily shown that although they can produce changes in the magnitude of a power difference, single indicators are immune to producing speciously revealed or hidden transitions. Nonetheless, depending on the question at hand, researchers sometimes prefer composite indicators due to other problems with single indicators, such as data availability and validity issues for certain states.
6. For example, *CINC* errors are not restricted to measurement of time-dependent phenomena.
7. For our purposes, we have greatly simplified power transition theory. In fact, the theory is more nuanced and includes additional elements such as states' satisfaction with the status quo. See Organski (1958), Organski and Kugler (1980), and Kugler and Lemke (1996) for richer expositions on power transition theory.
8. A straightforward dyadic comparison that ignores *Z*'s capabilities altogether would avoid this problem entirely. If a comparison group is theoretically irrelevant to the dyadic power distribution, and if there is no need to assess individual states' performances outside the dyad, we recommend that researchers simply use the head-to-head calculation.
9. Explicit mathematical proofs for ratio-detected transitions are available from the authors.
10. A state's entry into the referent group can cause transitions and its exit can mask transitions in a way unrelated to the *CINC* measure itself (see de Soysa et al., 1997).
11. In the mid 19th century, Jevons similarly argued that the geometric mean was best suited for isolating changes in prices due to shifts in the value of gold (1865, p. 122).
12. We leave for future research an analytic discussion of the econometric effects of predicting conflict with the error-prone *CINC* power measure. We suspect that such an analysis is complex because although some elements of the measurement error are systematic (predictable based on *CINC*'s functional form), others are random (e.g., based on the idiosyncrasies of coding schemes) or mimic randomness (systematic errors that sometimes mask transition and sometimes reveal artificial ones or sometimes hasten transitions and sometimes delay them). A fuller discussion of how these types of measurement error in explanatory variables produce inefficient estimates and weaken statistical relationships can be found in King et al.'s (1994) work.
13. According to COW rules, the major powers comprise the United States, 1898–1985; the United Kingdom, 1816–1985; France, 1816–1940 and 1945–1985; Germany, 1816–1918 and 1925–1945; Austria-Hungary, 1816–1918; Italy, 1860–1943; Russia/USSR, 1816–1917 and 1922–1985; China, 1950–

- 1985; and Japan, 1895–1945 (Singer and Small 1972; Small and Singer 1982).
14. No power transition is associated with more than 1 war.
 15. Difference of proportion tests produced low z-scores.
 16. $df = 22$, $t = 2.46$
 17. This was not a problem for the occurrence assessments because the non-entries were substantively equivalent to zeros. That is, conflict did not occur. We cannot make that inference here; a non-entry is not associated with any particular amount of time passing between a power transition and a conflict.

REFERENCES

- Bennett, D. Scott (1996). "Security, Bargaining, and the End of Interstate Rivalry." *International Studies Quarterly*, Vol. 42, pp. 157–183.
- Bennett, D. Scott (1997). "Testing Alternative Models of Alliance Duration, 1816–1984." *American Journal of Political Science*, Vol. 41, pp. 846–878.
- Bennett, Robert W., and Joseph Zitomersky (1982). "The Delimitation of International Diplomatic Systems: 1816–1970. The Correlates of War Project's Systems Reconsidered." In J. Zitomersky, ed., *On Making Use of History*, Lund: Esselt Studium, pp. 67–129.
- Bremer, Stuart (1992). "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816–1965." *Journal of Conflict Resolution*, Vol. 36, pp. 309–341.
- Bremer, Stuart (1993). "Democracy and Militarized Interstate Conflict, 1816–1995." *International Interactions*, Vol. 18, pp. 231–249.
- Bueno de Mesquita, Bruce (1981). *The War Trap*, New Haven: Yale University Press.
- Bueno de Mesquita, Bruce and David Lalman (1992). *War and Reason*, New Haven: Yale University Press.
- Bueno de Mesquita, Bruce, James Morrow, and Ethan Zorick (1997). "Capabilities, Perception, and Escalation." *American Political Science Review*, Vol. 91, pp. 15–27.
- de Soysa, Indra, John Oneal, and Yong-Hee Park (1997). "Testing Power-Transition Theory Using Alternative Measures of National Capabilities." *Journal of Conflict Resolution*, Vol. 41, pp. 519–528.
- DiCicco, Jonathan and Jack Levy (1999). "Power Shifts and Problem Shifts." *Journal of Conflict Resolution*, Vol. 43, pp. 675–704.
- Dixon, William (1994). "Democracy and the Peaceful Settlement of International Conflict." *American Political Science Review*, Vol. 88, pp. 14–32.
- Doran, Charles and Wes Parsons (1980). "War and the Cycle of Relative Power." *American Political Science Review*, Vol. 74, pp. 947–65.
- Fazal, Tanisha (2000). "When Is a State a State? Revising the Correlates of War List of Members of the International System." Unpublished paper, Stanford University.
- Geller, Daniel (1992). "Power and Conflict Initiation." *Conflict Management and Peace Science*, Vol. 12, pp. 1–16.
- Geller, Daniel (1993). "Power Differentials and War in Rival Dyads." *International Studies Quarterly*, Vol. 37, pp. 173–193.
- Gleditsch, Kristian and Michael Ward (1999). "A Revised List of Independent States Since the Congress of Vienna." *International Interactions*, Vol. 25, pp. 393–413.
- Goertz, Gary and Paul Diehl (1995). "The Initiation and Termination of Enduring Rivalries: The Impact of Political Shocks." *American Journal of Political Science*, Vol. 39, pp. 30–52.
- Houweling, Henk and Jan Siccama (1996). "A Two-Level Explanation of World War." In J. Kugler and D. Lemke, eds., *Parity and War: Evaluations and Extensions of the War Ledger*, Ann Arbor: University of Michigan Press, pp. 107–116.
- Huth, Paul, Christopher Gelpi, and D. Scott Bennett (1993). "The Escalation of Great Power

- Militarized Disputes: Testing Rational Deterrence Theory and Structural Realism." *American Political Science Review*, Vol. 87, pp. 609–623.
- James, Patrick (2002). *International Relations and Scientific Progress: Structural Realism Reconsidered*, Columbus, Ohio: The Ohio State University Press.
- Jevons, William Stanley (1865). "On the Variation of Prices and the Value of the Currency Since 1782." *Journal of the Royal Statistical Society*, Vol. 28, pp. 294–325.
- Jones, Daniel, Stuart Bremer, and J. David Singer (1996). "Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns." *Conflict Management and Peace Science*, Vol. 15, pp. 163–213.
- Kadera, Kelly M. (2001). *The Power-Conflict Story*, Ann Arbor: The University of Michigan Press.
- Kadera, Kelly M. (1997). "Measuring Capabilities in Regional and Global Rivalries." A paper presented at the Annual Meeting of the International Studies Association, March 18–22, Toronto, ON.
- Kim, Woosang (1989). "Power, Alliance, and Major Wars, 1816–1975." *Journal of Conflict Resolution*, Vol. 33, pp. 255–273.
- Kim, Woosang (2002). "Power Parity, Alliance, Dissatisfaction, and Wars in East Asia, 1860–1993." *Journal of Conflict Resolution*, Vol. 46, pp. 654–671.
- Kim, Woosang and James Morrow (1992). "When do Power Shifts Lead to War?" *American Journal of Political Science*, Vol. 36, pp. 896–922.
- King, Gary, Keohane, Robert, and Sydney Verba (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton: Princeton University Press, pp. 150–182.
- Kugler, Jacek and Douglas Lemke, eds. (1996). *Parity and War: Evaluations and Extensions of The War Ledger*, Ann Arbor: The University of Michigan Press.
- Lemke, Douglas (1995). "The Tyranny of Distance: Redefining Relevant Dyads." *International Interactions*, Vol. 21, pp. 23–38.
- Lemke, Douglas and Suzanne Werner. (1996). "Power, Parity, Commitment to Change, and War." *International Studies Quarterly*, Vol. 40, pp. 235–260.
- Mansfield, Edward (1992). "The Concentration of Capabilities and International Trade." *International Organization*, Vol. 46, pp. 731–64.
- Mansfield, Edward (1994). *Power, Trade, and War*, Princeton: Princeton University Press.
- Maoz, Zeev and Bruce Russett (1993). "Normative and Structural Causes of Democratic Peace, 1946–1986." *American Political Science Review*, Vol. 87, pp. 624–638.
- Mitchell, Sara M., Gates, Scott, and Håvard Hegre (1999). "Evolution in Democracy-War Dynamics." *Journal of Conflict Resolution*, Vol. 43, pp. 771–92.
- Organski, A. F. K. (1958). *World Politics*, New York: Alfred Knopf.
- Organski, A. F. K. and Jacek Kugler (1980). *The War Ledger*, Chicago: University of Chicago Press.
- Powell, Robert (1996). "Uncertainty, Shifting Power, and Appeasement." *The American Political Science Review*, Vol. 90, pp. 749–764.
- Powell, Robert (1999). *In the Shadow of Power*, Princeton: Princeton University Press.
- Sample, Susan (1998). "Military Buildups, War, and Realpolitik: A Multivariate Model." *Journal of Conflict Resolution*, Vol. 42, pp. 156–175.
- Schampel, James H. (1993). "Change in Material Capabilities and the Onset of War: A Dyadic Approach." *International Studies Quarterly*, Vol. 37, pp. 395–408.
- Senese, Paul (1997). "Costs and Demands: International Sources of Dispute Challenges and Reciprocation." *Journal of Conflict Resolution*, Vol. 41, pp. 407–427.
- Singer, J. David and Melvin Small (1972). *The Wages of War, 1816–1965: A Statistical Handbook*, New York: John Wiley and Sons.
- Singer, J. David, Stuart Bremer, and John Stuckey (1972). "Capability Distribution, Uncer-

- tainty, and Major Power War, 1820–1965.” In *Peace, War, and Numbers*, B. Russett, ed. Beverly Hills: Sage, pp. 19–48.
- Siverson, Randolph and Harvey Starr (1994). “Regime Change and the Restructuring of Alliances.” *American Journal of Political Science*, Vol. 38, pp. 145–161.
- Small, Melvin and J. David Singer (1982). *Resort to Arms: International and Civil Wars, 1816–1980*. Beverly Hills, CA: Sage.
- University of Toronto Mathematics Network, May 22, (1997). “Applications of the Geometric Mean.” <http://www.math.toronto.edu/mathnet/questionCorner/geomean/html>: Internet.
- Werner, Suzanne (1999). “Choosing Demands Strategically: The Distribution of Power, the Distribution of Benefits, and the Risk of Conflict.” *Journal of Conflict Resolution*, Vol. 43, pp. 705–726.