AN IDENTIFIED SYSTEMIC MODEL OF THE DEMOCRACY–PEACE NEXUS

PATRICK JAMESa,*, ERIC SOLBERGb and MURRAY WOLFSOnb

a Department of Political Science, Iowa State University;
b Department of Economics, California State University-Fullerton

(Received 1 August 1997; Revised 1 April 1998;
In final form 4 April 1998)

In previous empirical work, the basis for the proposition that democratic countries do not fight each other has been a single equation regression of hostility on democracy and other variables. This approach is misleading for two reasons. First, peace and democracy are part of a simultaneous system of relations in which they foster each other. Before quantitative inferences which affect policy conclusions can be reached, a separate structural equation has to be estimated corresponding to each of these variables. The equations must be distinguishable from one another at the same time as they embody the interdependence between peace and democracy. Second, previous regression studies emphasized the statistical association between democracy and peace rather than focusing on the substantive magnitude of that effect. To demonstrate the sensitivity of the results to these concerns, we accepted the data and indexes of one of the most influential statements of the peace–democracy thesis, and studied the Cold War period where it is deemed strongest. We differed from the prevailing paradigm by estimating an identified, simultaneous two-equation interactive system. In this more properly specified model, the dyadic democracy–peace nexus generally was not statistically significant and, more important, was very small in its impact. The alternative, peace causing democracy, was much stronger. Neither equation in the simultaneous system explained more than six per cent of the variance, so that other factors are likely to be much more important. Deterrence of aggression and patient negotiation of differences may be a more important guide to public policy than attempts to export western, democratic institutions to other nations.

The analysis of these data show that it is more likely that the most important differences that arise between nations are specific to historic epochs and

* Corresponding author.
their political and socio-economic conjunctures. Various other indices have been suggested for conflict and democracy in the literature, as well as numerous modifications of the additional variables to be included in the regression equations. The results have tended to be variable with respect to these modifications, but the failure of single equation model specifications to deal with the problem of simultaneous causation makes their results subject to our fundamental methodological criticism and unable to support the burden of policy recommendations.

KEY WORDS: Democracy; Peace; Empirical tests

METHODOLOGICAL INTRODUCTION

As a result of accumulated statistical investigations, the Kantian thesis that democratic nations do not go to war with each other is said to be as close to a law of nature as can be achieved in international relations (Levy, 1989). We will demonstrate how shortcomings in the method common to virtually all of these studies reduces the purported law to an unverified hypothesis.

First, all of the literature estimates single-equation models in which a measure of peace depends on an index of democracy along with other explanatory variables.\(^1\) When these studies find a positive and statistically significant coefficient, an unwarranted inference is made that democracy determines the likelihood of peace. Yet peace also influences the level of democracy. Stated negatively, war and preparations for war may diminish and even destroy democratic government (Lasswell, 1977: 167, 171; Almond, 1990: 269; Layne, 1994: 45; Gates, Knutsen and Moses, 1996: 5). Peace and democracy are likely to be part of a larger system and feed back on each other.

Second, these studies do not make it clear whether a structural or reduced form equation has been estimated. A structural equation in a larger system captures the feed-back effects of other endogenously determined variables. Reduced-form equations are regressions on the exogenous variables only, excluding all the endogenous variables determined within a system of equations. Reduced-form equations

may serve as predictors, but their coefficients do not correspond to the causal effects or underlying theory.

Third, since these studies do not take a systemic approach. They do not inquire whether the equation being estimated is actually the structural equation the authors believe themselves to be studying. Failing an identification procedure, it is simply unknown whether one is estimating the effect of democracy on peace, peace on democracy, or some mixture of the two influences. As a consequence, even when the coefficient estimates are statistically significant, there is no basis for assigning substantive political meaning to them.

Finally, the current literature tends to focus on the statistical significance of the associations in the form of t-scores or beta coefficients, rather than the substantive significance given by the magnitude of effects. Small and unimportant influences may be highly correlated with outcomes, and even be good predictors, without influencing them in an important way. We need to know the "oomph" of the explanatory variables (McCloskey 1993: 32–3), measured by their partial derivatives normalized into comparable units. Those studies that do measure magnitudes frequently fail to properly normalize them into comparable units; we review one such study in this paper.²

We offer two competing hypotheses to the democracy–peace theory, and make them operational by specifying falsifiable "prior beliefs" corresponding to each:

1) War is an expression of systemic interactions among evolving economic and social forces within and between nations. Democracy in the political structure is determined simultaneously with the other variables (James and Wolfson, 1993). Feedback mechanisms, however, require more than merely "controlling" for these variables within a single equation. Their interaction must be captured in the way the model is specified to identify the reciprocal influences of democracy

and peace. We anticipate that even a first step toward a systemic model specification will yield results significantly different from the single-equation studies.

(2) Since the complex of forces evolves historically, we expect all of them, including such political forms as democracy and autocracy, to be part of fairly distinct social structures. Consequently, the conjuncture of the influences in each structure is likely to be the dominant determinant of the overall character of an epoch. We expect that studies of the influence of the divergent political forms of dyads within a given epoch will reveal effects much smaller than those of the epochs themselves. We anticipate that intra-epochal regression studies of war will exhibit large constant terms, small explanatory coefficients even when statistically significant, and low overall explanatory power.

While our methodological criticism of the democracy–peace thesis applies to this literature in general, we illustrate its significance in terms of a representative and noteworthy exposition (Oneal, Oneal, Maoz and Russett, hereafter Oneal et al., 1996, see also Maoz and Russett, 1992; 1993; Oneal et al., 1994). Their data base is the Militarized Interstate Disputes (MID) compilation for the Cold War era (1950–1986). We accepted the indices they constructed from this base in that study, and, for the most part, their choice of variables.

These measures are not free from controversy. For instance, Polachek has long argued that conflict and cooperation should best be measured by the Cooperation and Peace Data Bank (COPDAB) rather than the MID series: he finds that his openness measure of trade then crowds out democracy as a statistically significant predictor of this variable (Polachek and Robst, 1998). Beck et al. (1997) point out that the data are pooled time-series and cross section dyad-years. As a result the observations – especially in the presence of many zero values for the dependent variable – are not independently and identically distributed and serial dependence becomes a serious stumbling block to valid inference. Their corrections weaken the results of Oneal and Russett (1997). The measure of dyadic democracy has likewise been the subject of criticism and suggested alternatives. Dixon (1993, 1994) introduced the weakest link in a dyad as a preferred measurement for

---

3 We are grateful to them for supplying us with the data and indexes used in their original conference paper (Oneal et al., 1994). It differs in some minor respects from the published version (Oneal et al., 1996), but in private correspondence Professor Oneal has replicated our results with his revision.
joint democracy. Subsequently, Oneal and Russett (1997) and other studies feature this approach which they claim results in a showing that both trade and democracy contribute toward peace. Gartzke (1998) suggests United Nations Affinity Scores as the proper measure of jointness of purpose, and crowds out democracy as measured by either the 1996 measure or the 1997 Dixon–Oneal–Russett indices. Indeed, using different methods, both Gleditsch and Ward (1997) and Ling and Wolfson (1997) have concluded that the underlying Polity III indexes of individual (monadic) levels of democracy do not reflect the usual connotation of freedom of expression, but rather the degree of constraint on the individual executive whether exerted by parliament or junta. Finally, Wolfson (1996) has explored the consequences of attempting to confer scalar attributes upon essentially vector valued functions.

These are all issues of the sensitivity of particular measures to the stability of forecast. None of those questions is relevant to the central thesis of this paper, which is to illustrate the over-riding principle that the methodological failure to consider simultaneous causation in all the current literature undermines its explanatory power. Consequently, we have not entered into a discussion of whether the index newly adopted by Oneal and Russett that engrosses Polacheck’s strictures to the relevance of trade, is or is not actually a superior measure. Under these circumstances it is best to go with the original. A comparative study of the sensitivity of outcomes to various indices within a simultaneous system is worth undertaking, but clearly beyond the scope of the present inquiry. In any case, where there were difficult statistical decisions to be made in pursuing this revised modeling and estimation of the Oneal et al. (1996) results, we chose those that most strongly supported the Kantian thesis.

We estimated a two-equation model as a minimal first step away from the single equation specification. Our first equation expresses the Kantian thesis: democracy is the causal independent variable and conflict the dependent one. The second equation expresses the possibility that the level of conflict is the causal independent variable, and

---

4 The minimal score in a dyad is regarded as the “lowest common denominator” for interactions and is expected to determine the level of discourse. However appealing this may seem as an alternative measure, the weakest link is problematic because it depends on the controversial assumption that norms within a regime are projected outward in the dealings of its leadership with other states. For competing views based on data analysis and case studies, respectively, see Dixon (1998) and Elman (1997).
democracy is its consequence. The other variables in these equations are assumed to be exogenous; they are derived from the literature on international relations and serve to identify the equations.

We do not claim to have produced a definitive, completely statistically unbiased explanation of peace and democracy. We doubt whether that is possible, but that issue need not be resolved to justify our skepticism (Manski, 1995: 3). A negative proposition may be demonstrated in one of three ways: show a logical contradiction; provide contrary evidence; or show that the results are not warranted by the original data when properly analyzed. We chose the third method. Therefore, this study is not a replication of Oneal et al. in the usual sense of running their model over again, except diagnostically to satisfy ourselves that we had it working properly. We resisted the temptation (and many suggestions of colleagues) to do things differently, as if we were conducting the investigation from scratch. That would have confounded the methodological issue. We found that when the present method is improved even partially by constructing a two-equation simultaneous system, the estimated magnitudes support our counter-hypotheses more than the Kantian thesis. These results may or may not carry through to alternative measures and indices, but that is a subject for further substantive inquiry.

Our agenda consists of five steps: (1) estimate a peace-causes-democracy equation as well as the Kantian democracy-causes-peace equation; (2) make the identifying restrictions explicit; (3) measure the magnitudes of the effects of democracy on conflict and conflict on democracy as well as statistical significance; (4) discuss the received literature in light of the statistical results; and (5) conclude with implications for public policy and further research.

WHY SHOULD PEACE LEAD TO DEMOCRACY?

The rationale for dyadic democracy leading to peace is well known. The normative argument derives from the ties of common culture among democratic states. The institutional argument reflects parliamentary political structure and accumulated skills of negotiation and compromise that might carry over into the international arena.

The justification for the converse proposition that peace induces democracy is best stated negatively; war undermines democracy. The fact is that a population in society sometimes does opt for war, from motives that range from self-aggrandizement to pure ethnic hatred
(Hirschleifer, 1987; 1995; James and Wolfson, 1993; Wolfson, 1995; Wolfson, James and Solberg, 1998). Even when they choose war through democratic institutions, they frequently unleash forces which tend to reduce the level of democracy.\footnote{Although an extensive review is beyond the scope of this study, research confirms the historically negative impact of war and preparations for it on democracy. Examples include Mintz (1985a: 632, 1985b: 12) on the military industrial complex in Israel: Martin (1994: 398), Segal (1994: 375) and Dandekar (1994: 357–58) on security-related needs and the diminution of the rule of law in France, the United States and United Kingdom; and Downing (1992: 239–40) on European history, most notably in relation to the destruction of nascent constitutionalism as a result of resource mobilization for protracted warfare.} War is an extreme case of a public good producing an external benefit to a significant number of individuals in society. Individuals are required to sacrifice their personal liberties of speech, movement, dress, consumption preferences, money income – not to speak of their lives – for the collective interest. Like all public goods, war cannot be provided by individuals independently maximizing only their utility without some sort of public compulsory action.

Compulsion is not necessarily conscription. It may take the more benign form of taxation to hire the requisite labor and capital even if the fighting is done by a mercenary army. Whether or not these acts are carried out with representation, they are species of the larger genus of compulsory appropriation of life, liberty and property for the public purpose of war. To implement that compulsion, a majority might well vote to give up democratic institutions in the sense of Gurr, Jaggers and Moore (1989) to suppress a dissident minority.

Thompson's (1996: 147) historical analysis puts the preceding argument in more general terms: "Some level of 'peace,' especially as indicated by the abandonment of expansionist foreign policies, made the historical development of liberal institutions and democratization possible." Stated in the strongest possible way, "It is not democracies that create the possibilities of peace, but peace that creates the possibility of democracy." Along similar lines, Midlarsky (1995) links autocracy and expansionism to environmental conditions in a model far more complex than the mere connection of peace to prior democracy.

The dynamic through which peace and democracy are linked is necessarily more complex than any representation that could be feasible in a statistical study. The rise of Hitler and the collapse of the Weimar Republic is a canonical example of external conflict undermining democratic institutions, reinforcing in turn, the drive toward
war. In *The Economic Consequences of the Peace*, John Maynard Keynes denounced as vindictive the peace imposed by the democratic victors at Versailles (Keynes, 1919). He saw it as setting the stage for German revanchism, and saw his forecast materialize as fascist militarism against which Britain would have to fight for its life. Whatever misgivings Lloyd George may have had about Clemenceau’s program to destroy the German economy were swept away by British public opinion which refused to “let the Hun off.” Keynes hardly concealed his contempt for the ineffectual Woodrow Wilson, whose program for the League of Nations was rebuffed by Congress and American public opinion. The postwar German economy exploded into hyper-inflation under the weight of reparations (Kindleberger, 1993: 281–334). The democratic government in Germany shattered into factions and appeared to be incapable of even elementary law and order. As a result, Hitler was elected to power as chancellor of the Reichstag parliament, and subsequently was able to destroy democracy by justifying war. Lebensraum was to be conquered to propagate those of the highest “race value” (Birken, 1991). Popular assent and participation in dictatorship and war of conquest was based on the perceived public good encapsulated in the slogan, “Ein Reich, ein Volk, ein Fuhrer.”

**PROBLEMS OF IDENTIFICATION AND BIAS IN SIMULTANEOUS SYSTEM ESTIMATION**

Single equation regressions cannot capture the reciprocal linkages between peace and democracy we have described. To demonstrate this, we first express the Kantian explanation as a stochastic, linear equation:

\[ P = a + bD + cM + u \]  

(1)

The constant intercept term is \( a \). \( D \) is multiplied by \( b \), which is presumed positive. \( M \) is a vector of other variables, and \( c \) is a vector of their coefficients. The random error term is \( u \). More generally, \( P = f(D, M, u) \), and the democracy partial derivative is positive, \( \partial P/\partial D > 0 \).

Now suppose that democracy is also a dependent variable, positively related to peace. Then \( D = g(P, M, v) \), where \( v \) is another random error and \( \partial D/\partial P > 0 \). Inverting this expression to display \( P \) as a dependent variable, we have \( P = h(D, M, v) \) and \( \partial P/\partial D > 0 \). Expressed as a
stochastic, linear equation:

\[ P = d + eD + fM + v \]  \hspace{1cm} (2)

The current literature regresses \( P \) on \( D \) and \( M \), and on that basis, claims overwhelming support for the neo-Kantian thesis. Does that regression capture Eq. (1) or (2)? Does it reflect neither (1) nor (2), but some combination of the two equations? The true test equation is not identified.

Ordinary least squares (OLS) estimation of either equation in this system may also be subject to simultaneous equation bias. OLS requires that the explanatory variables in each equation be uncorrelated with the error term. The systemic model shows that is not necessarily the case. If a change in the error \( u \) in equation (1) gives rise to a change in \( D \) through the effect of \( P \) on \( D \) in equation (2), the expected values of the least squares estimators of the parameters are not equal to their population values. Then least squares estimates are not even consistent; the bias is not mitigated by increasing the sample size.

As a result of either the failure to identify the equations in specifying the test model, or deal with simultaneous equations bias, it is entirely possible that the coefficients reported in the literature might differ from their actual values in size and even in sign. Even if they both are positive, policy makers need to know the main driving force between \( P \) and \( D \). If the coefficient in the Kantian equation (1) is large relative to Eq. (2), then the pursuit of democracy as a policy instrument is a reasonable way for countries to proceed to preserve the peace. If it is small, then the pursuit of democracy in other countries might be a laudable goal, but it is likely to be best brought about as a consequence of policies to prevent war. A patient search for solutions to economic and social problems, negotiation of differences, and balanced deterrence of aggression, might be better policies than trying to directly restructure the politics of other nations. Counter-productive outcomes might ensue even if the Kantian effect is positive but small relative to the negative effect that conflict has on democracy. It may well be true, as Dixon and Senese argue (Dixon, 1998; Dixon and Senese, 1997), that democracies are more effective than others in negotiating their differences, yet countries might well react in a hostile manner to the imposition of what others judge to be democracy. Wilsonian policies aimed at "making the world safe for democracy" might lead to a greater level of hostility and ultimately less democracy as well.
AN IDENTIFIED MODEL

As they stand, Eqs. (1) and (2) are not identified because they cannot be distinguished statistically from each other. They are observationally equivalent since either could be estimated by regressions of \( P \) on \( D \) and \( M \). Since they "look alike to the data" we need to invoke further information to tell them apart.\(^6\) A *necessary* condition for identification of a structural equation in a system is that the number of exogenous variables excluded from it must be at least as large as the number of endogenous variables included in it. When this *order condition* is satisfied, the equation has at least one solution. A *rank condition* is sufficient for identification. As a rule of thumb, the rank and order conditions are satisfied if each equation in the system has an exogenous (or predetermined) variable unique to it. That is the case in our operational realization of the theoretical two-equation model (1) and (2) as Eqs. (3) and (4). Henceforth, our notational practice will be to reserve single letter symbols such as \( P \) and \( D \) for theoretical Eqs. (1) and (2), and acronyms for statistical endogenous and exogenous variables in (3) and (4).\(^7\)

The Kantian hypothesis is represented operationally by Eq. (3):

\[
Hostility = f(\text{Regime, Growth, Proximity, Ally})
\]  

(3)

Peace is replaced by a converse measure of hostility, *Hostility*. Democracy is measured by the Maoz-Russett (1993) metric of joint democracy between dyads, *Regime*. The exogenous variable *Growth* is a measure of the joint economic growth shared between dyads (Polacheck, 1994). Exogenous variables *Proximity* and *Ally* represent geographical closeness and alliance status respectively (Small and Singer, 1976).

---

\(^6\) There are various ways to identify an equation within a system. First, some coefficients may be normalized. An example is when the coefficient on one endogenous variable is set equal to unity because it is the "natural" dependent variable. Second, there may be identities - variable definitions or known relations - that imply that all of the coefficients in a particular equation are known. Third, the exclusion of a variable from an equation may identify that equation: exclusion amounts to setting the coefficient of the omitted variable to zero. Fourth linear restrictions on the structural parameters may rule out false structures. Fifth, knowledge of disturbances and the restrictions on the disturbance covariance matrix work like restrictions on slope parameters. Sixth, nonlinear restrictions on the coefficients, though complicating the analysis, can aid in identification (Greene, 1993: 585–598).

\(^7\) These variables correspond to the MID data and indices supplied by Oneal and his colleagues as follows: *Hostility* = HIHOST. *Regime* = JOINREG. *Growth* = STABILITY. *Proximity* = GEO1. *Ally* = ALLY1 and *Stability* = JOINSTB.
The hypothesis that democracy is influenced by peace (operationalized by its converse, Hostility) is represented by Eq. (4):

\[ \text{Regime} = g(\text{Hostility, Growth, Stability}) \]  

(4)

Explanatory exogenous variables are chosen for each equation in accordance with widely accepted theory and serve to identify them. Stability is the exogenous variable representing the joint stability of regimes. Now these equations look different to the data and are identified. Proximity and Ally appear in (3) but not in (4); Stability appears in (4) but not in (3). As Proximity or Ally vary in (3), they trace out points on (4) and serve to identify it. Similarly the variation that Stability induces in (4) traces out points on (3) and makes it visible to the regression procedure.\(^8\)

DESCRIPTION OF THE EMPIRICAL VARIABLES

The observational unit for all variables is the dyad-year, the standard approach within the literature on the democratic peace. Hostility is the highest level of hostility reached by either member of a dyad in a year classified in the following categories: no dispute, threat of force, display of force, use of force, and war. Hostility is based on militarized interstate disputes data assembled from the Correlates of War project (Gochman and Maoz, 1984) and extended to 1986 (Maoz and Russett, 1993: 627). It refers to dyads which “plausibly” might engage in conflict; they are commonly referred to in the literature as “politically relevant”. These consist of contiguous dyads and major powers that are able to project military power far beyond their own borders. Because of “doubts about the ordinality” of the levels of conflict which they previously had included in their studies, Oneal et al. (1996) turned to a binomial measure, Dispute. After analyzing the outcomes of the multilevel Hostility classification, we shall compare our results with the dichotomous Dispute measure.

A histogram (Fig. 1) of the categories of Hostility strikingly displays the prevalence of peace during the Cold War epoch. Of the 16,322 continuing series of observations between 1950 and 1986, 15,830 are

\(^8\) The model of Hostility to be estimated is also identified below by its multinomial logistic specification, which differentiates it from the Regime equation that is linear in its parameters.
classified as peace. Peace constitutes about 97 percent of the already restricted set of dyads for which conflict is coded as plausible. The remaining cases consist of 24 instances of threat, 70 of display of force, 335 of use of force, and 63 observations of war. This highly skewed distribution, with so few cases of any conflict, sends a clear warning. Any statistical method that attempts to differentiate the
non-peace categories from one another will encounter difficulty because of relative and absolute scarcity of observations associated with each category.

*Regime* is a measure of joint democracy created by Maoz and Russett (1993). A regime index \( r \) for a single country is defined as
\[
r = c(d - a);
\]
c is an index of concentration of political power reflecting society’s ability to enforce its law, \( d \) is an index of democracy, and \( a \) is an index of autocracy (Gurr, Jaggers, and Moore, 1989). The indexes \( c, d, \) and \( a \) are each measured along an eleven point scale from zero to ten. The variable *Regime* is defined as:

\[
Regime = \frac{(r_H + r_L)(r_H - r_L + 1)}
\]

The regime scores for the two countries in a dyad are \( r_H \) and \( r_L \) with \( r_H > r_L \). Both \( r_H \) and \( r_L \) take on discrete values that may range from −100 to +100. As a result, *Regime* may take on discrete values that range from −200 to +200.

Figure 2 is a histogram of *Regime* using intervals of 20 units. The minimum observed score is −160, and the maximum is 100. Over 92 percent of the observations are clustered within the range −20 to +20. There are 8,290 cases in the range from −20 to zero, and there are 6,810 cases in the range 0 to +20. 80 percent of the cases lie between −4 and +10; there is a spike of 341 cases in the range from −80 to −60, and there is another spike of 443 cases between +80 and +100. This distribution conveys additional warnings: Since almost all the cases are clustered at the center with little variation, *Regime* may not be very useful as an explanatory variable. Moreover, any statistical model will be challenged to explain the variation in *Regime* as a dependent variable.

The remaining variables are assumed to be exogenous for the purposes of this study. Their means and standard deviations are reported in Table 4. *Stability* is a measure of joint stability of a dyad constructed from the longevity of a political regime for each country measured in years. *Stability* is computed in a similar fashion to *Regime* (Gurr, 1974; Gurr *et al.*, 1989). *Growth* is a measure of joint economic growth for a dyad, the average annual percentage change in real gross domestic product of the dyad over a three-year period (Summers and Heston, 1988). *Ally* is a dummy variable that equals unity if the pair of countries are formally allied with each other, or if both are allied with the United States; otherwise *Ally* equals zero. *Proximity* is a dummy variable that equals one if the pair of countries are either contiguous, or can extend military force indirectly beyond contiguous states (United States, United Kingdom, France, China, Soviet Union); otherwise, it equals zero.
SPECIFICATION AND ESTIMATION OF THE EMPIRICAL MODEL

Having shown that Eqs. (3) and (4) are identified, we turn to the second problem, minimizing simultaneous equation bias. The classic approach is to construct instrumental variables to replace the
endogenous variables that are linked to the error terms. An instrumental variable should be correlated with the endogenous variable it replaces, yet be uncorrelated with the error term. The appropriate technique for creating instrumental variables depends on the nature and number of variables and estimating equations.

If the endogenous variables were continuous values, two step least squares (2SLS) would be the natural method. Instruments would be created by first regressing each of the dependent variables on all the exogenous variables in both equations; then these first-stage reduced-form equations would be used to predict values of Hostility and Regime; finally, the predicted values would be used as instrumental variables in the structural equations to obtain consistent coefficient estimates. The precision of the estimates would increase with the number of observations as the distribution of each estimator converges about its expected value.

2SLS is not appropriate in light of the ordinal, categorical, nature of the Hostility variable, as well as its highly skewed distribution. Instead, we used a two-step procedure to estimate the Regime structural Eq. (4). First, we estimated a multinomial logit model of the categorical levels of Hostility as a function of all the exogenous variables. Second, this “reduced form” logit equation was used to generate instruments in the form of predicted probabilities of dyads being at each conflict level. These probabilities served as instruments for the ordinal categories to obtain consistent estimators.

Estimation of Hostility, Eq. (3), presented more formidable problems. Our first attempt was to deal with simultaneous equations bias by the two stage procedure: regress Regime against the exogenous variables; calculate its predicted values; then use those predicted values as an instrument in estimating the multinomial logit Eq. (3). This procedure yielded implausible results because the lack of variance in Regime meant that it could not be adequately explained by the exogenous variables in order to create a usable instrument. Fortunately, simultaneous equations bias turned out not to be an insurmountable problem. The estimation of (3) by maximum likelihood may be sufficient to provide consistent measures of its parameters without the use of instrumental variables if Regime is not contemporaneously correlated with the error term. We performed a necessarily crude test⁶, and

⁶There are better tests available when the dependent variable is continuous, but our dependent variable is categorical. To test the assumption of no contemporaneous correlation, we estimated the multinomial logit model with both Regime and its
satisfied ourselves that the conditions are likely to have been met, and estimated (3) using Regime itself as an explanatory variable.

We proceeded to carry out the estimations having satisfied ourselves that both equations were identified by the exclusion of variables method, and that the simultaneous equations bias was minimized within the modest two-equation limitation we imposed on ourselves.

ESTIMATING THE STRUCTURAL EQUATION FOR HOSTILITY

HOSTILITY is a categorical variable that requires special treatment. It is the observed frequency of ordinally ranked qualitative descriptions of levels of hostility. Therefore we use multinomial and sequential logit models to estimate the probability that each of these observed levels will be determined by the independent variables. Since the statistical package being applied requires category designations beginning with zero, we recoded HOSTILITY to Level = HOSTILITY - 1. Level = 0 if peace; Level = 1 if threat of force; Level = 2 if display of force; Level = 3 if use of force; Level = 4 if war.

Our estimation procedure encompassed three dyadic conflict scenarios:

1. Non-sequential: Conflict is not necessarily presumed to have passed through lower levels to reach a given intensity. We estimate the probability that dyads might simply be at any of the five hostility levels.

predicted value (obtained from a first-stage reduced form model) as explanatory variables to obtain the model's log-likelihood. A log-likelihood was also obtained from the model when the predicted value of Regime was excluded from the equation. A chi-squared test was then conducted to see if the inclusion of the predicted value made a significant difference in the log-likelihood. The computed value of chi-squared was 0.88, but the critical value of chi-squared at a five per cent level of significance with four degrees of freedom (the number of slopes set to zero) is 9.49. Regime may be used as its own instrument since it is not significantly correlated with the error; the null hypothesis cannot be rejected that they are not significantly correlated.

10 Our logit models were estimated using LIMDEP version 7.0 Econometric Software, Inc. (Greene, 1995; Greene, 1997: ch. 19). Probit specifications are just as plausible, but in practice the logit and probit alternatives usually yield similar results. The probit (and gompit) study by Gates and McLaughlin (1996: 16) comes to the same conclusion as we do: "Democracies are less likely to engage in conflict at the dyadic level, but the marginal impact in comparison to the other variables is tiny."
II. Escalation: Dyads move up to the next higher level of hostility having achieved a given level. Dyads need not have proceeded through a prior sequence of the preceding levels. At each level we estimate the probability that hostility might increase.

III. Sequential: We estimate the probability that dyads may be at any hostility level, having passed through all of the lower levels.

I. Non-sequential

The multinomial logit specification is appropriate for a non-sequential categorical variable. In four equations it estimates the logarithm of the likelihood of Hostility being at a given level of hostility compared to a condition of peace.¹¹ For each level of conflict \( j = 1, 2, 3, 4 \) higher than peace \( j = 0 \), the log-odds (the logarithm of the ratio of the probabilities) is expressed as a linear function of the explanatory variables:

\[
\ln[\text{Prob}(L = j)/\text{Prob}(L = 0)] = b_j x
\]

Since the underlying logistic equation is non-linear, these successive expectations will not simply be shifts of a regression line, but will vary in slope from one another.

The estimated coefficients are reported in Table 1. A chi-squared test strongly rejects the hypothesis that the levels of conflict are independent of the explanatory variables considered simultaneously. As others have found, there is an association between the probability of conflict and the independent variables. However, the degree of explanation provided by all of these variables taken together is low. As a measure of goodness of fit, the pseudo-R-squared (analogous to R-squared) indicates that only 5.9 percent of the variation in the probabilities of dyads being in the various categories has been explained. Evidently levels of conflict are largely determined by other forces which are impounded in the constant terms of the regression, or result from omitted variables and complex dynamic processes not specified.

¹¹ This is a non-linear specification in which the dependent variable is the probability that Hostility will be at any given level. Prob(L = j):

\[
\text{Prob}(L = 0) = 1/[1 + \Sigma \exp(b_j x) ] \quad \text{for} \ j = 0
\]

\[
\text{Prob}(L = j) = \exp(b_j x)/[1 + \Sigma \exp(b_j x) ] \quad \text{for} \ j = 1, 2, 3, 4
\]

where the \( b_j \) represent the right-hand-side of equation (3). There are different coefficients \( b_j \) to be estimated for each of four categories with respect to the base category \( L = 0 \); the base coefficients are set at \( b_0 = 0 \) for the purpose of estimation.
Table 1  Estimated coefficients of structural logit, log-odds of moving from peace to a higher level of conflict

<table>
<thead>
<tr>
<th>Variable</th>
<th>Peace to threat coefficient</th>
<th>Peace to display coefficient</th>
<th>Peace to use coefficient</th>
<th>Peace to war coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-7.0406</td>
<td>-6.0445</td>
<td>-4.5518</td>
<td>-5.7150</td>
</tr>
<tr>
<td></td>
<td>(-17.993)</td>
<td>(-25.276)</td>
<td>(-40.509)</td>
<td>(-26.837)</td>
</tr>
<tr>
<td>Regime</td>
<td>-0.0104</td>
<td>-0.0053</td>
<td>-0.0068</td>
<td>-0.0028</td>
</tr>
<tr>
<td></td>
<td>(-1.343)</td>
<td>(-1.076)</td>
<td>(-2.863)</td>
<td>(-0.465)</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.0179</td>
<td>-0.0373</td>
<td>-0.0078</td>
<td>0.0360</td>
</tr>
<tr>
<td></td>
<td>(-0.177)</td>
<td>(-0.617)</td>
<td>(-0.288)</td>
<td>(-0.0605)</td>
</tr>
<tr>
<td>Proximity</td>
<td>1.4441</td>
<td>1.6974</td>
<td>1.9721</td>
<td>1.0974</td>
</tr>
<tr>
<td></td>
<td>(+3.033)</td>
<td>(+6.036)</td>
<td>(+15.423)</td>
<td>(+4.039)</td>
</tr>
<tr>
<td>Ally</td>
<td>-0.3662</td>
<td>-0.4576</td>
<td>-0.9754</td>
<td>-1.4035</td>
</tr>
<tr>
<td></td>
<td>(-0.755)</td>
<td>(-1.652)</td>
<td>(-7.239)</td>
<td>(-3.631)</td>
</tr>
</tbody>
</table>

Note: t-ratios in parentheses; Log-likelihood = -2511.487; Restricted (Slopes = 0) log-L = -2674.622; Pseudo R-squared = 0.059; Chi-Squared (16) = 326.2708.

by the model, or simply reflect the dominance of random effects that swamp systematic explanation.

Within this frame work of weak explanation, the influence of democracy is even more questionable. For the log-odds of moving from peace to threat, ln(Prob₁/Prob₀), only Proximity has a coefficient statistically different from zero. Moving from peace to display, ln(Prob₂/Prob₀), only Proximity remains significant, and Ally becomes significant at the 10 percent level. Moving from peace to use of force, ln(Prob₃/Prob₀), the log-odds of the variable Regime does have a statistically significant negative coefficient, and both Proximity and Ally remain significant. Finally, for the log-odds of moving from peace to war, ln(Prob₄/Prob₀), Regime is no longer statistically significant, but Proximity and Ally remain significant explanatory variables. As the current literature suggests, the overall results indicate that on average an increase in democracy is related to a decrease in the probability of moving directly from peace to any higher level of conflict. Nevertheless, the effect is statistically insignificant for four of the levels of conflict, including the most important movement from peace to war.

Policy makers concerned with the magnitude of the influences on conflict need to know the marginal effects of each explanatory variable.
on the probability of *Hostility* being at a given level. The partial derivatives are evaluated at their means and reported in Table 2. Holding the other variables constant, they show how much the probability of being in each conflict level changes as a result of a unit increase in one of the explanatory variables from its average.\(^{12}\)

These marginal effects are stated in original units, but to interpret their magnitudes they must be converted into comparable units. The elasticity statistics in Table 2 are the ratio of the percentage change in the probability of being in a conflict category for a one percent change in an explanatory variable.\(^{13}\) As ratios of pure percentages, elasticities are dimension-free and permit comparisons regardless of original units (Cioffi-Revilla and Starr, 1995: 466–67). Elasticities are most useful for measuring *substantive* effects. For Eq. (3) they compare the percent change in the probability of *Hostility* being at a given level to the percent change in explanatory variables *Regime* or *Growth*.\(^{14}\) The elasticity measure also permits us to compare the influence of explanatory variables *between* equations. We shall thus be able to compare the percent change in *Regime* with percent changes in the various explanatory variables in Eq. (4).

We find statistical significance in the marginal effects only for the two categories with the largest number of observations: Peace and Use of Force. Substantively, we find that a one percent increase in *Regime* will increase the probability of peace, but by only one half of one thousandths of a percent. A one percent increase in *Regime* will decrease the probability of the use of force by 8 ten-thousandths of a percent, and will lower the probability of war by about 15 thousandths of a percent.

To compare these results with the dichotomous *Dispute* variable O'Neal *et al.* now prefer, we use a binomial logit model to estimate the

\(^{12}\) The partial derivatives of the multinomial logit model are given by:

\[
\frac{\partial P}{\partial x_i} = P_i [b_j - \Sigma_j P_j b_j].
\]

For any \(x_i\), \(\frac{\partial P}{\partial x_i}\) need not have the same sign as \(b_i\) (Greene, 1993: 664–672).

\(^{13}\) Elasticity is not an appropriate measure for categorical variables for which per cent changes are meaningless.

\(^{14}\) The elasticity may be estimated at any point of a regression equation and does not have to be estimated from the original data. For instance in the linear equation \(Z = a + bX + c1\), the elasticity of the effect of per cent changes in \(X\) on the per cent change in \(Z\) is \((\Delta Z/Z)/\Delta X/X\). The measure is thus equal to the slope of the regression in the \(XZ\) plane given by \(b\) multiplied by \(X/Z\) at any point. In general if \(Z = \ell(X, Y)\) the elasticity at each point is equal to \((\ell Z/\ell X)X/Z\). By convention, we evaluate the elasticity at the means of the variables.
Table 2  Elasticities and marginal effects of explanatory variables on the probabilities of hostility level from multinomial logit model

<table>
<thead>
<tr>
<th>Change in</th>
<th>Elasticity</th>
<th>Marginal effect</th>
<th>Std. error</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable = Prob(Peace):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regime</td>
<td>4.5865E-05</td>
<td>1.3480E-04</td>
<td>4.3548E-05</td>
<td>3.095</td>
</tr>
<tr>
<td>Growth</td>
<td>2.6110E-04</td>
<td>1.3494E-04</td>
<td>4.9449E-04</td>
<td>0.273</td>
</tr>
<tr>
<td>Proximity</td>
<td></td>
<td>-3.8702E-02</td>
<td>2.3660E-03</td>
<td>-16.358</td>
</tr>
<tr>
<td>Ally</td>
<td></td>
<td>2.0302E-02</td>
<td>2.5646E-03</td>
<td>7.916</td>
</tr>
<tr>
<td>Dependent variable = Prob(Threat):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regime</td>
<td>-2.6607E-03</td>
<td>-1.2094E-05</td>
<td>3.5936E-05</td>
<td>-0.337</td>
</tr>
<tr>
<td>Growth</td>
<td>-2.6379E-02</td>
<td>-2.1084E-05</td>
<td>1.5258E-04</td>
<td>-0.168</td>
</tr>
<tr>
<td>Proximity</td>
<td></td>
<td>1.6630E-03</td>
<td>9.8076E-03</td>
<td>0.170</td>
</tr>
<tr>
<td>Ally</td>
<td></td>
<td>-4.0894E-04</td>
<td>5.0901E-03</td>
<td>-0.080</td>
</tr>
<tr>
<td>Dependent variable = Prob(display):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regime</td>
<td>-1.2904E-03</td>
<td>-1.6814E-05</td>
<td>2.6997E-05</td>
<td>-0.623</td>
</tr>
<tr>
<td>Growth</td>
<td>-5.3124E-02</td>
<td>-1.2172E-04</td>
<td>1.9961E-04</td>
<td>-0.610</td>
</tr>
<tr>
<td>Proximity</td>
<td></td>
<td>5.4283E-03</td>
<td>6.3875E-03</td>
<td>0.850</td>
</tr>
<tr>
<td>Ally</td>
<td></td>
<td>-1.4303E-03</td>
<td>3.2850E-03</td>
<td>-0.435</td>
</tr>
<tr>
<td>Dependent variable = Prob(utility of force):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regime</td>
<td>-1.5646E-03</td>
<td>-9.7196E-05</td>
<td>3.6658E-05</td>
<td>-2.651</td>
</tr>
<tr>
<td>Growth</td>
<td>-1.0245E-02</td>
<td>-1.1191E-04</td>
<td>3.9020E-04</td>
<td>-0.287</td>
</tr>
<tr>
<td>Proximity</td>
<td></td>
<td>2.8108E-02</td>
<td>4.1288E-03</td>
<td>6.808</td>
</tr>
<tr>
<td>Ally</td>
<td></td>
<td>-1.3884E-02</td>
<td>2.6585E-03</td>
<td>-5.222</td>
</tr>
<tr>
<td>Dependent variable = Prob(war):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regime</td>
<td>-7.3592E-04</td>
<td>-8.6972E-06</td>
<td>2.8738E-05</td>
<td>-0.303</td>
</tr>
<tr>
<td>Growth</td>
<td>5.7634E-02</td>
<td>1.1977E-04</td>
<td>1.9654E-04</td>
<td>0.609</td>
</tr>
<tr>
<td>Proximity</td>
<td></td>
<td>3.5021E-03</td>
<td>5.9773E-03</td>
<td>0.586</td>
</tr>
<tr>
<td>Ally</td>
<td></td>
<td>-4.5778E-03</td>
<td>3.6792E-03</td>
<td>-1.244</td>
</tr>
</tbody>
</table>
joint probability of conflict, regardless of its intensity. The results are summarized in two Appendix tables and are consistent with the richer multinominal model. The effect of Regime on the probability of conflict, while negative and statistically significant \((t = 3.1)\) is tiny, whether measured by marginal effect or elasticity (Table A1). The effect of increasing the probability of conflict on Regime is also negative, but not statistically significant (Table A2). While the marginal effects and elasticities are small, the effect of Dispute on Regime is larger than the effect of Regime on Dispute. The former elasticity is \(-0.0122\), roughly ten times the size of the latter which is \(-0.00156\).

Actually, when normalized, the Oneal et al. (1996) results also show that the influence of democracy on peace is negligible. They compare the probability of a binary Dispute measure with a variation in the independent variable Regime by one standard deviation. The increase of Regime from its mean by one standard deviation \((25.8)\) reduces the probability of Dispute from 0.63 percent to 0.58 percent (Table II p. 20). The reduction in the probability of Dispute is 0.05 percent or as more customarily written, 0.0005. The mean probability of Dispute is 0.0271 and its standard deviation is 0.162 (p. 28). These statistics confirm our conclusions in three ways. First, the change in the probability of Dispute by 0.0005 requires a change in Regime by one standard deviation. Consequently, it would require a reduction in Regime by 20 standard deviations to increase the probability of Dispute by one percentage point, the 0.01 change from 0.063 to 0.0163. Clearly the probability of such an increase in any reasonable distribution of Regime is nil. Second, the elasticity measure applied to the 0.0005 increase in Dispute resulting from the 25.8 increase in Regime is \((0.0005/0.063)/(25.8/2.72) = 0.0007\). This small elasticity is of the same order of magnitude as our estimate. Third, the change in both variables may be normalized by dividing by their respective standard deviations. The resulting statistic is \((0.0005/0.162)/(25.8/25.8) = 0.003\). Each increase in Regime by one standard deviation reduces the probability of Dispute by 0.003 percent. Reducing the probability of Dispute by one standard deviation would require 333 standard deviations increase in Regime.

In view of the substantive insignificance of democracy in affecting the probability of peace, it is difficult to make a credible case that democracy is inversely related to the levels of hostility.\(^{15}\)

\(^{15}\)In a pioneering article, Senese (1997) estimated a single equation multivariate ordered logit model of the peace–democracy thesis and found even more strongly negative results. Joint democracy is significantly related to escalation of hostility in disputes.
Table 3  Calculated coefficients from structural logit, log-odds for moving to next-highest level of hostility

<table>
<thead>
<tr>
<th>Variable</th>
<th>Peace to threat coefficient</th>
<th>Threat to display coefficient</th>
<th>Display to use coefficient</th>
<th>Use to war coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-7.0406</td>
<td>0.9961</td>
<td>1.4927</td>
<td>-1.1632</td>
</tr>
<tr>
<td>Regime</td>
<td>-0.0104</td>
<td>0.0051</td>
<td>-0.0015</td>
<td>0.0041</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.0179</td>
<td>-0.0194</td>
<td>0.0295</td>
<td>0.0439</td>
</tr>
<tr>
<td>Proximity</td>
<td>1.4441</td>
<td>0.2533</td>
<td>0.2747</td>
<td>-0.8747</td>
</tr>
<tr>
<td>Ally</td>
<td>-0.3662</td>
<td>-0.0914</td>
<td>-0.5178</td>
<td>-0.4281</td>
</tr>
</tbody>
</table>

II. Escalation
Turning to increments in hostility, does greater democracy reduce the likelihood that conflicts will become worse? We calculate the log-odds that dyads will escalate one level of hostility to the next higher level from the differences in the parameters in the previous non-sequential model. The estimates for moving sequentially from one level to the next reported in Table 3 are derived from the non-sequential estimates.\textsuperscript{15} The influence of Regime on moving from peace ($L=0$) to threat ($L=1$) is negative but not statistically significant.\textsuperscript{17} The influence it has on the probability of moving from threat ($L=1$) to display of force ($L=2$) is positive but small and likewise statistically insignificant. The peace-democracy thesis is supported by the negative and marginally significant influence Regime has on moving from display of force ($L=2$) to use of force ($L=3$). However, the most important effect for this inquiry, the influence that Regime has on moving from use of force ($L=3$) to war ($L=4$), is positive and marginally significant. Increases in democracy appear to be consistent with an increased probability of the escalation of conflict into war.

Policy makers might hope to be guided by the most probable functional relationship between democracy and a given level of conflict. To estimate this relationship we first write out Eq. (5) explicitly. The

\textsuperscript{15} The log-odds that the conflict will escalate to the next higher one when they are one level apart is $\ln[P_j/P_{j-1}] = x(b_j - b_{j-1})$ when $j-k=1$. Estimates of $b_j$ and $b_k$ are drawn from Table 2.

\textsuperscript{17} The statistical significance depends on the precision of the estimated coefficients in Table 2. More formally, for two estimated coefficients $a$ and $b$ the variance of the difference is $\text{Var}(a-b) = \text{Var}(a) + \text{Var}(b) - 2 \text{Cov}(a,b)$. We did not compute these variances to find the standard errors and $t$-ratios for the parameter estimates of Table 3. We more
log-odds of dyads being at any level of hostility $j$ are:

$$\ln(\text{Prob}_j / \text{Prob}_0) = b_0 + b_1 \text{Regime} + b_2 \text{Growth} + b_3 \text{Proximity} + b_4 \text{Ally}$$

(5)

To associate $\text{Regime}$ with the movement of $\text{Hostility}$ from the base level to a higher level, we fix $\text{Growth}$, $\text{Proximity}$, and $\text{Ally}$ at their means. Then we replace the logarithm of the expected probabilities of the conflict levels on the left hand side with their observed frequencies, and solve for the value of $\text{Regime}$ associated with that ratio of probabilities. Said differently, we first generate a point estimate of the log-odds of levels of conflict at the observed frequencies of those levels, and we then reason backward to get the corresponding estimate of $\text{Regime}$. This is a great deal of leverage to place on the basic data, requiring that these calculations be evaluated with considerable caution.\textsuperscript{18} The results are shown in Fig. 3. In view of the importance of geographical contiguity (Vasquez, 1993: 123–52), we repeated the calculation for $\text{Proximity} = 0$ and $\text{Proximity} = 1$.

Figure 3 shows how unstable $\text{Regime}$ is as an influence on $\text{Hostility}$. For the middle case where $\text{Proximity}$ equals its mean, the only transition consistent with the Kantian thesis is the slight reduction in $\text{Regime}$ associated with the increase in $\text{Hostility}$ from Level 1 to Level 2. Thereafter there is little appreciable change in $\text{Regime}$ associated with the remaining higher levels of $\text{Hostility}$. This pattern suggests that, on average, democracy is directly related to increased hostility at lower levels of hostility, inversely related at intermediate levels of hostility, and not very important at the highest levels. When $\text{Proximity}$ is zero, referring to non-contiguous countries, the same pattern becomes more pronounced. When $\text{Proximity}$ is set at unity, the pattern reverses itself completely, and increased democracy appears consistent with higher levels of hostility.

If there is any consistency at all to the pattern, it is where the three curves cross. This is where the escalation from threat ($L = 1$) to display of force ($L = 2$) is associated with the neutral level of democracy where $\text{Regime}$ is close to zero. In view of the weak significance of the coefficients, and the unstable pattern of behavior they suggest, there is little basis for the assertion that democracy causes either increased or decreased change in levels of hostility.

\textsuperscript{18}Detailed worksheet calculations are available from the authors on request.
III. Sequential

Certain conflicts might best be described as arising from a progression of events, the outcome of a sequence of increasingly severe confrontations. We represent this process as an ordered logit expression of Eq. (3). The procedure estimates nested threshold parameters at which Hostility switches from one level to the next. Nested threshold values, $0 < \mu_1 < \mu_2 < \mu_3$ are estimated along with the structural coefficients, $b$. The independent variables, $x$, determine the five levels of conflict, as the variable Level takes on values $L = 0$ through $L = 4$. If $bx \leq 0$, $L = 0$; if $0 < bx \leq \mu_1$, $L = 1$; if $\mu_1 < bx \leq \mu_2$, $L = 2$; if $\mu_2 < bx \leq \mu_3$, $L = 3$; finally, if $\mu_3 < bx$, $L = 4$. The results are summarized in Table 4.
Table 4 Estimated ordered (sequential) logit model of hostility level

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-ratio</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>0.0528</td>
<td>4.892</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.2231</td>
<td>9.669</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>2.1052</td>
<td>17.481</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.0628</td>
<td>-44.421</td>
<td>1.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Regime</td>
<td>-0.0064</td>
<td>-2.125</td>
<td>0.33000</td>
<td>23.208000</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.0062</td>
<td>-0.258</td>
<td>1.876761</td>
<td>2.009659</td>
</tr>
<tr>
<td>Proximity</td>
<td>1.7830</td>
<td>17.322</td>
<td>0.321100</td>
<td>0.466902</td>
</tr>
<tr>
<td>Ally</td>
<td>-0.9054</td>
<td>-8.14</td>
<td>0.279380</td>
<td>0.448710</td>
</tr>
</tbody>
</table>

Pseudo R-squared = 0.055

Note: The mus are threshold values which, along with the constant, divide the distribution into five regions of conflict.

Responsiveness of Prob(Level) due to partial change in variable

<table>
<thead>
<tr>
<th>Category</th>
<th>Marginal effect</th>
<th>Elasticity</th>
<th>Change from 0 to 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regime*</td>
<td>Regime**</td>
<td>Proximity</td>
</tr>
<tr>
<td>Peace</td>
<td>1.42E-04</td>
<td>4.79E-05</td>
<td>-0.059660</td>
</tr>
<tr>
<td>Threat</td>
<td>-6.97E-06</td>
<td>-2.35E-06</td>
<td>-0.038639</td>
</tr>
<tr>
<td>Display of force</td>
<td>-2.03E-05</td>
<td>6.83E-06</td>
<td>0.0008283</td>
</tr>
<tr>
<td>Use of force</td>
<td>-9.65E-05</td>
<td>3.19E-05</td>
<td>0.040706</td>
</tr>
<tr>
<td>War</td>
<td>-1.80E-05</td>
<td>5.93E-06</td>
<td>0.007853</td>
</tr>
</tbody>
</table>

*Calculated for one unit increase above its mean.
**Calculated for a one per cent increase above its mean.
Note: Effects not calculated for growth since its coefficient is not statistically significant.

Ordered logit explains the observed levels of conflict by the independent variables with random error $\epsilon$. The probability of each level of conflict is determined with random variation by $x$ given the parameters $b$ and the threshold parameters $\mu$. 19 Figure 4 displays the probability that any level of hostility will pass sequentially through level $L=0$ to $L=1$, from $L=1$ to $L=2$, and so on to $L=4$. The area within each range is the probability of Level falling in that category. 20 Since about 97 percent of the dyads were classified as peace, the range

19 The whole distribution shifts at rates determined by $b$ as $x$ changes.
20 The probabilities can be calculated as a function of $x$ given the parameters $\mu$ and $b$: $\text{Prob}(L=0) = F(-bx)$; $\text{Prob}(L=1) = F(\mu_1 + \mu_2 - bx) - F(-bx)$; $\text{Prob}(L=2) = F(\mu_1 - bx)$; $\text{Prob}(L=3) = F(\mu_2 - bx)$; $\text{Prob}(L=4) = 1 - F(\mu_3 - bx)$. 
Figure 4  Probabilities for the ordered logit when Prob(\(L=0\)) is about 0.97.

\(L=0\) covers the first 97 percent of the distribution. This places a great burden on any estimation technique to discriminate among the other four categories crowded into the three percent right-hand tail of the distribution.

A chi-squared test indicates that Level is not independent of the explanatory variables taken together. As in the previous multinomial logit model, the pseudo-R-squared is low: only 5.5 percent of the variation in the actual categories is explained by the estimating equation.\(^{21}\) All the estimated coefficients are statistically significant except Growth. In particular, Regime is significantly negatively related to Hostility, reflecting the statistical correlation that has been noted in the literature.

The problem is to evaluate the substantive significance of the association between Regime and Hostility. Even though the error is frequently made, the coefficients of the ordered logit cannot be interpreted as the

\(^{21}\) We also estimated this ordered logit model with the reduced form predicted value of Regime added as an explanatory variable. Its estimated coefficient was not significantly different from zero, a result that is consistent with the use of Regime as its own instrumental variable.
marginal effects of changes in the regressors. This is so because the probability density function shifts as any one explanatory variable changes. For any given regressor, some of the partial effects must be positive and others negative. As the probability of being in any category increases, the sum of the probabilities of being in the other categories must decrease. Since the direction and magnitude of a change of any explanatory variable cannot be determined apart from its effect on the others, the sign of its marginal effect cannot be inferred simply by the sign of its coefficient.\(^{22}\)

The correct marginal effect of a change from its mean of Regime on the probability of Hostility being at any level is reported in Table 4. It also shows the elasticity of response to Regime, the percent change in the probability of being at each level of Hostility resulting from a one percent change in Regime.\(^{23}\) For all levels of conflict except the peace category, an increase in democracy does cause a decrease in the probability of moving to the next higher level. However, the predicted changes in probabilities are extremely small both in original units and elasticities. Since democracy as measured by Regime would have to increase between one hundred thousand to a million percent to bring about a one percent decrease in the probability of conflict reaching a given level of Hostility, its effect can be safely neglected for policy purposes.

ESTIMATING THE STRUCTURAL EQUATION FOR REGIME

Now that we have estimated the effect of democracy on conflict, we turn to the influence of conflict on democracy. The specification of Eq. (4) is influenced by the fact that Level is a categorical variable.

\(^{22}\) The correct partial derivatives are:

\[
\begin{align*}
\frac{\partial \text{Prob}(L = 0)}{\partial x} &= [(-\frac{1}{2}bx)\mu_1]b \\
\frac{\partial \text{Prob}(L = 1)}{\partial x} &= [(1-bx)-\frac{1}{2}(\mu_1-bx)]b \\
\frac{\partial \text{Prob}(L = 2)}{\partial x} &= [(\mu_1-bx)-\frac{1}{2}(\mu_2-bx)]b \\
\frac{\partial \text{Prob}(L = 3)}{\partial x} &= [(\mu_2-bx)-\frac{1}{2}(\mu_3-bx)]b \\
\frac{\partial \text{Prob}(L = 4)}{\partial x} &= [(\mu_3-bx)]b
\end{align*}
\]

\(^{23}\) Similar calculations were not performed for the other continuous explanatory variable, Growth, because its coefficient was not statistically significant. The effects of changing each dummy variable are calculated for an increase from zero to unity while all other variables equaled their means. Elasticities are not meaningful for these categorical variables.
Table 5  Estimated equation for regime

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-ratio</th>
<th>Marginal effect</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peace ($L = 0$)</td>
<td>2.00470E+01</td>
<td>15.952</td>
<td>19.4428</td>
<td>57.1441</td>
</tr>
<tr>
<td>Threat ($L = 1$)</td>
<td>9.24970E+03</td>
<td>11.644</td>
<td>13.6008</td>
<td>0.0618</td>
</tr>
<tr>
<td>Display ($L = 2$)</td>
<td>-1.14360E+04</td>
<td>-16.498</td>
<td>-49.0456</td>
<td>-0.6390</td>
</tr>
<tr>
<td>Use of force ($L = 3$)</td>
<td>2.16630E+03</td>
<td>17.177</td>
<td>44.4611</td>
<td>2.7620</td>
</tr>
<tr>
<td>War ($L = 4$)</td>
<td>-6.52120E+03</td>
<td>-16.259</td>
<td>-25.1705</td>
<td>-0.2975</td>
</tr>
<tr>
<td>Other variables:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>2.22660E+01</td>
<td>1.276</td>
<td>-0.0646</td>
<td>-0.0052</td>
</tr>
<tr>
<td>Stability</td>
<td>-1.13840E+00</td>
<td>-8.066</td>
<td>-1.2866</td>
<td>-1.9632</td>
</tr>
<tr>
<td>Growth-squared</td>
<td>3.23560E-02</td>
<td>1.569</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth x Stability</td>
<td>-1.33270E-01</td>
<td>-4.192</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stability-squared</td>
<td>1.66090E-02</td>
<td>6.821</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adj. R-squared = 0.0453.

Note: Standard errors corrected to heteroskedasticity.

measuring an ordinal ranking of five distinct levels of hostility. In our two-step procedure we use as regressor instruments the predicted probability of Hostility being at each level based on a reduced-form logit estimation against all the exogenous variables in the system. These predicted probabilities are highly correlated with the dummy variables, but not with the error term.\(^{24}\) Equation (4a) is the empirical expression of Eq. (4):

\[
\text{Regime} = b_0 \text{Prob}(L = 0) + b_1 \text{Prob}(L = 1) + b_2 \text{Prob}(L = 2)
\]
\[
+ b_3 \text{Prob}(L = 3) + b_4 \text{Prob}(L = 4) + b_5 \text{Growth} + b_6 \text{Stability}
\]
\[
+ b_7 (\text{Growth})^2 + b_8 (\text{Growth} \times \text{Stability}) + b_9 (\text{Stability})^2 \quad (4a)
\]

The results are summarized in Table 5.\(^{25}\)

\(^{24}\) The independence of predicted probabilities from the error was tested by regressing the predicted error term from the model that used dummy variables on the predicted probabilities obtained from the reduced form logit model. The resulting t-values were uniformly small.

\(^{25}\) To make sure that we did not exaggerate the impact of the probability of levels of conflict on democracy, we attempted to capture non-linear effects of the competing variables by including squared terms for Growth and Stability in the regression equation. We also added an interaction term Growth times Stability to allow for the presumed relationship between economic growth and democratic political institutions that might exist in market economies. These other variables also help to identify the equation.
Figure 5  A crude 95% interval for regime at each hostility level

The adjusted R-squared is low: less than five percent of the variation in the dependent variable has been explained by regression. Figure 5 presents "crude" 95 percent confidence intervals for predicted values of Regime being at each level of conflict.²⁶ As they stand, all but one interval overlaps the value zero. Nevertheless, most of the estimated coefficients, marginal effects and elasticities are statistically significant.

The direction of influence was somewhat irregular, but the polar cases were in striking comparison to the negligible effect of democracy on peace. Even the smallest elasticity of the effect of peace on democracy was ten-thousand times as large as the largest effect of democracy on peace. The largest elasticity showed that a one percent increase in

²⁶ Each midpoint was calculated by fixing the values of the exogenous variables at their means and weighting respective level coefficients by their frequencies. The intervals are constructed by adding and subtracting two standard errors of regression to each midpoint while also perturbing the respective level coefficients by ±2 of the coefficient's standard errors before weighting it by its frequency. We call this a crude 95 per cent interval because we did not consider variation in the coefficients in the exogenous variables. Thus an actual 95 per cent interval would be larger for each category.
the probability of being at peace (Prob $L=0$) increased democracy by more than 57 percent. In contrast, a one percent increase in the probability of a dyad being in a state of war decreased democracy by approximately 0.3 percent. The anomalous case was the large 2.8 percent increase in democracy resulting from an increase in the probability of the use of force. Nevertheless, the comparison with the negative war elasticity is striking. Evidently the use of force short of war has served to reinforce joint political democracy in the Cold War era, but as we anticipated, the outbreak of war served to undermine it. In contrast to this sizable substantive effect of peace on democracy, it takes between one hundred thousand to a million percent change in democracy to effect a one percent change in the probability of a level of hostility.

OVERALL EXPLANATORY POWER OF THE DEMOCRACY–PEACE THESIS AND FUTURE DIRECTIONS

Contrary to the received literature, little of the preservation of dyadic peace can be explained by the influences of democracy. The main reason for this outcome is evident from the data itself. Put plainly, there was very little dyadic hostile action to be explained.\footnote{The arbitrary dyadic level of aggregation may distort the impact of peace and democracy on each other. Levels of both peace and democracy between dyads may be directly or indirectly influenced by their conflicts or cultural or economic interactions with other states. A more flexible general system of aggregation is required and is the subject of further research (Wolfson, 1992: 60–80, 1996).} Overwhelmingly, the best predictors of the level of Hostility are the constant term in the regression, and the dummy variables Proximity and Ally that serve to shift the intercept. Since the coefficients for democracy are small even when statistically significant, the best predictor of the level of hostility is simply its mean value over the time period under review.

Our two general hypotheses have been borne out by this investigation. First, war is an expression of systemic interactions between and within nations. When we treat the internal political regimes of the nations of the world system as endogenous, and when the marginal effects are measured appropriately, our results differ markedly from the received single equation literature. The view that democracy causes peace is not strongly supported by the empirical results. There is somewhat stronger support for the inference that peace affects democracy.
Second, our results are consistent with the hypothesis that hostility is likely to vary more widely *between* historical epochs than *within* them. To be sure we have not undertaken a comparative study of different epochs which might demonstrate that view more conclusively than we have attempted here. Nevertheless, the fact that the constant terms in our equation turn out to be the dominant elements in explaining conflict rather than the variation of the variables particular to the era under study, suggests that other forces are at work which can be expected to vary across wider time spans and social conjunctures.

Our fundamental theme has been methodological. We showed that the addition of only one additional plausible relation to the single equation research design of the dominant literature results in strikingly different conclusions. We did not allow that simple message to be diluted by investigating the alternative data and indices. These considerations are under ongoing review by numerous researchers. Our substantive results cannot be more reliable than the data from which they are constructed and must therefore be regarded as more tentative than our methodological conclusion. Greater realism may be achieved with better data, more satisfactory means of aggregation into indices, and without doubt the construction and estimation of larger, more inclusive systemic models of simultaneous international interaction. Nevertheless, to the extent that this study can claim to reflect reality as well as method, it suggests three policy consequences:

First, the inability of either equation to explain much of the variation in the dependent variables makes it clear that, even when statistically significant, the contribution to peace of policies promoting democracy is likely to be swamped by other forces. Therefore, policy should be based on the dynamic interaction of peace and democracy with each other, with variables not yet properly studied, and with the larger epochal picture in which they are embedded. We need to expect the unexpected and avoid dogmatism in an evolving, complex international environment.

Second, there might be a trade-off between promoting peace or promoting democracy as rival uses of scarce political capabilities and economic resources. We should not expect that the rate of trade-off will be constant either relative to each other, over time, or between epochs. The apparent greater effect of peace on democracy which we have suggested in the Cold War era is not a law of nature any more than the converse proposition. The effects may not always be additive or even mutually consistent. Woodrow Wilson's proposition that
World War I was a “war to end all wars” and also a war to “make the world safe for democracy” failed on both counts.

Third, policy makers need to focus their attention on the underlying complex conjuncture of causes of conflict as well as the institutions through which it is implemented. The Cold War conjuncture has changed, and new policies need to emerge to preserve the peace.28 Our results illustrate the necessity of facing these new situations by studying systemic conjunctures of mutual causation. Perhaps the extended debate between structural realism and its opponents might yield to a more fruitful discussion about which variables are usefully treated as endogenous or exogenous in an identified, interactive system (Waltz, 1986; Keohane, 1986; Grieco, Powell and Snidal, 1993; Polacheck, 1994; Midlarsky, 1995).

Despite the apparent small causal, substantive effect of dyadic democracy on peace, the statistically significant correlations between them makes the loss of democracy a warning symptom of a systemic war syndrome. The appearance of authoritarian regimes puts us much in the same position as the physician whose patient presents with a spiked fever which might result from many causes. We are on notice that something is going wrong. To be sure, the fever itself can be so life-threatening that it needs to be reduced, even at the risk of masking other symptoms or even aggravating the illness. Without a protocol for treatment and cure, that may be the best that the physician can offer, but in general, palliative treatment without a differential diagnosis is not enough. Like many medical pathologies, the war syndrome does not always derive from a single pathogen, but may be the combination of several interacting causes. These components may not be simply summable as contributing to the total outcome as if they were multiple variables in a single equation model of disease, but may interact in complex and as yet inadequately understood ways. Then we must face the limitations of our knowledge. Peace and democracy are features of much more complex systems that have yet to be adequately modeled and measured, but if we have contributed anything at all in this study it is our suggestion that our preferred method of treatment should derive from understanding the war syndrome as well as respond to one of its symptoms.

28 We believe the restriction of the theory to “interstate disputes” to be unduly narrow in that it disallows civil wars from the data set. Somewhat artificially, as these warring ethnic groups claim national independence (as in the former Soviet Union) they will fall under the rubric of “militarized interstate disputes” while formerly they would be excluded.
REFERENCES


APPENDIX

Table A1  Estimated binomial logit model of conflict (level > 0)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>$t$-ratio</th>
<th>Marginal effect</th>
<th>$t$-ratio</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-4.0633E+00</td>
<td>-45.415</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regime</td>
<td>-6.3730E-03</td>
<td>-3.213</td>
<td>-1.4214E-04</td>
<td>-3.238</td>
<td>-1.5583E-03</td>
</tr>
<tr>
<td>Growth</td>
<td>-6.6780E-03</td>
<td>-0.295</td>
<td>-1.4895E-04</td>
<td>0.295</td>
<td>-9.2883E-03</td>
</tr>
<tr>
<td>Proximity</td>
<td>1.7842E+00</td>
<td>17.082</td>
<td>3.9796E-02</td>
<td>19.162</td>
<td></td>
</tr>
<tr>
<td>Ally</td>
<td>-9.0143E-01</td>
<td>-7.986</td>
<td>-2.0106E-02</td>
<td>-8.026</td>
<td></td>
</tr>
</tbody>
</table>

Pseudo R-squared = 0.0691.

Table A2  Alternative estimated equation for regime

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>$t$-ratio</th>
<th>Marginal effect</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflict ($L &gt; 0$)</td>
<td>-4.42060E+00</td>
<td>-0.766</td>
<td>-0.1332</td>
<td>-0.0122</td>
</tr>
<tr>
<td>Other variables:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-5.71730E-02</td>
<td>-0.139</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>1.01380E+00</td>
<td>5.624</td>
<td>0.4270</td>
<td>0.0345</td>
</tr>
<tr>
<td>Stability</td>
<td>-2.52600E-01</td>
<td>-2.157</td>
<td>-0.4171</td>
<td>-0.6363</td>
</tr>
<tr>
<td>$(Growth)^2$</td>
<td>-4.88150E-02</td>
<td>-2.382</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth $\times$ Stability</td>
<td>-1.31610E-01</td>
<td>-3.775</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(Stability)^2$</td>
<td>1.34590E-02</td>
<td>5.633</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adj. R-squared = 0.0074.

Note: Standard errors corrected for heteroskedasticity.
COMMENT: WHY "AN IDENTIFIED SYSTEMIC MODEL OF THE DEMOCRACY-PEACE NEXUS" DOES NOT PERSUADE*

JOHN R. ONEAL‡,§ and BRUCE RUSSETT¶

‡Professor of Political Science, Political Science Department, University of Alabama, Tuscaloosa, AL 35487-0213
¶Dean Acheson Professor of International Relations, Political Science Department, Yale University, New Haven, CT 06520-8301

(Received 19 April 1999; Revised 15 October 1999; In final form 15 October 1999)

In their article in this journal, James, Solberg and Wolfson (1999) challenge our findings that two states are more likely to have peaceful relations if they are both democratic. They claim to develop a simultaneous system of two equations showing that peace and democracy foster each other, and that the effect of peace in encouraging democracy is stronger than that of democracy on peace. Their analysis, however, is flawed. Their research design employs measures of dispute and joint democracy that are inferior to those now common in the literature, and their equation for predicting peace is not properly specified. These problems distort their results. Even so, their results provide evidence of the pacific benefits of democracy. Analyses we conduct with a more completely specified model reveal stronger support for the democratic peace. Furthermore, a test of the effect of interstate conflict on democracy should be done at the national (or monadic) level of analysis; but James et al. perform a dyadic analysis. In a monadic test using vector autoregression, we find that disputes make no contribution to explaining the character of regimes. Even with their dyadic method, their finding that peace promotes democracy is not robust. Including a crucial control variable, the ratio of militarily relevant national capabilities, that James et al. omitted, dramatically alters their findings.

KEY WORDS: Democracy; Democratic peace; War

*We thank Birger Heldt, Zev Maoz and Allan Stam for comments, and the National Science Foundation and the World Society Foundation of Switzerland for financial support.
†Corresponding author. E-mail: joneal@tenhoo.as.ua.edu.
James, Solberg, and Wolfson (1999) begin with the premise that peace and democracy foster each other. They argue that all previous research supporting the democratic peace makes an unwarranted inference because the influence of international conflict on the character of political regimes has not been simultaneously estimated. They seek to show the implications of this error by reconsidering two of our published studies (Maoz and Russett, 1993; Oneal, Oneal, Maoz and Russett, 1996). The possibility of a reciprocal relation between conflict and the character of government is an important question, as Russett (1993, p. 137) has noted; but the results the authors report do not show that our previous analyses of the democratic peace are flawed. Rather, they re-affirm our findings. Further, vector-autoregressive analyses at the national level indicate that militarized interstate disputes do not have a significant effect on regime type.

The authors' central claim is that the democratic peace must be tested within a simultaneous system of equations. They report, however, that their first effort to use a two-stage, instrumental-variable procedure to identify the influence of joint democracy on the likelihood of conflict produced "implausible results" (p. 15). Using the test described in note 9 (p. 16), James et al. satisfied themselves that there is no contemporaneous correlation between Regime and the error term in their Equation 3 and concluded that, in the jargon, "Regime may be used as its own instrument." Consequently, they assessed the consequences of democracy for interstate relations "using Regime itself as an explanatory variable" (p. 16) in the same straightforward manner as we and others have done previously. In short, a two-stage procedure of estimation is not needed and was not used.\(^1\) Thus, the methods we have used previously are valid, and our conclusions stand: There was a separate peace among democracies in the cold war era (Maoz and Russett, 1993; Oneal, Oneal, Maoz and Russett, 1996; Oneal and Russett, 1997; Oneal and Russett, 1999a) and in earlier periods (Oneal and Russett, 1999b; Oneal and Russett, 1999c). Indeed, democracies were more peaceful, ceteris paribus, at the national (or monadic) level of analysis, too (Oneal and Russett, 1997).

The authors' research suffers from three fundamental problems:

1. They offer various tests of the democratic peace, using our data, but do so in ways that depart crucially from anything we have published.

\(^1\)Despite this, the authors claim in their conclusion that they "treat the internal regimes of the nations of the world system as endogenous" (p. 30).
While admitting that their "study is not a replication of Oneal et al." (p. 6), they assert that, "where there were difficult statistical decisions to be made...we chose those that most strongly supported the Kantian thesis" (p. 2). But in reality, their departures from our methods stack the deck against the liberal position. Their tests were destined to fail. Even so, they support our results better than we might have expected!

2. When the authors find evidence of the democratic peace, they claim to show that these benefits are inconsequential. In Oneal et al. (1996) and subsequent reports (Oneal and Russett, 1997; 1999a; 1999c; Oneal and Ray, 1997; Russett, Oneal and Davis, 1998), we provide estimates of the probability that a dyad (or pair of states) would be involved in a militarized dispute under various conditions. These show in simple, easily understood terms that democracies are much more pacific than other states. The authors' alternative calculations are convoluted and misleading.

3. The alternative thesis they propose – that the effects of conflict on the character of regimes are greater than the peace-promoting benefits of democracy – is not properly evaluated. There are critical disjunctures between their thesis and their test. War may weaken democratic institutions, as Lasswell (1977) feared; but James et al. assess this hypothesis regarding individual nations with dyadic data on militarized disputes, only a small fraction of which are wars. Moreover, the results they report are extremely sensitive to alterations in the specification of the test; and a more plausible two-equation system produces results that are diametrically contrary to those they report.

TESTS DESTINED TO FAIL

The analyses that James et al. undertook were unlikely to provide strong support for our findings because they use the militarized disputes data inappropriately, rely on an out-dated measure of joint democracy, and do not include a control for relative power, the central element of realist theory:

(a) The authors assess the liberal thesis using measures of the nature of regimes – both for the individual states, and more importantly, in the computation of a joint regime score for each dyad – that were superseded more than three years ago. As the authors note, our Regime variable was criticized for reasons we have discussed in detail (Oneal and Russett, 1997; Oneal and Ray, 1997). Consequently, we adopted new specifications that are better informed theoretically than
those originally used by Maoz and Russett (1993). Tests using these new specifications provide unambiguous support for the pacific benefits of democracy (Oneal and Russett, 1997; 1999a; 1999b; 1999c). We shared the earliest of these results in manuscript form with the authors in May 1996, along with a commentary on an earlier version of their article. They chose not to use these improved specifications. It is hard to see how this gives the Kantian thesis a fair test.

(b) The authors consider three alternative scenarios of dyadic conflict, the non-sequential, escalatory, and sequential. The first two are assessed using multinomial logistic regression analysis, in which the probability of a dyad being at one of five levels of Hostility (peace; threat, display, or use of force; or war) is estimated. Multinomial analyses bias the test against the Kantian thesis because of the rarity of conflict: There are only 24 threats, 70 displays of force, 335 uses, and 63 observations of war in the data they used; 97% of the 16,322 cases (a pair of states in a year) are peaceful. This skewness makes it very likely a priori that an estimation procedure utilizing the individual levels of Hostility will produce insignificant results.

The authors themselves note the problem. "Any statistical method," they say on pp. 12–13, "that attempts to differentiate the non-peace categories from one another will encounter difficulty because of [the] relative and absolute scarcity of observations associated with each category." When reporting their results, they urge caution because of the "great burden on any estimation technique" of trying to discriminate among the four categories "crowded into the three percent right-hand tail of the distribution" (p. 26). Indeed, James et al. conclude that the reason for their weak results "is evident from the data itself. Put plainly, there was very little hostility to be explained" (p. 30).² There are indeed few enough disputes without subdividing them. Why do it? To try to explain the causes of 24 threats among more than 16,000 cases (0.15% of the total) seems pointless. Tellingly, they do find statistically significant marginal effects "for the two categories with the largest number of observations: Peace and Use of Force" (p. 19, emphasis added).

The alternative to the method adopted by James et al. is to use a binomial indicator of dispute involvement, which takes a value of 1 if

²Using the latest Correlates of War data (Bremer, 1996), there are 947 dyad-years of disputes among the politically relevant dyads, 1950–85, almost twice as many as in our older data set. We sent our new data to the authors in November 1997, but they have continued to use the old data. Our new data are available at http://bama.ua.edu/~joneal/jpr_data or www.yale.edu/unsy/democ.htm.
there was a threat, display, use of force, or a war; if peace, this variable equals 0. This is the approach we have relied upon in all our work.\(^3\) The use of a dichotomous variable is preferable not only because of the paucity of cases at particular levels of Hostility but because of serious doubts about the ordinality of this variable, as James et al. note (p. 11). Stuart Bremer, the principal compiler of the revised militarized disputes data, has said that many uses of force involve the unopposed seizure of a civilian fishing vessel. Such relatively minor events are coded 3 on the 0–4 Hostility ranking, whereas a threat of general or even nuclear war would be coded only 1. For this reason, too, researchers should rely primarily on a dichotomous indicator of disputes, as Bremer (1992; 1993) has done.

James et al. do report one test of the democratic peace using the dichotomous indicator of disputes. Not surprisingly, the results – relegated to an appendix (James et al., Table A-1) – are much stronger than in any of their other analyses: The coefficient of Regime is very significant \((t = -3.213)\). Moreover, in the Appendix (James et al., Table A-2), the effect of Dispute on Regime, which they characterize as “negative, but not statistically significant,” is miniscule \((t = 0.77)\).

(c) James et al. evaluate the democratic peace controlling only for economic growth rates, alliances, and contiguity. This is a specification we have never employed. Maoz and Russett (1993) included GDP per capita as well as economic growth rates in their tests. Oneal et al. (1996) controlled for economic interdependence. Most importantly, both of these studies included the ratio of the dyadic members’ national military capabilities (Singer and Small, 1995). Indeed, all our published and unpublished research reports use a measure of relative power, a variable representing the central thesis of realism. Others (e.g., Bremer, 1992; 1993) also routinely control for the balance of power, and for good reason. Tests of the democratic peace would not have been given any credence otherwise. The authors’ omission of our measure of relative power is doubly surprising, because they do not explain or even note the change and because in place of a policy of promoting democracy internationally they suggest the “balanced deterrence of aggression” (p. 9).\(^4\)

\(^3\)Maoz and Russett (1993) used ordered logistic analysis in one test to show that the results were consistent with their binomial analyses.

\(^4\)James et al.'s advocacy of balanced deterrence is inconsistent with our results and those of others (e.g., Kugler and Lemke, 1996; Bremer, 1992; 1993) that have consistently shown that it is a preponderance of power, not a balance, that leads to peace.
Table 1  Estimated ordered (sequential) logit models of hostility level, without and with \textit{CapRatio}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>z-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>0.0528</td>
<td></td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.2231</td>
<td></td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>2.1049</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>4.0624</td>
<td>45.640</td>
</tr>
<tr>
<td>Regime</td>
<td>-0.0064</td>
<td>-3.201</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.0069</td>
<td>-0.305</td>
</tr>
<tr>
<td>Proximity</td>
<td>1.7738</td>
<td>16.991</td>
</tr>
<tr>
<td>Ally</td>
<td>-0.9021</td>
<td>-7.997</td>
</tr>
</tbody>
</table>

$N = 16,371$

Pseudo $R$-squared = 0.056

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>z-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>0.0476</td>
<td></td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>0.2207</td>
<td></td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>2.0893</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.6539</td>
<td>35.066</td>
</tr>
<tr>
<td>Regime</td>
<td>-0.0070</td>
<td>-3.496</td>
</tr>
<tr>
<td>Growth</td>
<td>-0.0121</td>
<td>-0.524</td>
</tr>
<tr>
<td>Proximity</td>
<td>1.4169</td>
<td>12.546</td>
</tr>
<tr>
<td>Ally</td>
<td>-0.9207</td>
<td>-7.998</td>
</tr>
<tr>
<td>CapRatio</td>
<td>-0.0029</td>
<td>-5.032</td>
</tr>
</tbody>
</table>

$N = 15,884$

Pseudo $R$-squared = 0.066

Here we replicate the analyses James \textit{et al.} report in their Table 4, with and without the capability ratio \textit{(CapRatio)}, to show the effect of omitting this variable from their Equation 3.\textsuperscript{3} There are some minor differences between the authors' results and those reported at the top of our Table 1, where we reproduce their analysis. For some reason, 41

\textsuperscript{3}We do not use Beck, Katz and Tucker's (1998) correction so that the results will be comparable with those of James \textit{et al.} James \textit{et al.} say (p. 4) that Beck \textit{et al.}'s work weakens our results, but actually it strengthens our finding about democracy. The pacific benefits of democracy are also confirmed in analyses using an alternative correction for temporal dependence - the General Estimating Equation - that we prefer (Oneal and Russett, 1999a; 1999c). For a response to Garzke's (1998) critique of the democratic peace, see Oneal and Russett (1999b) where we estimate the influence of democracy on the likelihood of conflict controlling for states' preferences.
cases in our data are not included in their analysis; and our estimation provides slightly stronger results for Regime than does theirs. The results at the bottom of Table 1 show clearly, however, that the capability ratio does make a statistically significant contribution to explaining the likelihood of international conflict. The significance level of Regime increases slightly when CapRatio is included in the regression. Excluding the capability ratio from the system of equations has much more serious consequences when the effect of conflict on regimes (James et al.'s Equation 4a) is estimated, as we show below.

SUBSTANTIVE IMPORTANCE OF THE DEMOCRATIC PEACE

Despite their use of an inferior specification, James et al. themselves do find that two democratic states are more likely to be at peace (James et al., Table 2) and to be less involved in militarized disputes (James et al., Table A-1). They try to dismiss these findings with a discussion of substantive significance. On this point at least, we and the authors agree. We, too, are concerned with the substantive importance of our findings, not merely their statistical significance. We acknowledge that the variance explained in our models, as measured by the pseudo-$R^2$, is small. This is at least in part a result of analyzing a data set in which the vast majority of the observations of the dependent variable are zero. But the substantive significance of our results is not well represented by this statistic. “Peace,” represented by zeros in our data, is actually a multifaceted phenomenon, in different circumstances involving close co-operation, non-violent contention, or indifference between states. In principle, the full range of co-operative and conflictual behavior is subject to analysis along the lines we have suggested, but in this instance we and James et al. are looking at only

---

6 Note that the coefficients for the binomial logistic model on Dispute, reported in James et al.'s Appendix Table A-1, are virtually indistinguishable from those reported for their sequential model of Hostility (Table 4). This is further indication that the only statistically meaningful distinction is between dispute and non-dispute. There are simply not enough cases to permit discrimination among the levels of Hostility.

7 Certainly, we are open to suggestions regarding influences affecting the likelihood of conflict that should be included in our model; and the explanatory power of our analyses has increased as our research project has progressed. The pseudo-$R^2$ of the best-fitting specification in Oneal and Russett (1999a) is 0.46.
a truncated part of the full range of interstate behavior – with the consequent difficulty of explaining a large portion of the variance.

A more enlightening representation of substantive significance is to consider the difference a change in regimes makes in the probability that a dispute will occur. The pacific benefits of democracy are clearly shown in Oneal and Russett (1997, p. 284), where we report the probabilities that various dyads will become involved in a militarized dispute. With all continuous explanatory variables set at their means, the annual probability of a dispute for a contiguous pair of non-allied states is 0.086. But for a pair of democracies, ceteris paribus, the likelihood is only 0.054 – a reduction of 37 percent. A dyad composed of a democracy and an autocracy, on the other hand, has a likelihood of conflict equal to 0.137, more than two-and-half times that of a pair of democratic states.⁶

Medical researchers frequently estimate the influence of various risk factors on the probability that a person will suffer a heart attack in a particular year – a rare event, with only a 0.003 annual probability for U.S. citizens – and report that exercise or control of diet can reduce the probability of an attack by 20 or 30 percent. These findings are considered important for good reason. Even a small reduction in risk is significant if the event in question – a heart attack or militarized interstate dispute – has potentially fatal consequences. Moreover, what may be a marginal reduction in danger for an individual can, for a large population, be very beneficial. There are 574 observations of dispute among the contiguous dyads, 1950–85, using the latest data. A reduction of 37 percent in the incidence of disputes, the implied result if all states had been democratic, would have meant 212 fewer dyad-years of conflict during these years.

The authors seek to minimize the pacific benefits of democracy in several ways. First, they use our older measures, specifications, and data. Second, they use elasticities to detail the substantive significance of political regimes, which obscures rather than clarifies. For example, they evaluate the Regime elasticity of Hostility at the mean value of Regime, which in the data they use is 0.33. Thus, when evaluating the consequences of a 1-percent change in Regime, they are

⁶Since two autocracies fight with greater frequency (.071) than two democracies (.054) and autocracies fight democracies at the same rate that democracies fight autocracies (.137), democracies at the national level of analysis are, other things being equal, more peaceful.
estimating the effect of an increase of 0.0033 units. No wonder they find that such a change reduces the likelihood of conflict by a small amount! In fact, changes in Regime smaller than a single unit are not meaningful given the character of the underlying Polity III data. Even an increment of 1 unit is trivially small since the actual range of Regime is +100 to −160 and the standard deviation is 23. Third, in discussing the results reported in Oneal et al. (1996), they consider only our analyses of the complete set of politically relevant dyads. As the authors say (p. 21), a one standard-deviation increase in Regime reduces the likelihood of a dispute in this group from 0.0063 to 0.0058; but the majority of the politically relevant dyads are non-contiguous pairs of states for whom the likelihood of conflict is small. A change of a standard deviation in Regime among the contiguous pairs reduces the incidence of disputes from 0.0563 to 0.0521 – an equivalent percentage drop in a much larger probability. Finally, James et al. do not mention that the regression analyses in Oneal et al. (1996; also Oneal and Ray, 1997) were also estimated using a dichotomous measure of joint democracy in place of Regime. These results indicate a reduction in the probability of conflict of 80 percent for two democracies.

A FLAWED TEST OF THE ALTERNATIVE HYPOTHESIS

James et al. argue that war undermines democracy; but they fail to justify their specification of the equation explaining Regime, they evaluate a theory about phenomena at the national level of analysis using dyadic data, and they report results that are extremely sensitive to alternations in the test. Indeed, a more plausible instrumental-variable test, in which the capability ratio is incorporated in the explanation of conflict, produces results completely contrary to those that James et al. report:

(a) The authors provide little justification for their test of Lasswell’s (1977) thesis regarding the rise of the garrison state. Their hypothesis concerns the corrosive effect of war on democratic institutions; but they test it with data on militarized disputes, only 13 percent of which are wars. And their selection of control variables is not theoretically grounded. Several influences on regime type are ignored (see Gasiorowski, 1995, for example); and Growth and Stability are introduced without explanation. It is likely that economic growth promotes or sustains democracy; but, ironically, their equation may then suffer
from a problem of endogeneity and reciprocal causation, because several recent studies suggest that democracy increases economic growth rates (Leblang, 1997; Feng, 1997; Oneal and Silver, 1997). Moreover, the relationship between political stability and democracy is uncertain. Stability does not seem to be either a cause of democracy or a characteristic of democracies alone. Gurr (1974) and Lichbach (1984) reported that thoroughgoing democracies and autocracies both tend to be long-lived; it is "incoherent" politics with mixed political characteristics that are the least enduring. In any event, for Gurr and Lichbach, it was the character of the regime that determined stability, not the other way round. These are serious problems for James et al. because Stability and the terms involving it are the only exogenous variables unique to their Equation 4.

(b) James et al.'s test of the influence of conflict on regime type incorporates the odd assumption that dyadic peace affects dyadic democracy, a problem obscured by the names the authors have chosen for their variables — Regime, Stability, etc. — which suggest that these are characteristics of individual states. It is one thing to contend that a country at peace with all states is likely to be more democratic than is a country involved in serious hostilities with any one; but to assume that peace within a single dyad is apt to make both members of the dyad more democratic, regardless of their relations with all other states, is quite another. The magnitude of this problem is indicated by recalling that our data on politically relevant dyads encompass all contiguous pairs of states and dyads created by matching each of the major powers with all other states. Thus, every country is a member of at least 5 dyads, unless there are missing data; the United States participated in 160 dyads in 1985.10

As a preliminary test of Lasswell's thesis at the national level, we assess the effect of a country's involvement in militarized disputes on the character of its regime using vector autoregression or Granger-causality testing (Granger, 1969; Freeman, Williams and

10Empirically, Stability is unrelated to Regime. In a bivariate regression, the coefficient of Stability is not statistically significant (p < .05). Adding Stability1, as James et al. did, does not help; the R² remains .00.

10The difficulty of using our older dyadic data to assess a monadic relationship is compounded by the construction of the Regime variable, which measures not only how democratic the two members of the dyad are (given by the sum of the individual states' regime scores in the numerator) but also how different the regimes are (indicated by the difference of the regime scores in the denominator).
Lin, 1989). Thus, rather than assume that the reciprocal effects of democracy and conflict are simultaneous, we estimate an equation that includes lagged terms to capture the effect of conflict on regime. After all, it makes sense first to sort out the temporal sequence of the causal links if democracy and conflict are thought to be mutually related. We count the number of disputes ($NumDisp_{it}$) a state had in each year with all other states and calculate the democratic character of its regime ($Dem_{it}$) in the manner recommended by Jaggers and Gurr (1995; 1996). $Dem_{it}$ ranges from +10 for states that are thoroughly democratic to −10 for extreme autocracies. Both variables are measured monadically. We include three lags of each variable in the Granger test. The appropriate number of lags was determined by minimizing the final prediction error (Mahdavi and Sohrabian, 1991; Burkhart and Lewis-Beck, 1994).\footnote{There is no universally accepted method for determining the optimal number of lagged terms, but minimizing the final prediction error (FPE) is common. If the number of lags is too large, the estimated coefficients will be unbiased but inefficient, running the risk that useless or misleading information is incorporated in the estimates. If too few lags are used, the estimates will be biased (Freeman, Williams, and Lin 1989). We followed a two-step procedure. First, we estimated a set of autoregressive equations for $Dem_{it}$ using one through five lagged values for this variable over the same set of observations. Three lags minimized $FPE(m)$, which equals $[(N+m+1)/(N-m-1)] * [SSR(m)/N]$ where $N$ is the number of observations, $m$ is the number of lags, and SSR is the sum of squared residuals. We then added an equal number of lags of $NumDisp_i$. This equation minimized the $FPE(m, m)$ among all equations with equal numbers of lags for the two variables. $FPE(m, m)$ equals $[(N+2m+1)/(N-2m-1)] * [SSR(m)/N]$. For all lag lengths, our tests indicate that conflict does not “Granger-cause” democracy.}

The equation we estimate is:

\[
Dem_{i,t} = \beta_0 + \beta_1 * NumDisp_{i,t-1} + \beta_2 * NumDisp_{i,t-2} + \beta_3 * NumDisp_{i,t-3} + \beta_4 * Dem_{i,t-1} + \beta_5 * Dem_{i,t-2} + \beta_6 * Dem_{i,t-3}
\]

As seen in our Table 2, the autoregressive effects of the past values of the regime score dominate the results. The number of disputes in which a state was involved makes no statistically significant contribution to explaining $Dem_{it}$. The coefficients of $NumDisp$ are not jointly different from zero at the 0.37 confidence level. Ultimately, only theoretically specified, multivariate analyses that link democracy to peace,
Table 2  VAR analysis of national level of democracy ($Dem_t$)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>NumbDisp$_{t-1}$</td>
<td>0.0225</td>
<td>0.453</td>
</tr>
<tr>
<td>NumbDisp$_{t-2}$</td>
<td>0.0630</td>
<td>1.193</td>
</tr>
<tr>
<td>NumbDisp$_{t-3}$</td>
<td>-0.0873</td>
<td>-1.748</td>
</tr>
<tr>
<td>Dem$_{t-1}$</td>
<td>0.9131</td>
<td>19.969</td>
</tr>
<tr>
<td>Dem$_{t-2}$</td>
<td>-0.0081</td>
<td>-0.186</td>
</tr>
<tr>
<td>Dem$_{t-3}$</td>
<td>0.0707</td>
<td>2.285</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0140</td>
<td>-0.440</td>
</tr>
</tbody>
</table>

$N = 3040$
$R^2$-squared $= 0.943$

and peace to democracy, at the national level of analysis can establish if there are important reciprocal effects; but there is no evidence in these results that involvement in militarized disputes affects the character of political regimes. In addition, Mousseau and Shi (1999) report that countries are not more likely to become autocratic before the onset of a war.

(c) Finally, we return to the omission of the dyadic capability ratio in the authors' test of the democratic peace (James et al., Equation 3). As noted above, a measure of relative power is the *sine qua non* of any effort to explain international conflict, especially for authors who reject liberal theory in favor of realist prescriptions. Here we consider the consequences of omitting the capability ratio for James et al.'s estimation of the effect of hostility on regime type. If they had included this measure of the dyadic balance of power in their Equation 3, they would also have had to use it in the first-stage estimation of the instrumental variables *Peace-War* entered into their Equation 4a. Only in this way could they insure that the instruments are uncorrelated with the error term. Including *CapRatio* in the first-stage estimation produces dramatically different results in the second.

To show this, we first perform a multinomial logistic regression of *Hostility* on the exogenous variables in James et al.'s Equations 3 and 4, just as the authors did, and then estimate Equation 4a. We report in Table 3 the estimated coefficients of Equation 4a when *CapRatio* is omitted. As can be seen by comparing these results to James et al.'s Table 5, we have indeed replicated their analysis. The standard errors we report are substantially larger because we adjust for
**Table 3** Estimated equation for *Regime* without the capability ratio (*CapRatio*) in first stage

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peace (<em>L = 0</em>)</td>
<td>1.9746E+01</td>
<td>3.901</td>
</tr>
<tr>
<td>Threat (<em>L = 1</em>)</td>
<td>9.2983E+03</td>
<td>3.129</td>
</tr>
<tr>
<td>Display (<em>L = 2</em>)</td>
<td>-1.1378E+04</td>
<td>-4.177</td>
</tr>
<tr>
<td>Use of force (<em>L = 3</em>)</td>
<td>2.1433E+03</td>
<td>4.216</td>
</tr>
<tr>
<td>War (<em>L = 4</em>)</td>
<td>-6.4213E+03</td>
<td>-3.944</td>
</tr>
<tr>
<td>Other variables:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>2.4288E-01</td>
<td>0.718</td>
</tr>
<tr>
<td>Stability</td>
<td>-1.1390E+00</td>
<td>-2.250</td>
</tr>
<tr>
<td>Growth-squared</td>
<td>3.1697E-02</td>
<td>0.929</td>
</tr>
<tr>
<td>Growth × Stability</td>
<td>-1.3856E-01</td>
<td>-2.622</td>
</tr>
<tr>
<td>Stability-squared</td>
<td>1.7566E-02</td>
<td>2.281</td>
</tr>
</tbody>
</table>

*N = 16,371*

Adj. *R*-squared = 0.0444

*Note*: Standard errors adjusted for heteroskedasticity taking into account clustering of the data by dyad.

**Table 4** Estimated equation for *Regime* with the capability ratio (*CapRatio*) in first stage

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peace (<em>L = 0</em>)</td>
<td>-1.6635E+00</td>
<td>-0.835</td>
</tr>
<tr>
<td>Threat (<em>L = 1</em>)</td>
<td>1.2973E+03</td>
<td>0.996</td>
</tr>
<tr>
<td>Display (<em>L = 2</em>)</td>
<td>-8.3092E+02</td>
<td>-1.096</td>
</tr>
<tr>
<td>Use of force (<em>L = 3</em>)</td>
<td>2.8957E+01</td>
<td>0.285</td>
</tr>
<tr>
<td>War (<em>L = 4</em>)</td>
<td>6.4491E+02</td>
<td>2.427</td>
</tr>
<tr>
<td>Other variables:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth</td>
<td>6.3866E-01</td>
<td>1.717</td>
</tr>
<tr>
<td>Stability</td>
<td>6.2674E-02</td>
<td>0.119</td>
</tr>
<tr>
<td>Growth-squared</td>
<td>-4.2079E-02</td>
<td>-1.207</td>
</tr>
<tr>
<td>Growth × Stability</td>
<td>-1.1396E-01</td>
<td>-2.005</td>
</tr>
<tr>
<td>Stability-squared</td>
<td>8.5240E-03</td>
<td>1.207</td>
</tr>
</tbody>
</table>

*N = 15,884*

Adj. *R*-squared = 0.0135

*Note*: Standard errors adjusted for heteroskedasticity taking into account clustering of the data by dyad.
heteroskedasticity using Huber (or White's) correction and, unlike the authors, take into account the clustering of our observations by dyads.\footnote{We do not report Huberized standard errors for the ordered logistic analysis in our Table 1 because the Stata 5.0 statistical package does not estimate these for this type of regression; but in a logistic analysis of the dichotomous dependent variable that takes clustering into account, the $z$-score of the $Regime$ variable is $-3.24$ ($p < .001$) when $CapRatio$ is included in the equation. As noted in footnote 3, the coefficients of the ordered and binomial logistic estimations are virtually identical.}

Next we include $CapRatio$ among the exogenous variables in the first step of the estimation process and then re-estimate Equation 4a. Our Table 4 shows these results. They are very different indeed. The coefficients of $Threat$, $Display$, and $Use$ are now far from statistical significance; as a consequence, the adjusted $R$-square has dropped by more than two-thirds. More interesting, the signs of the coefficients of both $Peace$ and $War$ have switched; and the latter is quite significant. It now appears that the most severe interstate conflict promotes, not undermines, democracy! Maoz (1997) reports the same finding using a three-stage least-squares estimation of a system of simultaneous equations. This is precisely what one would expect if democracies are more likely than autocracies to win the wars they wage (Bueno de Mesquita, Siverson and Woller, 1992; Lake, 1992; Stam, 1996) and leaders who lose wars are likely to be removed from office (Bueno de Mesquita and Siverson, 1995). War may be a means whereby, over time, autocratic governments are defeated and replaced by democracies, as Mitchell, Gates, and Hegre (1999), too, report. Thus, the predominant effect of war on regimes in this analysis is the opposite of what James et al. contend.

The coefficients of $Peace-War$ change so dramatically when $CapRatio$ is added to the estimation of the instruments because it makes an important contribution to explaining $Hostility$. The coefficients of $CapRatio$ were statistically significant in the multinomial estimation for three of the four levels of $Hostility$. Only for threats, of which there are just 24 instances, was it insignificant. But the estimation of Equation 4a is sensitive to even minor changes in specification. Simply adding $Growth^2$, $Growth \times Stability$, and $Stability^2$ to the first-stage, as the authors should have done given their inclusion in Equation 4a, makes an important difference. The coefficient of $War$, for example, falls from $-6421$ (Table 3) to $-960$. Since Equation 4a is being estimated
CONCLUSION

We are confident that democracy reduces the likelihood of conflict. Our own research (Russett, 1993; Oneal, Oneal, Maoz and Russett 1996; Oneal and Russett, 1997; 1999a; 1999b; 1999c; Oneal and Ray 1997; Russett, Oneal and Davis 1998) indicates that this effect is substantively important, highly significant statistically, and robust. Moreover, our results are reinforced by a large body of analyses from other times, regarding other social systems, and employing a wide variety of methods (see Russett, 1993; Ray, 1995; Weart, 1998; and Russett and Starr, 2000 for reviews of the literature). By contrast, the effect of conflict on the character of political regimes has yet to be compellingly demonstrated. The tests of James et al. are not theoretically justified, and their results are not robust.

We agree with the authors that we must acknowledge the limits of our knowledge. Democracy, peace, economic interdependence, and other aspects of interstate relations are undoubtedly interrelated in complex ways (Russett, 1998). But our prescriptions for public policy can not be deferred until our knowledge is certain, for verification is beyond the capability of science. Inevitably, our conclusions are probabilistic statements of belief, and we have no choice but to play the odds. Accepting that democracies are less prone to become involved in interstate conflict may involve some risk; but rejecting it, given the weight of the evidence, is to run the greater danger of discarding powerful opportunities to reduce international conflict.

---

13The variables Peace, Threat, Display, Use of force, and War are extremely collinear. With James et al.'s specification, the lowest correlation is .635 (between Threat and War). The highest is −.998 (between Peace and Use). Consequently the standard errors reported are deflated, as a boot-strap analysis indicated.
REFERENCES


DEMOCRACY AND PEACE: REPLY TO ONEAL AND RUSSETT

PATRICK JAMES\textsuperscript{a}, ERIC SOLBERG\textsuperscript{b} and MURRAY WOLFSON\textsuperscript{c}

\textsuperscript{a}Professor of Political Science, Iowa State University and Visiting Lady Davis Professor of International Relations, Hebrew University, Jerusalem
\textsuperscript{b}Professor of Economics, California State University-Fullerton, Fullerton, California
\textsuperscript{c}Professor of Economics, California State University-Fullerton and Adjunct Professor of Economics and Global Peace and Conflict Studies, University of California-Irvine, Irvine, California

(Received 24 August 1999; In final form 24 August 1999)

The criticism of James, Solberg and Wolfson (JSW) (1999) by Oneal and Russett (OR) is not responsive to the methodological issues at stake. JSW argued that war is an endogenous feature of the world political and economic system. If its causes are to be measured, it must be as a structural equation in a simultaneous system. Wedded to the idea that "democracies never fight each other," OR rely on a single equation to justify their view. JSW claim that such an equation may be an \textit{ad hoc} reduced form with no causal implications unless the equation is explicitly identified as a structural equation. JSW expand the model to explain democracy and conflict as two endogenous variables. JSW do not claim to have discovered the true relationships between these variables by their minimal expansion of the structural relation. They do show that unless these (and other) variables are treated as part of a system, the results are unstable, contradictory, of minimal size and not a reliable guide to public policy.

KEY WORDS: Peace-democracy; International relations; Simultaneous system; Identification problem

Professors Oneal and Russett (2000) (OR) have not addressed the methodological focus of our study. Perhaps they have obscured this fact from themselves as well, since they continue to gather data, debate

\textsuperscript{*}Corresponding author. Department of Political Science, 503 Ross Hall, Ames, IA 50011–1204. Tel.: 515–294–0798. E-mail: pjames@iastate.edu.

215
and revise indices – as they and many others in this area of research have in the past decade – without seriously asking whether they are going about the problem in the right way.\footnote{OR acknowledge our observation that 97% of the cases (i.e., pairs of states in a year from 1950 through 1985) are peaceful. Moreover, they acknowledge our recognition that this places a great burden on any statistical technique to discriminate between those cases and the remaining 3%. OR err, however, in their subsequent claim that the use of a binomial indicator is better and researchers should rely primarily on the dichotomous indicator of disputes. There are many different kinds of disputes; diplomatic protest, for example, is quite different from armed conflict where lives are lost. The use of a single indicator ignores those differences, and runs the risk of committing aggregation bias in the statistical estimation of marginal effects. Nevertheless, our results are robust to the dichotomous treatment of conflict as shown in the appendix to our previous article.} Given the intensity with which the matter of democracy has been debated over the years, perhaps it is too much to expect more than a generally dismissive reaction. Still, there is progress to report in practice, since OR’s response to James, Solberg, and Wolfson (JSW) (1999) follows some of the methodological guidelines set out in that study. If our work has brought peace and conflict studies closer to an operational convergence of approach, we will have made the contribution we intended, and the way then becomes open to a less dogmatic examination of the data wherever it leads.

Our response to OR unfolds in five stages. First, we review the context and intended contributions of our previous study. The second stage provides an overview of OR’s reaction to this. Third, we try to draw some conclusions from the ongoing debate over the issues between OR and ourselves. In the fourth stage we address some specific and important issues related to measurement. Fifth, and finally, we try to sum up the debate and suggest where attention might be focused in the future.

1 JSW: CONTEXT AND INTENDED CONTRIBUTIONS

Our efforts were aimed at contesting the methodology through which the notion that “democratic countries never fight each other” became the ruling paradigm for many leading students of international relations (Levy, 1988; Maoz and Russett, 1993; Ray, 1995; Weart, 1998). JSW is a commentary on how a faulty methodology caused the ‘rabbit to get into the hat.’ Two questions about this occurrence arise at the outset; a brief answer to each query will help to place what follows in a more general context: First, why did numerous social scientists initially regard the democracy-peace thesis as implausible, only to be con-
vinced by apparent statistical evidence? Second, can a change in research methodology facilitate a return to a more balanced view of international relations?

For an answer to the first question, which pertains to a massive shift in worldview among a wide range of scholars, there is one obvious time and place to start looking: the sociology of knowledge about international relations at the winding down of the Cold War. The apparent failure of realism to explain the central problem of international relations, war and its causes (Vasquez, 1983; 1993; 1998), put the field in a position to experience a Kuhnian “paradigm shift” (Kuhn 1962). Widely regarded as moribund, both classical and structural realism looked as if they needed replacement—and quickly, given the surprisingly peaceful devolution and final conclusion of the Cold War. Earlier manifestations of neo-Kantianism (Doyle, 1986) and empirical tests (Babst, 1972; Small and Singer, 1976; Rummel, 1983) found a newly receptive audience in the research community. From Maoz and Abdolali (1989) onward, the neo-Kantian testing program, which featured the dyad-year as the standard unit of analysis, expanding data sets and increasingly sophisticated statistical methods, moved forward. Thus the answer to the first question posed above is straightforward and easily phrased in Kuhnian terms: Without something to replace the realist paradigm, “normal science” could not continue, and the democratic peace, which promised a better way, promptly took over in the increasingly introspective field of international relations.

Given the answer to the first question, the response to the second seems almost forced, namely, “maybe.” It is not clear what, at this point, would be sufficient to satisfy advocates of the democratic peace thesis that their worldview might need to change. The democratic peace has moved in less than two decades from a narrow hypothesis derived from a somewhat mechanical reading of Kant to the point of departure for virtually all discussion of either theory or policy in the field of international relations. It is likely to take extraordinary efforts to restore a sense of balance to a discipline still recovering from such a thorough transformation. Our study was intended as an initial step in that direction.

The reader will recall that JSW attempted, first, to persuade researchers that war and peace were endogenous features of a much more complex interdependent, historically evolving political, economic and social system. The statistical corollary to this thesis is that, to infer policy-relevant conclusions, estimates must be made of the causal relations between interlocked endogenous variables (which are themselves creatures of the international system). That is to say, a structural model
of those simultaneous relations must be deduced from theory and data. To estimate simultaneous structural equations from data generated by an interdependent system, they must be disentangled from one another. This is the classic question of identification.\(^2\)

2 THE REACTION OF OR TO JSW

OR see the preceding methodological issues quite differently and most recently have turned to what looks like a more purely inductive approach, namely, Vector Autoregression (VAR) models. To be sure, \textit{ad hoc} estimation of reduced form equations in which the endogenous variables are regressed against all the others may predict something of interest, but this method admits of no causal interpretation. \textit{A fortiori} this is true of VAR models and the somewhat misnamed Granger causality representations with which OR are now experimenting.\(^3\) An overview of VAR models should help to explain why they do not solve the problems raised by JSW with regard to the statistical analysis offered by Oneal, Oneal, Maoz and Russett (1996).

Within an \(N\)-variable vector autoregression of order \(p\), each of the \(N\) variables is regressed on \(p\) lags of itself and \(p\) lags of every other variable. Trends, seasonals, and other exogenous variables may also be included, as long as they appear in every equation and as long as no endogenous variables are included. Thus, VAR models are similar to reduced-form equations, and they have limited usefulness in estimating structural relationships. (If other endogenous variables are included in VAR models, estimation by seemingly-unrelated regression is required to obtain good statistical properties.) In VAR estimation, so-called “predictive causality” is an issue. In an unrestricted VAR, \textit{everything causes everything else} because lags of every variable appear on the right hand side of every equation. Predictive causality, however, contains little, if any, information about causality in the scientific sense.

While VAR models can be very useful for forecasting, they are not at all useful for assessing causality in a structural equation setting. We

---

\(^2\)The question initially was posed to economic statisticians 73 years ago. Working (1926) reminded his readers that prices and quantities were simultaneously determined by the interactive system of supply and demand. Consequently a simplistic “statistical demand curve” regression of prices on quantities might be (a) a demand curve; (b) a supply curve; or (c) some “reduced form” combination of the two.

\(^3\)See Kinsella and Chung (1998) for an insightful use of these techniques.
DEMOCRACY AND PEACE: REPLY TO ONEAL AND RUSSETT 219

find it curious that OR report neither a VAR equation for their hostil-
ity measure, *NumDisp* (the number of disputes) nor the results of a
formal test of predictive causality. In a VAR model, it is entirely pos-
sible for a variable *y* to "cause" variable *z* and, simultaneously, for *z* to
"cause" *y*. If this should occur, is it then logical to conclude that both
cause each other? Or is each being influenced by still other events?

Readers will notice that OR rely heavily on the above-noted tech-
nique in their reply to JSW. To be sure, in the first stage of the two-stage
method we have suggested, OR's estimates of the causal influence of
conflict (HOSTILITY) on democracy (REGIME) differ from ours; this
is so because OR opt to add a variable, CAPRATIO (a measure of the
dyadic balance of power) to our specification. CAPRATIO is the "ratio
of the stronger side's capability index to that of the weaker member in
a dyad" (Oneal and Russett, 1997a, p. 276). At least three problems,
however, arise from an introduction of this variable into our system.

First, it is not clear from OR's introduction of CAPRATIO what
role this variable is playing within the system. We are never told
whether CAPRATIO is presumed to be exogenous or endogenous.
How might it be modeled?

Second, OR gratuitously tax us for being "realists" even though the
hoary realist-liberal debate plays no part in our study. Having stereo-
typed us by this *ad hominem* argument, they go on to complain that we
have *not* constructed a realist model by not introducing CAPRATIO.
We are damned either way.

Third, it is not obvious why this variable, among the many capability-
based indicators that might have been selected, is the one designated
to represent realism. While CAPRATIO has 'face validity', it has not
been introduced as part of a network of variables from one or more
branches of realism; the "structural" and "neotraditional" schools, for
example, emphasize different variables while sharing an emphasis on
international conflict and competition (Vasquez, 1983; 1998). Since at
least one taxonomy of the wide range of capability-based variables
already exists (James, 1993; 1995), and the thrust of the argument by
JSW is in the direction of an identified system, the introduction of
CAPRATIO as an isolated representation of realism seems at once
too little and too much.

---

*The Correlates of War (COW) Project military capabilities index is "composed (in
equal weights) of a country's share of the system's total population, urban population,
energy consumption, iron and steel production, military manpower, and military expend-
itures" (Oneal and Russett, 1997a, p. 276).*
Whatever role might be played by CAPRATIO or other power-related variables in some future model, one thing stands out regarding the results obtained by OR: With respect to the crucial, policy-relevant influence of democracy on peace, which motivates this whole line of research (i.e., the effect of REGIME on HOSTILITY), OR arrive at the same estimates as we do.5

3 IDENTIFICATION, BIAS AND MAGNITUDE

It is not necessary to explain to readers of Defence and Peace Economics that the issue of identification of equations within a simultaneous system is not the same as dealing with the simultaneous equations bias that may also result. A structural equation is identified when it is specified to be “sufficiently different” from (combinations of) the others in the system. This is not a data question, nor a matter of minimizing statistical bias, but it is a theoretical specification that permits the researcher literally to know what, if anything, the researcher is measuring. Distinct from the identification issue, frequently a statistical issue of biased estimation occurs when the error terms in simultaneous systems are not independent of one another. Frequently – but not always – a two stage procedure may be attempted to develop “instrumental variables” to minimize the bias.

OR contest our technique in dealing with the problem of bias. This certainly is a matter that may profitably be discussed further, but it has nothing to do with the identification issue. We are not pleased by the seemingly ad hominem suggestion, apparently drawn from that confusion, that we have been less than fair to their point of view.6 Neverthe-

---

5 It is not clear why OR choose to replicate only the ordered logit model of the HOSTILITY equation from JSW. They do not report the multinomial logit version, the specification that we prefer. Nevertheless, it is reassuring that they get nearly the same results even when they include the CAPRATIO variable. To be more precise, coefficients and z-ratios reported for the variable REGIME are (-0.0064, -3.201) and (-0.0070, -3.496) when CAPRATIO is absent and present, respectively.

6 We would sympathize more with OR’s complaint that it was difficult to work through the analysis and exposition if we could understand how correctly calculated partial (marginal) effects and elasticities are “convoluted and misleading.” OR claim that we “stack the deck against the liberal position”, but their results and ours do seem very similar with respect to the key issue related to the impact of democracy on disputes; to be more precise, regardless of intentions or underlying beliefs about the democratic (or liberal) peace thesis, each research team finds a substantively modest connection.
less, the point is that identification, not bias, is the methodological issue at stake. It describes how the international system is put together politically. As the reader can see, apart from the VAR excursus, OR now follow our simultaneous model and the "democracy-peace" thesis dissolves into a more realistic view of multiple interactive causation.

We attempted to persuade researchers to address the issue of magnitudes of incremental (marginal) effects, not just the signs and p-values of statistical significance of the regression coefficients. OR claim that we challenge their finding that pairs of states (dyads) are more likely to have peaceful relations if they are both democratic. We do not make that claim. In fact, our results in the REGIME-HOSTILITY equation (which neither party disputes) tend to confirm the positive influence that democracy has on peace. However, we find that the influence of democracy on peace is so small as to be trivial.

OR's response, however, is to change the universe of discourse. They argue that since the danger of war involves such high costs, the virtually unlimited use of political and economic resources to prevent it is justified even if substantial influence seems unlikely on the basis of reproducible evidence.\(^7\) Clearly this is a normative question and not responsive to the positive issues we have raised about methods of inquiry. OR certainly are entitled to their value judgments but, in fact, the unbounded use of instruments to achieve highly improbable returns is impossible to implement. Unless the 'insurers' are of a purely altruistic turn of mind, neither individuals nor countries can protect themselves completely against all catastrophic outcomes without inducing other disasters such as economic or political bankruptcy. The issue becomes one of how much risk must be borne. How much should scarce political, economic and military resources be dedicated to the democratic transformation of regimes, and how much to other instruments and goals such as public health and safety, economic

\(^7\)A separate issue, beyond the scope of this study, concerns the reductionism inherent in assuming that continuing increases in the proportion of democratic dyads in the system will reduce disputes (or other conflicts, for that matter) incrementally. Macro-micro linkages, which are excluded by definition from any purely micro-level model (Bunge, 1996), may have unanticipated effects. To cite one possibility, which is developed at greater length in James (1999), the presence of at least one autocratic and potentially threatening great power may be essential for further increments of democraticness among dyads to continue having the effects anticipated by the democratic (or liberal) peace thesis. In other words, inter-democratic cohesion at the dyadic level may depend on a factor that always has existed at the system level, so a more fully democratic international polity may be prone to conflicts for reasons not built into the current models.
development, environmental restoration and education? Weart (1998, p. 291), in a study favorable to the democratic peace, offers this cautious observation: "crusades to impose democracy, from Athenians in the Aegean to Americans in the Caribbean, have usually made more enemies than friends." Thus it is important to bear in mind that, in a complex system of variables, it may be better to promote objectives such as those noted a moment ago, with the hope that governments will improve as steps forward take place in other aspects of life.

Furthermore, we faced a methodological and practical problem of our own. Since the issues we raised were fundamentally matters of principle, it was pointless (as well as logistically impossible) to generate a new data base, as some demanded of us. The question was, and is, this one: what is to be done with the data at hand? Consequently, we used the data provided to us by OR to estimate a two equation simultaneous model.

OR claim that we used measures of dispute and joint democracy that are now superseded. However, the new measures appear still not to have reached a steady state. (This point will be established in some detail during the discussion of how joint democracy is measured.) It is true that there are alternative measures, many of which may be better than those used in our study as well as theirs. However, we intentionally used the data set and measures that OR and others relied on at the time we began our study. We did this to avoid any complaint that differences in results could be due to different definitions and data. The point of principle at which we arrived is that that results of the proper regression studies and normalization show that the influences of all the variables – taken together as well as separately – are trivially small, and unstable, and that the constant term turns out to be the most significant element. As the reader can see from their reply, these outcomes are not contested by OR.

One of the results of our study was that, using the data at hand, the influence of HOSTILITY on REGIME appeared to be much greater than the reverse effect. OR seem to conclude that this was our main thesis. In fact, as we repeatedly explained to the reader, any such find-

---

4 OR are incorrect about the chronology of our discussions, but that, too is a side issue.

5 The more general issue of whether significance tests should be used when interpreting results that appear to be based on a population of cases has not been addressed by JSW and will not be pursued at length here. However, the consensus against such use of significance tests is well-established and causes further difficulties for the democratic peace thesis, since the coefficients it relies on almost always look better in that context than with regard to substantive impact.
ing has to be regarded with extreme caution. This was not a consequence of any desire to “fudge” our result, but instead to illustrate that even a small improvement in specification—in this instance, to a two-equation simultaneous model—was capable of yielding results that contradict the prevailing paradigm. We did not say that the true nature of the relations between these variables had been established, although we continue to believe that a negative impact of HOSTILITY on REGIME is more than merely plausible.

If OR are not convinced by our line of reasoning and findings as related to REGIME, neither are we. We would no more make international policy on the basis of our R-sq. of 0.0453 than their R-sq of 0.0135.

4 WHAT ABOUT MEASUREMENT?

For reasons we have explained, we chose not to put too fine a point on our skepticism about the measures that were used at the time of Oneal, Oneal, Maoz and Russett (1996) and still largely characteristic of practice today. The underlying data-base for describing political regimes as democratic or undemocratic is drawn from the Polity III data. A cardinal index is constructed from subjective ordinal judgments. It has become increasingly apparent, due to the work of Gleditsch and Ward (1997) and Wolfson, James and Madjd-Sadagi (1999), that the standard measurement needs to be reconsidered. The de facto measurement of democracy turns out to be a matter of the degree to which the executive is or is not institutionally constrained, rather than the usual concept of civil liberties.

Consider this issue from the point of view of legal history. William H. Rehnquist, Chief Justice of the United States Supreme Court (1998), has written an account of the limitations placed on civil liberties in the course of the US Civil War and the First and Second World Wars. He finds that the executive branch was concerned to “get on with the war” even at the cost of suspension of habeas corpus, freedom of expression and rights of association, dwelling and privacy. Sometimes they were assisted by the judiciary, and sometimes opposed. The interesting facts, in the context of the democratic peace thesis, are two-fold: First, the restraining decisions were made almost entirely after hostilities ceased, whence they entered as precedent in the case law. Second, the executives involved were several of the presidents usually associated with politically liberal causes and individual rights—Abraham Lincoln, Woodrow Wilson and Franklin Roosevelt.
Let us return now to the more immediate issue of how to measure democracy on a cross-national basis: the Polity description of loss of democracy might be leading or trailing indicators of war, but this has more to do with preparations and conduct of war than the influence of regimes on the likelihood of conflict. The data are pooled time-series and cross-sectional data with unresolved questions of their appropriate use in these regression studies (Beck, Katz and Tucker, 1998). The indices constructed from this data were non-linear, and some possessed undefined regions of singularity unless adjustments were made. Finally, and most important to our thesis, the development of a two-equation model is illustrative as a counter-example and was never intended to be a definitive description of the enormously complex international system. (For example, the role of STABILITY in the system, as OR observe, certainly could use more work.) The whole point of the exercise was to show that an unidentified model, with trivial substantive outcomes and minimal explanatory power, is unstable and not a proper guide to public policy.

OR believe that they can rebut us by also using the (almost) common data to arrive at different results. As we have explained above, this time they use our data (really, of course, theirs) and specifications, adding only CAPRATIO to get different outcomes. Freeman's (1989, p. 61) analysis of temporal aggregation helps to make a compelling point about the apparent reversal in effects uncovered by OR with respect to the impact of HOSTILITY on REGIME. Freeman noted that empirical political theories generally devote only limited attention to determining the natural time unit for testing. The options, in principle, are almost infinite, from assessing virtually instantaneous transmission of effects to focusing on centuries or even millennia. What, then, is right for testing the neo-Kantian, or democratic, or liberal, peace?

Consider the lesser-known side of the relationship, namely, the effect of international conflict (or peace) on democracy. Is the true story the one offered by OR (Table 2), in which the national level of democracy in a given year is predicted by three lagged annual values for number of disputes and democraticness? Or, is the defeat of Germany in 1918, followed by its politico-economic decline and fall into dictatorship and war – a process that took over two decades to complete and presented by JSW as a “canonical example” – closer to the mark? Or, again, is it possible that the operation of complex systems, as Freeman might have guessed, will look different depending upon the amount of time that is granted status within the model as an interval? Freeman's (1989, p. 72) compelling treatment of these issues also
advised that results at minimum and maximum aggregation by time be compared, which seems especially appropriate for the area of research under review, where the dyad-year continues to account for the overwhelming majority of available data-based knowledge.

What is the proper conclusion to be drawn from the various stages of this "experiment", meaning both those of OR and JSW? While many observations might be made, one stands out. The influence of HOSTILITY on REGIME may not be robust with respect to model specifications. This does not surprise us at all. Actually the striking result of OR's stage of the experiment is that the influence of REGIME on HOSTILITY appears to be unaffected.

OR seem to imply that we should write a new paper every time they change their data, definitions, or introduce new "controls" for various variables. As will become apparent from a brief review of further work in the same tradition as Oneal, Oneal, Maoz and Russett (1996), this is both impractical and inappropriate.

It is not practical to keep up with all of the iterations in measurement and research design within a very active program. Take, for example, the point concerning JSW's use of Maoz and Russett's (1993) JOINREG, supposedly a candidate for "Jurassic Park" in the measurement of democracy at the dyadic level. If not JOINREG, however, what then? One option might be DEMt, the lower score for DEMOC-AUTOCC (from Polity III), which is used by Oneal and Russett (1997a, p. 274) and represented as a theoretical step forward because of its consistency with the "weak link" assumption developed by Dixon (1994).

---

10 We are pleased that OR were able to replicate our REGIME equation. They err in claiming that we did not appropriately calculate standard errors. We too applied White's technique to generate asymptotically correct standard errors in the presence of heteroskedasticity, as is clearly indicated at the bottom of each table in JSW. Their claim that we "should have added" the quadratic terms to the reduced-form multinomial logit model of the first stage is unwarranted. We used a fully specified reduced form equation with all of the exogenous variables from both structural equations on the right-hand side. Even if we had not, the logit model for HOSTILITY is a non-linear equation estimated by maximum likelihood and the equation for REGIME is a linear regression model estimated by least squares. A quadratic specification for REGIME can be used to capture any non-linear relationship; the same cannot be said of a logit equation.

11 The idea of the weak link is that the likelihood of conflict will be "primarily a function of the degree of political constraint experienced by the less constrained state in each dyad" (Oneal and Russett, 1997a, p. 274). DEMp, the higher score in the dyad, also is reported in the results from Oneal and Russett (1997a), but DEMp is granted pride of place throughout the analysis and singled out for its theoretical significance.
Other possibilities also arise, however, with respect to the measurement of democracy. There is more at work here than just separating out low and high scores within a dyad. Consider JOINDEM and the other continuous variables that are created from DEM (i.e., DEMOC − AUTOC) and used in a parallel study by Oneal and Russett (1997b, p. 758). JOINDEM relies on the original Maoz and Russett (1993) formula. DEMSUM is the sum of the democracy scores in a dyad. DEMGEO, or \((DEM_i + 11)(DEM_j + 11)\), is their geometric mean.\(^{12}\) DEM\(_{min}\) and DEM\(_{max}\) also are used in this study.\(^{13}\) In addition to these continuous variables, “coherent democracy” is used to create a dichotomous measurement. A coherent democracy is defined as “a country without significant autocratic characteristics; it is operationally defined as a country for which DEMOC – AUTOC is greater than six” (Oneal and Russett, 1997b, p. 759).\(^{14}\)

Given all of the preceding possibilities for measuring joint democracy, which are taken from just two articles published a year after Oneal, Oneal, Maoz and Russett (1996), it is not unreasonable for JSW to have aimed for stability in trying to build on the findings of the latter study.

Ultimately, it might be possible to work with each and every one of the measurements for democracy that appeared in Oneal and Russett (1997a, b) and other contemporary studies. To grant priority to any one of them (or a subset), however, would require evidence of true superiority. But consider this summary of findings from Table 2 of Oneal and Russett (1997b, pp. 762–763):

There is somewhat less support for the democratic peace when the geometric mean of the regime scores is used in place of Maoz and Russett’s measure. It, too, is associated with a reduced incidence of conflict if all cases are considered \((p < 0.10)\); but the coefficient is much closer to zero and far from statistical significance for the contiguous states \((p < 0.52)\). The results using the sum of regime scores are weaker yet. It is not significant with either the complete set of cases \((p < 0.48)\) or the contiguous dyads \((p < 0.84)\).

\(^{12}\)As Oneal and Russett (1997b, p. 758) point out, 11 is added in each instance to prevent negative values.

\(^{13}\)An interaction term, DEM\(_{min}\)∗DEM\(_{max}\), also appears in one equation.

\(^{14}\)The coding of category six as noted is derived from the Polity III Project (Jaggers and Moore, 1995).
Later in the same study it is noted that DEM$_{ld}$ produces statistically significant results for the both the data as a whole ($p < 0.001$) and contiguous pairs ($p < 0.03$). By contrast, DEM$_{m}$'s coefficients ($p < 0.07$ and $p < 0.24$, respectively) are associated with dispute involvement. While Oneal and Russett present an explanation for this reversal, it adds further to the sense that results are very sensitive to measurement and that no 'clear winner' exists among the many available options.

Ironically, the measurement used by JSW, JOINREG, is hardly the worst among the options for advocates of the democratic peace thesis. Yet OR cite the choice of this indicator as evidence that JSW conducted "tests destined to fail": "It is hard to see how using our earlier specification", they observe, "gives the Kantian thesis a fair test." Taken together, the preceding details would appear to provide ample justification for the more conservative approach toward pursuit of cumulative knowledge adopted by JSW.

5 CONCLUSIONS

Let us sum up the outcome of the OR enterprise: They have contributed to demonstrating the general sensitivity of the model to changes in specification. In their experiment, it appears that this instability is confined to the least policy-relevant part of the model. They have confirmed our findings about the minute significance of democracy as related to peace. In these respects OR have protested too much; they have supported our thesis even more strongly than we would have done ourselves.

Even while actually following JSW in practical terms, OR have not responded explicitly to the methodological-identification point at issue. Our method derives from the interactive complexity of the international system, not imperfect measures. Consequently, they continue to believe that, if the single equation behavioral representation is refined sufficiently, it ultimately will produce a correct result that silences the critics. This cannot happen. The issues are conceptual rather than data-based, so no matter how hard they try, OR will never "out" the "damned spot". That is why we neither followed all of the iterations of their various versions of the democracy-peace thesis nor made an issue of our deep distrust of the data and indices. We did not set out to criticize them personally or minimize the contributions they have made over the years. Rather, we are critical of the paradigm for which they are arguably the most prominent and best spokesmen.
REFERENCES


