

DEMOCRACY AND PEACE: REPLY TO ONEAL AND RUSSETT

PATRICK JAMES*^a, ERIC SOLBERG^b and MURRAY WOLFSON^c

^a*Professor of Political Science, Iowa State University and Visiting Lady
Davis Professor of International Relations, Hebrew University, Jerusalem*

^b*Professor of Economics, California State University-Fullerton,
Fullerton, California*

^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 *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

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

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.¹ 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-

¹OR 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.

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, *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. *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, *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

²The 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.

³See Kinsella and Chung (1998) for an insightful use of these techniques.

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*

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-

⁵It 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.

⁶We 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.

find it curious that OR report neither a VAR equation for their hostility measure, *NumDisp* (the number of disputes) nor the results of a formal test of predictive causality. In a VAR model, it is entirely possible 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 technique 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).⁴ 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 stereotyped 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.

⁴The 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 expenditures” (Oneal and Russett, 1997a, p. 276).

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.⁸ 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.⁹ 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-

⁸OR are incorrect about the chronology of our discussions, but that too is a side issue.

⁹The 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.

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 *t*-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

⁷A 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.

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

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 O Neal, O Neal, 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 twofold: 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.

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_{HI} and DEM_{LO} 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 ($p < 0.84$).

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

¹³An interaction term, $DEM_{LO} * DEM_{HI}$, also appears in one equation.

¹⁴The coding of category six as noted is derived from the Polity III Project (Jagers and Moore, 1995).

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.¹⁰ That 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 O Neal, O Neal, 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 DEM_L , the lower score for DEMOC – AUTO (from Polity III), which is used by O Neal 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).¹¹

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" (O Neal and Russett, 1997a, p. 274). DEM_H , the higher score in the dyad, also is reported in the results from O Neal and Russett (1997a), but DEM_L is granted pride of place throughout the analysis and singled out for its theoretical significance.

REFERENCES

- Babst, Dean V. (1972) A Force for Peace. *Industrial Research*, 14, 55–58.
- Beck, Nathaniel, Jonathan Katz and Richard Tucker. (1998) Beyond Ordinary Logit: Taking Time Seriously in Binary-Time-Series-Cross-Section Models. *American Journal of Political Science*, 42, 1260–1288.
- Bunge, Mario. (1996) *Finding Philosophy in Social Science*. New Haven, CT: Yale University Press.
- Dixon, William J. (1994) Democracy and the Peaceful Settlement of International Conflict. *American Political Science Review*, 88, 1–17.
- Doyle, Michael. (1986) Liberalism and World Politics. *American Political Science Review*, 80, 1151–1169.
- Freeman, John R. (1989) Systematic Sampling, Temporal Aggregation, and the Study of Political Relationships. In James A. Stimson, (ed.), *Political Analysis, Volume 1*. Ann Arbor, MI: University of Michigan Press.
- Gleditsch, Kristian S. and Michael D. Ward. (1997) Double Take: A Re-Examination of Democracy and Autocracy in Modern Polities. *Journal of Conflict Resolution*, 41, 361–383.
- Jagers, Keith and Ted Robert Gurr. (1995) Tracking Democracy's Third Wave with the Polity III Data. *Journal of Peace Research*, 32, 469–482.
- James, Patrick. (1993) Neorealism as a Scientific Research Enterprise: Toward Elaborated Structural Realism. *International Political Science Review*, 14, 123–148.
- James, Patrick. (1995) Structural Realism and the Causes of War. *Mershon International Studies Review*, 39, 181–208.
- James, Patrick. (1999) *Elaborated Structural Realism and Crises in World Politics*. Unpublished manuscript.
- James, Patrick, Eric Solberg and Murray Wolfson. (1999) An Identified Systemic Model of the Democracy-Peace Nexus. *Defence and Peace Economics*, 10, 1–37.
- Kinsella, David and Sam-man Chung. (1998) The Long and Short of an Arms Race. In Murray Wolfson (ed.), *The Political Economy of War and Peace*. Norwell, MA: Kluwer Academic Press.
- Kuhn, Thomas S. (1962) *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Levy, Jack S. (1988) Domestic Politics and War. *Journal of Interdisciplinary History*, 18, 653–677.
- Maaz, Zeev, and Nasrin Abdolali. (1989) Regime Types and International Conflict. *Journal of Conflict Resolution*, 33, 3–35.
- Maaz, Zeev and Bruce Russett. (1993) Normative and Structural Causes of the Democratic Peace, 1946–1986. *American Political Science Review*, 87, 624–638.
- Oneal, John R. and Bruce M. Russett. (1997a) The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950–1985. *International Studies Quarterly*, 41, 267–294.
- Oneal, John R. and Bruce M. Russett. (1997b) New Tests of the Democratic Peace: Controlling for Economic Interdependence, 1950–85. *Political Research Quarterly*, 50, 751–775.

Later in the same study it is noted that DEM_{LO} produces statistically significant results for the both the data as a whole ($p < 0.001$) and contiguous pairs ($p < 0.03$). By contrast, DEM_{HI} '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.

- Oneal, John R., Frances H. Oneal, Zeev Maoz and Bruce Russett. (1996) The Liberal Peace: Interdependence, Democracy, and International Conflict, 1950–1985. *Journal of Peace Research*, 33, 11–28.
- Oneal, John R. and Bruce Russett. (2000) Why An Identified Systemic Model of the Democracy-Peace Nexus Does Not Persuade. *Defence and Peace Economics*, 11, 2.
- Ray, James Lee. (1995) *Democracy and International Conflict*. Columbia, SC: University of South Carolina Press.
- Rehnquist, William H. (1998) *All the Laws But One, Civil Liberties in Wartime*. New York: Knopf.
- Small, Melvin and J. David Singer. (1976) The War-Proneness of Democratic Regimes. *Jerusalem Journal of International Relations*, 1, 50–68.
- Rummel, R.J. (1983) Libertarianism and International Violence. *Journal of Conflict Resolution*, 27, 27–71.
- Russett, Bruce. 1993. *Grasping the Democratic Peace*. Princeton, NJ: Princeton University Press.
- Vasquez, John A. (1983) *The Power of Power Politics: A Critique*. New Brunswick, NJ: Rutgers University Press.
- Vasquez, John A. (1993) *The War Puzzle*. Cambridge: Cambridge University Press.
- Vasquez, John A. (1998) *The Power of Power Politics: From Classical Realism to Neotraditionalism*. Cambridge: Cambridge University Press.
- Weart, Spencer. (1998) *Never at War: Why Democracies Will Not Fight One Another*. New Haven, CT: Yale University Press.
- Wolfson, Murray, Patrick James and Zagros Madjd-Sadjadi. (1999) Insights from A Cluster Analysis of National Types. (unpublished manuscript).
- Working, E. (1926) What Do Statistical Demand Curves Show? *Quarterly Journal of Economics*, 41, 212–235.

