Online APPENDIX

Further Results for:

The Effects of the International Security Environment on National Military Expenditures: A Multi-Country Study

William Nordhaus*, John R. Oneal**, Bruce Russett***

October 27, 2011

* Department of Economics, Yale University (email: william.nordhaus@yale.edu)

** Department of Political Science, University of Alabama (email: joneal@ua.edu)

*** Department of Political Science, Yale University (email: bruce.russett@yale.edu)

In this appendix, we provide justification for our preferred specification of the liberal-realist model (LRM), given in the third column of Table 1 of the research note. It is this regression model that was used to generate our measure of the threat each nation faced in the international security environment over time, our *MID P-hat* variable. We include this ex ante measure of the external threat in various models of national military expenditures. We also provide mean values of *MID P-hat* and the ratio of military expenditures to gross domestic product for all states with populations greater than 500,000, 1950-2000. These are reported in Table A2. Note that in the figures and tables, we have substituted the shorter term *"Phat"* for *"MID P-hat"* to make them more concise.

Estimating the LRM with Observations for All Years

The standard approach to estimating the LRM has been to use only the onset of a dispute and omit in each dyadic time series observations that are continuations of the same conflict. This is appropriate when testing the hypotheses incorporated in the LRM but not for our purposes. To explain annual military expenditures we need estimates of the probability of conflict for each year. In addition, analyzing only the onset of disputes does not fully capture the severity of the external military threat. If states anticipate becoming involved in a protracted conflict, they would be expected to spend more on the military than if only a brief skirmish were expected. In accounting for national military expenditures, we need, therefore, a "continuation sample" that includes all years of all disputes when creating *MID P-hat*; but including *PeaceYears* in the LRM with a continuation sample produces biased estimates of the regression coefficients. It is contrary to the assumption that conflict and the years of peace are unrelated and an artifact of the construction of the variable. This is easily shown.

Suppose there is actually no relationship between the years a pair of states has been at peace and the occurrence of a MID. Then, regressing onsets on the years of peace would yield a coefficient of zero; but in the continuation sample, roughly half the observations coded one for a fatal dispute represent the second, third, or later years. After a year of conflict, the peace-years variable is set to zero, so for subsequent years there will be an inverse relationship between the years of peace and the probability of conflict. We confirmed the bias by examining relations between the United States and North Korea, 1950-2000. The estimated coefficient of the peace-years variable dropped from -0.15 (± 0.12) in the non-continuation sample to -0.34 (± 0.13) with all years included.

Thus, to obtain unbiased estimates of $p_{i,j}^{fatal}(t)$ and *MID P-hat* with the continuation sample, we must either omit the peace-years variable or create an instrumental variable (IV) for it If we solve the LRM using past values of the $p_{i,j}^{fatal}(t)$ variable, we obtain as appropriate instruments the lagged liberal and realist variables. We call the IV estimate of peace-years "PY-hat."

We also must consider the possibility that conflict will have reciprocal influences on the other independent variables. The onset of a serious dispute, for example, is expected to affect bilateral trade adversely; and the structure of government may change over the course of a major war. We addressed this potential problem by constructing a set of "historical instrumental variables" for each independent variable. These are equal to their actual values during peacetime and to their last peacetime values before a period of conflict. These historical IVs will be shown to be unnecessary so need not be discussed in detail.

In Table A1 we report five sets of estimated coefficients for the LRM with the continuation sample for 1950-2000. Column E, for reference purposes, includes the years of peace variable. Thus, the results in column E correspond to the second column of Table 1 in the research note. The only difference is that the continuation sample is employed in column E of Table A1, not just the first years of all fatal disputes. In columns A through D are four possible specifications for analyzing the continuation sample; IV variables are either included or excluded. In columns A and B are specifications with and without PY-hat; other independent variables take their actual values. In columns C and D, historical IVs are substituted for the explanatory variables of the LRM, and again PY-hat is either included or excluded.

Begin by comparing the estimated coefficient of PY in column E with those for the IV version in columns A and C and the coefficient for PY in the second column of

Table 1 in the research note. The coefficient in E is much more negative than the others, indicating that the bias discussed earlier is indeed present when analyzing the continuation sample. (The bias is even greater if we use the spline function as is common, instead of the simple count of the years of peace.) Note also that the peace-years IV is statistically insignificant in column A and marginally significant in C. This suggests that *PeaceYears* is significant in column E only because it is negatively correlated with additional years of conflict, not because it contains information about prior values of the other independent variables. Major differences appear between E and the other estimations for several of the independent variables, but there are no systematic differences in the estimated coefficients across equations A through D. Some differences are due to different samples. Using IVs reduces the number of observations. Figure A1 below shows the stability of the coefficients in the alternative specifications.

In the analyses of national military expenditures we report below, we focus primarily on the specification in column B of Table A1 for the following reasons. First, it is clearly desirable either to omit peace years or to use PY-hat, so that eliminates equation E. Second, the IV for peace years is statistically insignificant at the .05 level in columns A and C. Third, there are no substantial differences between the results in column D where the historical IVs are used and the analysis with the actual variables in B, but the latter are more precisely estimated. Apparently, the reciprocal effects of conflict on the theoretical variables of interest are a less important source of bias than is the peace-years correction. Finally, equation B has the maximum sample size, requiring fewer imputations in constructing estimates of the security environment.

We show our ex ante estimates of the annual probability of a fatal interstate dispute for eight representative countries in Figure A2. *MID P-hat B* was generated with our preferred specification B from Table A1. *MID P-hat E* is the specification in column E that includes *PeaceYears* and the continuation sample. *MID P-hat F* is derived from the second column of Table 1 in the research note, where peace years were used with the non-continuation sample.

The graphs in Figure A2 show the severity of the external threat of conflict faced by each country from 1950 through 2000. Differences in *MID P-hat B*, cross-nationally and through time, are purely the result of the predictors derived from liberal and realist theories; they do not include any country- or year-fixed effects. As can be seen, there are major differences between high-conflict countries like the United States, the USSR/Russia, China, and Israel and low-conflict countries such as Canada, South Africa, or New Zealand.

The problem with using the actual years of peace in estimating *MID P-hat* with the continuation sample is again evident in Figure A2. The resulting time series *(MID P-hat E)* move more erratically and are strongly influenced by the actual timing of disputes, not just their theoretical determinants. Leaving those biased estimates aside, the other measures are highly correlated. The average correlation coefficient among the *MID P-hat* variants A, B, C, and D is 0.965 for all countries and 0.958 for the largest 40 countries.

	Actual independent variables		Historical instrumental variables		Actual independent variables
	A	B	C	D	E
Dependent variable	fatinv_cont	fatinv_cont	fatinv_cont	fatinv_cont	fatinv_cont
Peace years					-0.0553
					0.0074
Peace years IV	0.0057		0.0271		
	0.0156		*0.0139*		
Small democracy	-0.0860	-0.0938	-0.1170	-0.1030	-0.0889
	0.0264	*0.0210*	*0.0338*	*0.0285*	*0.0193*
Large democracy	0.0430	0.0419	0.0308	0.0401	0.0532
	0.0163	*0.0134*	*0.0150*	*0.0154*	*0.0127*
Trade/GDP	-249.5000	-192.9000	-230.1000	-185.2000	-99.9900
	77.4900	*63.3400*	*71.4900*	*65.0500*	*35.1200*
Contiguity	1.6990	1.1980	1.4000	1.6950	0.9460
	0.4500	*0.3030*	*0.4240*	*0.4170*	*0.3120*
Distance	-0.7850	-0.6650	-0.7660	-0.7410	-0.6200
	0.1780	*0.1490*	*0.1660*	*0.1670*	*0.1320*
Ratio of GDPs	-0.5870	-0.5030	-0.4880	-0.7250	-0.3440
	0.5690	*0.4830*	*0.5140*	*0.5490*	*0.4580*
Allies	-1.0060	-0.9850	-1.3630	-0.8300	-0.4030
	0.3780	*0.2100*	*0.3370*	*0.2160*	*0.1950*
GDP relative to	11.7400	11.4200	9.9500	11.9100	11.7200
world GDP	*2.7370*	*1.9840*	*2.5250*	*2.0570*	*1.7130*
System size	-0.9690	-1.3870	-1.3140	-0.9290	-1.3850
	0.3790	*0.2450*	*0.3460*	*0.3090*	*0.2460*
Constant	0.1370	-0.1050	-0.3540	-0.1960	0.3420
	1.3440	*1.0510*	*1.2970*	*1.2260*	*0.9670*
Sample period	1950-2000	1950-2000	1950-2000	1950-2000	1950-2000
Observations	371,080	406,067	371,062	405,923	406,067
Pseudo R-sq	0.267	0.252	0.259	0.255	0.297
Pseudo log likelihood	-3710.5	-4556.2	-3667.4	-3866.5	-4285.8

Each coefficient is shown with standard error of the coefficient below in italics.
Dependent variable (fatinv_cont) is a binary variable reflecting whether a dyad has a fatal militarized interstate dispute (MID) in a year. The sample *includes* "continuations," that is, second and further years of a continuing dispute.

Table A1. Alternative specifications of LRM with continuation sample

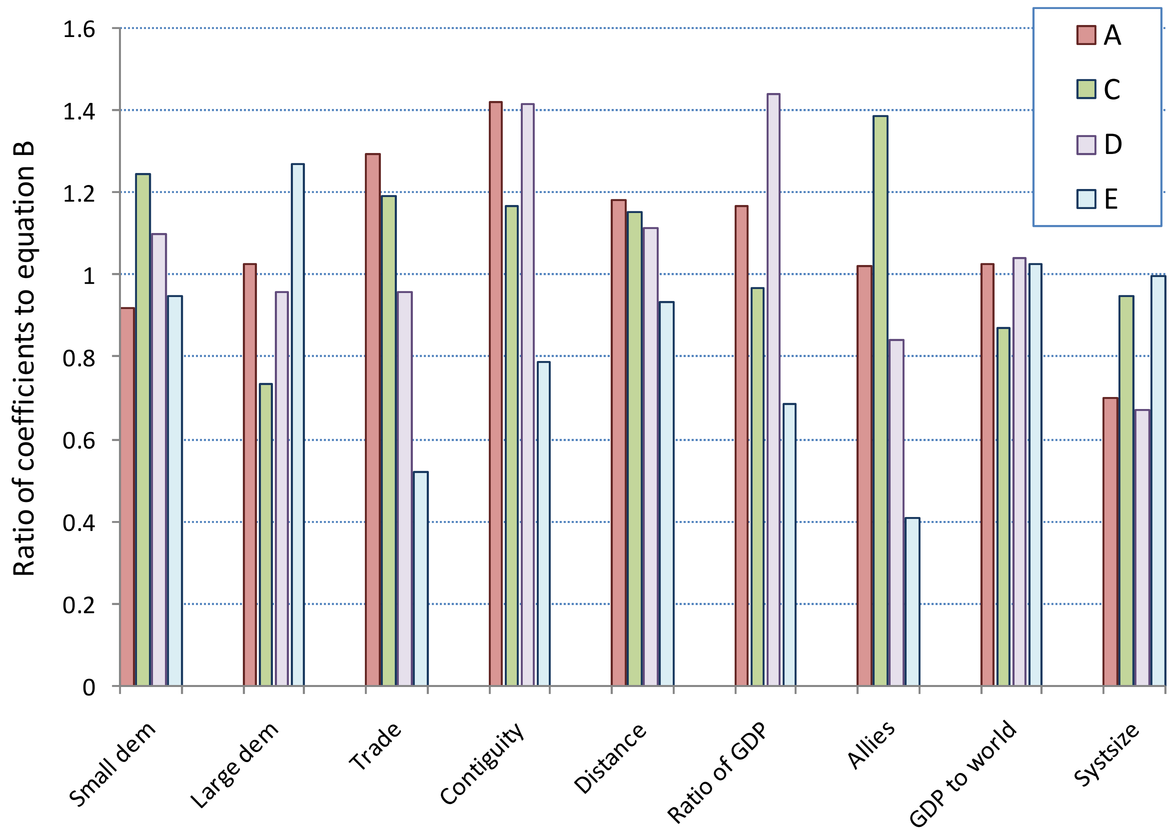

Figure A1. Stability of coefficients in Table 2: Ratio of coefficient in specification A, C, D, or E to specification B

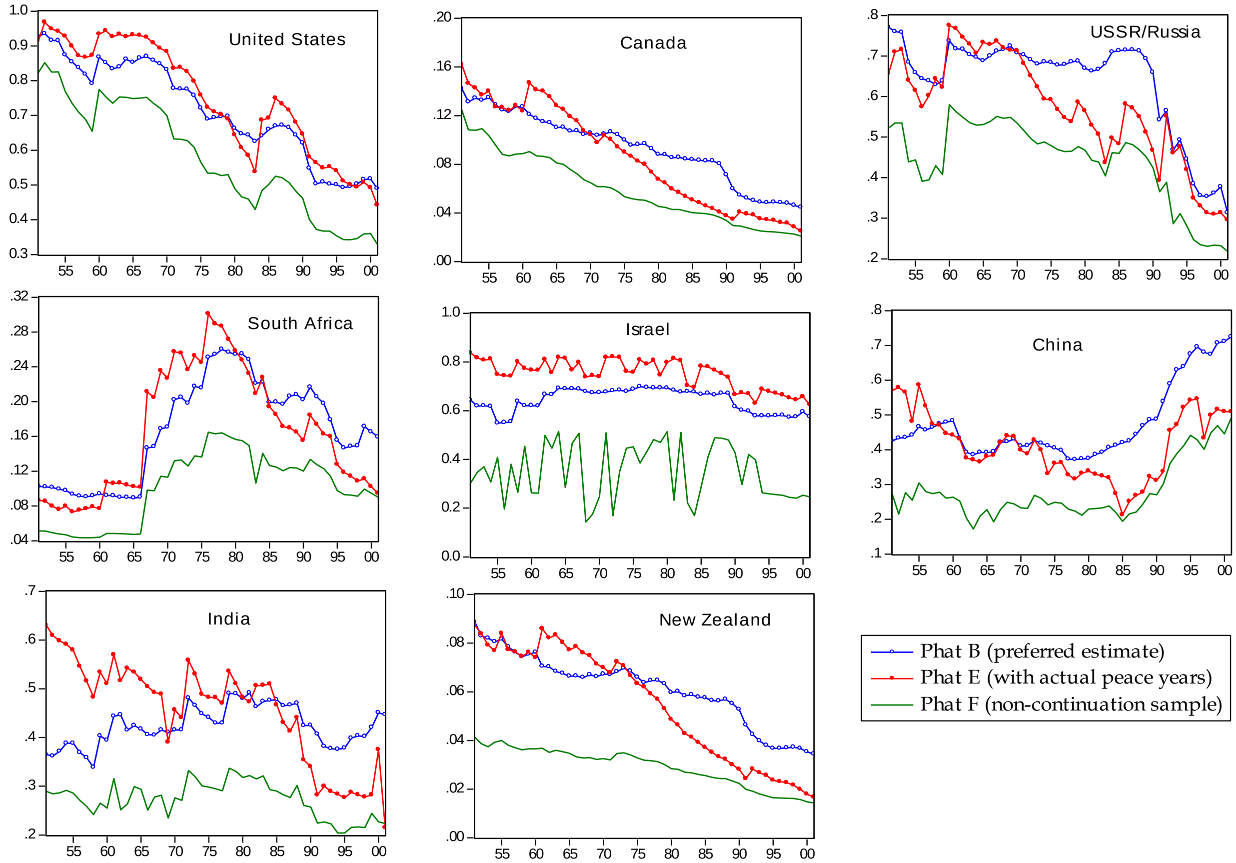

Figure A2. Calculated probability of conflict for eight countries, 1950 – 2000

These graphs show the estimated probability of conflict (fatal MID) for eight countries through time. Note the differences in the left-hand scale. Three estimates are shown for each country. The preferred estimate excludes peace years and uses the actual independent variables. The variant with actual peace years has excessive volatility (see Israel); the series using the non-continuation sample with actual peace years is even noisier.

Table A2. Probability of conflict (fatal militarized interstate dispute) and military spending as percentage of GDP, ranked by country, 1950-2000

Country	Probability of conflict (phat), % per year	Military spending/GDP, %
United States	71.7	5.5
Israel	64.2	11.2
Congo The Democratic Republic	63.9	1.3
Russian Federation	63.6	12.0
Congo	60.2	5.4
China	47.4	6.4
Yugoslavia/Serbia	45.7	5.3
Jordan	45.3	11.6
India	42.3	1.6
Syrian Arab Republic	40.8	14.5
Turkey	40.3	3.8
German Democratic Republic	39.4	6.6
Iran Islamic Republic of	37.4	3.2
Albania	34.7	5.4
Guinea	33.3	1.3
Lao People's Democratic Republ	33.1	3.6
Saudi Arabia	32.9	6.1
Bulgaria	32.8	5.0
Mozambique	32.5	3.2
Croatia	32.4	6.1
Italy	32.4	2.1
Germany	32.4	3.0
Egypt	32.0	5.9
Cameroon	31.7	1.6
Afghanistan	31.4	1.2
Korea Democratic People's Republic	31.2	37.6
Belarus (Byelorussia)	31.0	1.6
Pakistan	30.8	3.1
Greece	30.8	4.5
Myanmar	30.3	5.4

Country	Probability of conflict (phat), % per year	Military spending/GDP, %
Republic of Vietnam	30.0	20.2
Thailand	29.6	2.3
?	29.5	23.3
Sudan	29.4	3.2
Uzbekistan	28.6	1.6
Niger	28.5	0.9
Zambia	28.4	4.1
Azerbaijan	28.3	2.8
Ethiopia	28.2	3.8
Nigeria	28.0	1.6
Cambodia	27.9	3.4
Hungary	27.9	3.9
Tanzania United Republic of	27.9	2.6
Turkmenistan	27.5	2.1
Uganda	27.4	2.6
Rwanda	27.2	2.1
Chad	27.0	2.6
Austria	26.9	1.4
Cuba	26.7	4.4
None	26.4	11.3
Spain	26.0	1.2
Burkina Faso	25.8	1.8
Algeria	25.8	2.1
Denmark	25.7	2.8
Iraq	25.5	9.6
Gabon	25.5	1.4
Viet Nam	25.3	3.8
Romania	25.3	7.5
Sierra Leone	24.9	1.2
Togo	24.9	2.3

Country	Probability of conflict (phat), % per year	Military spending/GDP, %
Benin	24.8	1.9
Senegal	24.8	2.2
Korea Republic of	24.7	3.0
Swaziland	24.7	1.0
Mali	24.4	2.7
Tajikistan	24.3	1.3
Lithuania	24.2	0.6
Equatorial Guinea	24.1	2.9
Armenia	23.9	3.3
Ghana	23.8	2.2
Libyan Arab Jamahiriya	23.8	5.0
Cte D'ivoire	23.5	1.5
Nepal	23.2	0.7
Tunisia	23.1	2.2
Poland	22.6	5.2
Morocco	22.6	3.5
Qatar	22.5	11.2
Lebanon	22.5	4.5
Central African Republic	22.5	2.4
Latvia	22.4	0.7
Bosnia Herzogovinia	22.3	37.5
Bhutan	22.2	3.9
France	22.2	4.3
Cyprus	22.0	3.4
Japan	21.5	1.0
Kazakhstan	21.3	1.1
Norway	21.3	4.0
Angola	21.3	11.1
Zimbabwe	21.1	4.5
Kyrgyzstan	21.1	2.5

Country	Probability of conflict (phat), % per year	Military spending/GDP, %
Kuwait	21.0	5.8
Ukraine	21.0	2.5
Burundi	20.8	3.0
Malawi	20.5	1.4
Liberia	20.1	3.7
Botswana	19.7	1.4
Bahrain	19.3	5.3
None	19.1	5.5
None	19.1	33.0
Djibouti	18.8	6.6
Kenya	18.8	1.9
Macedonia	18.8	1.9
Sweden	18.1	4.1
Finland	18.0	2.2
Georgia	18.0	1.6
Oman	18.0	18.1
Somalia	17.9	2.6
Gambia	17.9	0.6
Mongolia	17.6	13.7
Indonesia	17.6	1.7
Malaysia	17.5	3.4
Luxembourg	17.0	1.8
Haiti	17.0	1.2
Mauritania	16.8	4.8
United Kingdom	16.8	4.2
South Africa	16.7	1.8
Nicaragua	16.6	1.6
Estonia	16.5	1.0
Mexico	16.3	0.5
Portugal	15.8	3.0

Country	Probability of conflict (phat), % per year	Military spending/GDP, %
Bangladesh	15.4	0.7
Dominican Republic	15.3	1.7
Brazil	15.2	1.3
Guinea-Bissau	15.2	3.0
Switzerland	15.1	2.3
Argentina	14.9	1.5
Honduras	14.5	2.0
United Arab Emirates	14.2	4.9
Guyana	13.6	1.8
Ireland	13.6	1.8
Lesotho	13.5	1.2
Jamaica	13.3	1.0
Colombia	13.3	1.5
Taiwan Province of China	13.1	6.7
Slovinia	12.9	1.5
Moldova (Moldovia)	12.8	0.9
Philippines	12.3	1.1
Paraguay	12.1	1.6
Netherlands	12.0	3.1
Belgium	11.9	3.3
El Salvador	11.9	1.4
Costa Rica	11.7	0.7
Bolivia	11.6	1.6
Guatemala	11.6	1.2
Venezuela	11.5	1.9
Uruguay	11.4	2.3
Comoros	11.1	2.0
Sri Lanka	10.7	1.5
Papua New Guinea	10.3	1.2
Panama	10.3	0.9

Country	Probability of conflict (phat), % per year	Military spending/GDP, %
Peru	10.2	1.9
Iceland	10.0	0.0
Madagascar	9.8	1.5
Canada	9.4	3.1
Mauritius	9.0	0.2
Ecuador	8.6	1.9
Chile	8.5	3.4
Trinidad And Tobago	8.2	0.9
Australia	7.0	2.9
Singapore	6.9	4.8
Solomon Islands	6.2	0.1
New Zealand	6.1	2.2
Fiji	5.2	1.4

Results of a Bootstrap Analysis for:
The Effects of the International Security Environment on National Military Expenditures: A Multi-Country Study

Xi Chen,* William Nordhaus,** and John Oneal***

*Department of Sociology, Quinnipiac University (email: Xi.Chen@quinnipiac.edu)

** Department of Economics, Yale University (email: william.nordhaus@yale.edu)

*** Department of Political Science, University of Alabama (email: joneal@ua.edu)

We have undertaken a bootstrap re-estimation of the central equation in our assessment of the impact of the security environment on military spending. We focused on row 2 in Table 2 as this was the one that could be implemented in STATA.

The procedure is complicated because it is a multi-stage calculation and there is no obvious technique to undertake this. We can write the system as follows (boldface are matrices or vectors):

(1) $\boldsymbol{p} = \boldsymbol{Ax} + \boldsymbol{e}$

(2) $\boldsymbol{Ph} = g(\boldsymbol{ph})$

(3) $\boldsymbol{m} = \boldsymbol{B}\ \boldsymbol{Ph} + \boldsymbol{Cz} + \boldsymbol{u}$

Equation (1) has $\boldsymbol{p}$ = the dyad-year value for a fatal MID or for peace, $\boldsymbol{x}$ is the vector of independent variables, and **e** is the error. $\boldsymbol{A}$ is a matrix of coefficients. We then generate $\boldsymbol{ph}$ as the forecasts of the estimates from equation (1).

Equation (2) is the product function (identity) that combines the dyadic predictions, ***ph***, into state-year predicted probabilities, ***Ph***, which are n x t vectors of predicted probabilities.

Equation (3) is the second stage of our system, in which we have exogenous variables **z** (such as GDP) along with the predicted probabilities generated from equations (1) and (2). ***B*** and ***C*** are matrices of coefficients, and ***u*** is the error.

The current paper uses ***Ph*** as generated in equations (1) and (2) in equation (3). This has the shortcoming of treating ***Ph*** as a fixed variable and not taking into account that ***Ph*** is a predicted variable and has a distribution.

There are many potential approaches to fixing the shortcoming (see Gary King, Michael Tomz, and Jason Wittenberg), but bootstrapping is the easiest to implement (B. Efron and R. Tibshirani.). We would generally think that the independent variables in equations (1) and (3) are fixed rather than random, but treating them as fixed tends to underestimate standard errors if there are misspecifications, so we treat all variables as random (John Fox). The procedure takes the following steps:

a) The first step is to generate a distribution of the ***Ph***. This is done by doing a bootstrap estimate of equation (1), generating a sequence of $\boldsymbol{ph^b(\theta)}$, $\theta = 1,\ldots,N$, N is large, each realization of $\boldsymbol{ph(\theta)}$ having approximately ½ million observations. We add a superscript "b" after a variable to indicate that it is a bootstrapped replication.
b) Then for each of these realizations, we generate $\boldsymbol{Ph^b(\theta)}$, $\theta = 1,\ldots,N$, where there are 5000+ observations for each of the bootstrapped realizations of $\boldsymbol{Ph^b(\theta)}$.
c) We then combined the bootstrapped $\boldsymbol{Ph^b(\theta)}$ into N replications of equation (3):

$$(3')\ \boldsymbol{m = B\,Ph^b(\theta) + Cz + u(\theta)},\ \theta = 1,\ldots,N.$$

d) We then take a single bootstrap of equation (3'):

$$(3'')\ \boldsymbol{m^b(\theta) = B(\theta)(Ph^b)^b(\theta) + C(\theta)z^b(\theta) + u^b(\theta)},\ \theta = 1,\ldots,N.$$

We then estimate equation (3′′) and treat that as a standard bootstrap. Assuming we have the proper specifications, the distribution of the parameters of $\boldsymbol{B(\theta)}$ will provide a consistent estimate of $\boldsymbol{B}$. (It is somewhat complicated to explain why this second bootstrap, and only a single run of the second bootstrap, are needed. The reason can be seen intuitively as follows. Assume that equation (1) were a perfect fit. Then if equation (3′) is not bootstrapped, the $\boldsymbol{Ph^b(\theta)}$ would be identical across different estimates of (3′) and the estimated variance of $\boldsymbol{B}$ would be zero. By doing the second bootstrap, we ensure that we resample both the $\boldsymbol{Ph^b(\theta)}$ and the $\boldsymbol{z}$.)

One further difficulty is the separate estimation of equations (1) and (3). The original table ran (1) in STATA and (3) in EViews because the preferred techniques for the two equations were not available in either program. We choose STATA for our bootstrapping because of sample size limitations in EViews. To conform to STATA, we used a variant of the equation in the second row of Table 2 in [***]. This equation is the pooled estimate with an AR correction. However, we cannot easily do an AR correction in a bootstrapping framework (because it involves re-sampling rows which would scramble the time periods), but we can get an estimate that is very close to that equation by omitted the AR term and adding a second lagged dependent variables. This leads to coefficient estimates which are close to those in row 2 of Table 5.

Our results are shown in the attached figure (for 1000 replications following Efron). We concentrate on the bootstrap for ***Ph***. The estimated standard error in the original regression was .07828. The estimate in the bootstrapped standard error is 0.08366. This indicates that the standard errors for the impact of the conflict probabilities on military spending are underestimated by about 5.8 percent for this equation. We can also estimate the standard error of the standard error. The ratio of the bootstrap to the original minus and plus one standard error are [1.0516, 1.0856]. Figure

1 shows the ratio of the bootstrap to the original coefficient along with the error bars. The STATA code for these calculations are available from the authors.

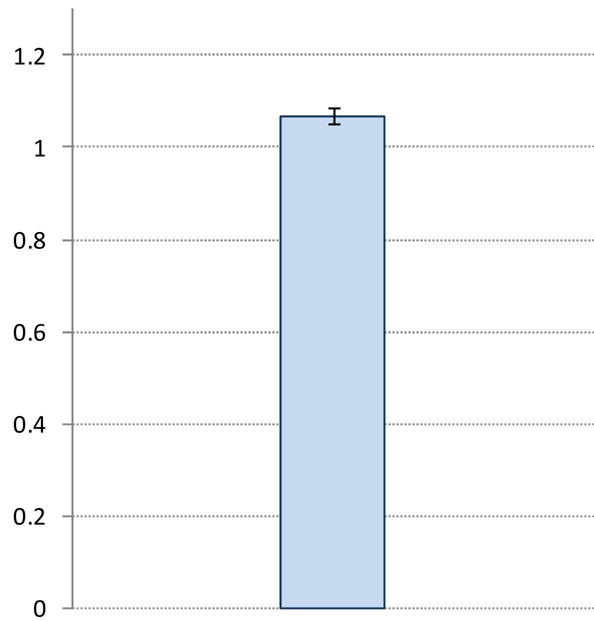

Figure 1. Ratio of coefficients of bootstrap to original regression with standard error of bootstrap coefficient as error bars

References

Bradley Efron, "Better Bootstrap Confidence Intervals," Journal of the American Statistical Association, Vol. 82, No. 397, (Mar., 1987), pp. 171-185.

Bradley Efron and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall, 1993.

John Fox, "Bootstrapping Regression Models," Appendix to An R and S-PLUS Companion to Applied Regression, January 2002, available at http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-bootstrapping.pdf.

Gary King, Michael Tomz, and Jason Wittenberg, Making the Most of Statistical Analyses: Improving Interpretation and Presentation," American Journal of Political Science, Vol. 44, No. 2, April 2000, Pp. 341–355.