Supporting Information

Choice or Circumstance?

Adjusting Measures of Foreign Policy Similarity for Chance Agreement

Frank M. Häge[1]

University of Limerick

Abstract

The supporting information in this online appendix presents additional replication results for the studies by Crescenzi (2007) and Salehyan (2008). In contrast to Gartzke's (2007) study discussed in the article, which relies on UN voting records, these studies use similarity measures based on alliance data. The results of these replications further support the notion that inferences drawn from statistical analyses can be sensitive to the choice of similarity measure.

[1] Email: frank.haege@ul.ie; web: www.frankhaege.eu.

Additional replication analyses

To investigate the consequences of replacing *S* by π or κ in statistical analyses, I conducted several replication analyses. In addition to the replication of Gartzke's (2007) study on the 'Capitalist Peace', which is reported in the main text of the article, I replicated two other recently published studies on the causes of international conflict: Crescenzi (2007) and Salehyan (2008). Both studies use *S* with an absolute distance metric as a control variable in the analysis, expecting a negative effect of similar foreign policy positions on international conflict. Both studies present the results of several regression models. My replications focus on what I consider to be the authors' most preferred, final, or complete model. In the first study, Crescenzi (2007) investigates how states' reputation for dealing with other states affects international conflict. The dependent variable is operationalized as the time until the onset of a militarized dispute between two dyad members. Crescenzi uses a semi-parametric Cox survival model to analyse the hazard rate of dispute onset. His sample includes all dyad-years from 1817 to 2000. While Crescenzi's findings support his main theoretical argument, the coefficient for the control variable *S* is positive and not statistically significant.

Using Crescenzi's original data, the regression results for Model 3 in Table 1 of Crescenzi (2007, 391) are easily replicated. The original estimate with 95 per cent confidence interval (0.22; CI -0.39, 0.82) is depicted in Panel (a) of Figure SI-1, together with the corresponding results of the replications.[2] Unfortunately, I was unable to replicate Crescenzi's results with my *S* measure that I had calculated directly from the Correlates of War Project (2003) alliance data.[3] Repeating Crescenzi's analysis with my *S* measure, the estimated coefficient is smaller and negative, but remains statistically insignificant (-0.07; CI

[2] The figure only presents the results for the regression coefficients of the different similarity measures. The full regression results are reported in Table SI-A1 in the appendix.

[3] My measure is almost indistinguishable from the measure provided by EUGene (Bennett and Stam 2000). The replication results using the measure provided by EUGene are virtually identical to those reported here.

-0.66, 0.53).[4] To make sure that any discrepancies between the estimation results for *S* and the estimation results for the chance-corrected agreement indices are not just due to different distance metrics, I also report estimation results for *S* calculated with a squared distance metric. The regression coefficient for the squared distance version of *S* is positive but very small (0.03; CI -0.58, 0.63); the difference to the absolute distance version of *S* is minimal. In contrast, the coefficients for π (-0.21; CI -0.54, 0.12) and κ (-0.23; CI -0.57, 0.11) are both negative and considerably larger in absolute value. The uncertainty surrounding these estimates is also smaller than those surrounding the two *S* measures. Still, the confidence intervals of the regression coefficients include zero. Thus, in this case, the statistical inferences are not affected by the choice of similarity measure.

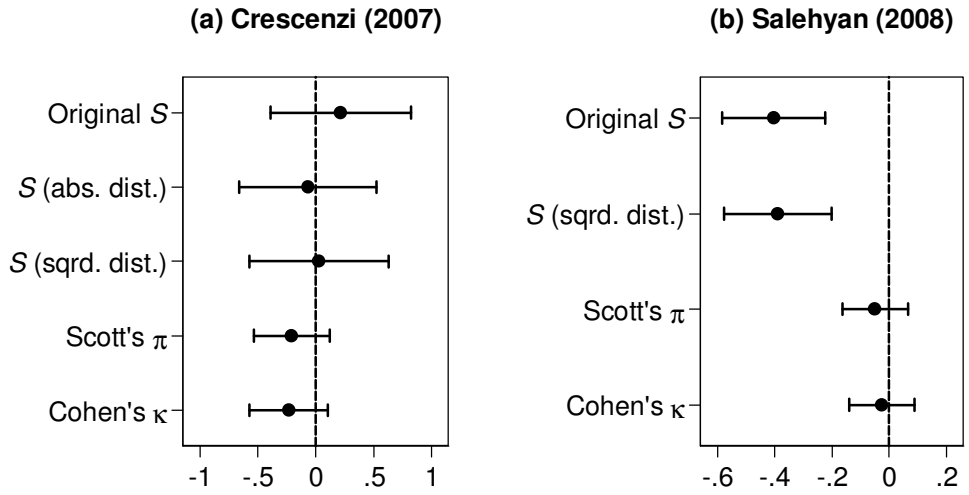

Fig. SI-1 Empirical Comparison of Similarity Measures. The two panels present the parameter estimates and 95 per cent confidence intervals for the coefficients of the original *S* measure and alternative similarity indices of two replication studies. Panel (a) presents the results for the replications of Model 3 in Table 1 of Crescenzi (2007, 391). Panel (b) presents the results for the replications of Model 2 in Table 1 of Salehyan (2008, 797). The former uses the unweighted *S* measure, while the latter uses the capability-weighted *S* measure. Scott's π and Cohen's κ are calculated on unweighted data in both cases. The complete regression tables of the replication studies are presented in the appendix.

[4] All other coefficient estimates in the model remain remarkably robust.

The choice of similarity measure is more consequential in the study by Salehyan (2008). This study is also interested in the causes of international conflict. Salehyan's main hypothesis is that refugee flows increase the likelihood of militarized interstate disputes. In contrast to Crescenzi (2007), Salehyan uses the weighted *S* measure in his analysis. The sample and type of statistical model also differ. Salehyan's analysis is restricted to politically relevant dyads, which consist of dyads of contiguous states and of dyads that contain at least one major power, between 1955 and 2000. The author examines the onset of militarized interstate disputes through a probit regression model.

Panel (b) of Figure SI-1 presents the results of the replications of Model 2 in Table 1 of Salehyan (2008). I was able to reproduce Salehyan's results exactly, not only with his original similarity measure, but also with my own.[5] The estimation results for the original, weighted *S* measure with absolute distance metric indicates a statistically significant and substantially large negative effect (-0.41; CI -0.58,-0.23) of foreign policy similarity on the likelihood of a militarized interstate dispute. Replicating Salehyan's analysis with the squared distance version of *S* does not change these findings in any substantive way (-0.39; CI -0.58,-0.20). However, the estimated coefficients for π (-0.05; CI -0.16,-0.07) and κ (-0.03; CI -0.14,-0.09) are both much closer to zero. Even though the uncertainty surrounding those estimates is smaller than the uncertainty surrounding the *S* measures, the 95 per cent confidence intervals include zero. On the basis of these findings and contrary to the original results, we would conclude that we cannot reject the null hypothesis that foreign policy similarity has no effect on international conflict. In addition to the similarity variable, one of the main independent variables relating to refugee flows in a dyad also loses statistical significance (see Table SI-A2 in the appendix). Echoing the results of the replication analysis of Gartzke (2007) presented in the article, these findings demonstrate that the choice of

[5] As the two are virtually identical, I only report the estimation result for Salehyan's original *S* measure.

similarity measure can have a profound impact on the conclusions drawn from a statistical analysis.

Appendix

Table SI-A1 Replication of Crescenzi's (2007) Cox survival analysis of dispute onset

	(a) *Crescenzi's S*	*(b)* *S (abs. dist.)*	*(c)* *S (sqrd. dist.)*	*(d)* *Scott's π*	*(f)* *Cohen's κ*
Similarity	0.22	-0.07	0.03	-0.21	-0.23
	(0.31)	(0.30)	(0.31)	(0.17)	(0.17)
Reputation	-13.23*	-13.34*	-13.30*	-13.41*	-13.40*
	(1.46)	(1.46)	(1.47)	(1.37)	(1.36)
Interaction History	-3.15*	-3.16*	-3.15*	-3.15*	-3.15*
	(0.22)	(0.22)	(0.22)	(0.22)	(0.22)
Reputation x	-30.62*	-30.78*	-30.72*	-31.06*	-31.05*
Interaction History	(4.31)	(4.30)	(4.30)	(4.33)	(4.32)
Contiguity	2.78*	2.81*	2.80*	2.85*	2.85*
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
Minor Powers	-1.29*	-1.26*	-1.27*	-1.25*	-1.25*
	(0.13)	(0.14)	(0.13)	(0.14)	(0.14)
Capability Ratio	-0.12*	-0.12*	-0.12*	-0.12*	-0.12*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Regime	-0.00*	-0.00*	-0.00*	-0.00*	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	565534	565534	565534	565534	565534
Failures	1998	1998	1998	1998	1998
Log likelihood	-13139.48	-13140.67	-13140.79	-13136.14	-13135.38
Chi^2	2483.58*	2410.36*	2414.24*	2497.57*	2500.88*

The table presents replication results for Model 3 in Table 1 of Crescenzi (2007, 391), using different measures of foreign policy similarity. The table reports coefficients with robust standard errors clustered on dyad in brackets. The sample includes all dyad-years from 1817 to 2000. Two-tailed significance tests: * = $p < 0.05$.

Table SI-A2 Replication of Salehyan's (2008) probit regression of dispute onset

	(a)	*(b)*	*(c)*	*(d)*
	Salehyan's S	*S (sqrd. dist.)*	*Scott's π*	*Cohen's κ*
Similarity	-0.40*	-0.39*	-0.05	-0.03
	(0.09)	(0.10)	(0.06)	(0.06)
Refugee Stock in Initiator	0.05*	0.05*	0.04	0.04
	(0.02)	(0.02)	(0.02)	(0.02)
Refugee Stock from Initiator	0.06*	0.06*	0.05*	0.05*
	(0.01)	(0.01)	(0.01)	(0.01)
Ref. in Initiator x Capability	0.09	0.09	0.11	0.11
	(0.06)	(0.06)	(0.06)	(0.06)
Ref. from Initiator x Capability	0.05	0.05	0.03	0.03
	(0.05)	(0.05)	(0.05)	(0.05)
Civil War in Initiator	0.22*	0.23*	0.20*	0.20*
	(0.04)	(0.04)	(0.04)	(0.04)
Civil War in Target	0.14*	0.14*	0.11*	0.11*
	(0.04)	(0.04)	(0.05)	(0.05)
Democratic Initiator	-0.02	-0.02	0.08	0.08
	(0.05)	(0.05)	(0.05)	(0.05)
Democratic Target	0.16*	0.16*	0.26*	0.26*
	(0.06)	(0.06)	(0.05)	(0.06)
Both Democratic	-0.42*	-0.43*	-0.57*	-0.58*
	(0.08)	(0.08)	(0.08)	(0.08)
Transitional Initiator	-0.19	-0.20	-0.21*	-0.21*
	(0.10)	(0.10)	(0.10)	(0.10)
Transitional Target	-0.03	-0.03	-0.04	-0.04
	(0.09)	(0.09)	(0.09)	(0.09)
Both Transitional	0.33	0.34	0.37	0.37
	(0.25)	(0.25)	(0.25)	(0.25)
Contiguity	0.76*	0.76*	0.64*	0.64*
	(0.06)	(0.06)	(0.06)	(0.06)
Colonial Contiguity	0.31*	0.31*	0.32*	0.32*
	(0.08)	(0.08)	(0.07)	(0.07)

Table SI-A2 Replication of Salehyan's (2008) probit regression (continued)

	(a)	*(b)*	*(c)*	*(d)*
	Salehyan's S	*S (sqrd. dist.)*	*Scott's π*	*Cohen's κ*
Capability Share	0.25*	0.25*	0.24*	0.24*
	(0.05)	(0.05)	(0.05)	(0.05)
Initiator's Trade Dependence	0.00	-0.00	-0.02	-0.03
	(0.20)	(0.20)	(0.26)	(0.27)
Target's Trade Dependence	-0.64	-0.64	-0.69	-0.70
	(0.39)	(0.39)	(0.39)	(0.40)
Shared IGO Membership	0.01*	0.01*	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
Peace Years	-0.21*	-0.21*	-0.21*	-0.21*
	(0.01)	(0.01)	(0.01)	(0.01)
Spline 1	-0.00*	-0.00*	-0.00*	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
Spline 2	0.00*	0.00*	0.00*	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
Spline 3	-0.00*	-0.00*	-0.00*	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	-1.83*	-1.83*	-1.98*	-1.98*
	(0.10)	(0.11)	(0.09)	(0.09)
N	86497	86497	86497	86497
Log likelihood	-4407.79	-4411.90	-4442.85	-4443.45
Chi^2	1631.68*	1624.68*	1571.91*	1569.07*

The table presents replication results for Model 2 in Table 1 of Salehyan (2008, 797), using different measures of foreign policy similarity. The table reports coefficients with robust standard errors clustered on dyad in brackets. The sample includes politically relevant dyads from 1955 to 2000. Two-tailed significance tests: * = $p < 0.05$.

References

Bennett, D. Scott, and Allan C. Stam. 2000. EUGene: A conceptual manual. *International Interactions* 26(2):179-204.

Correlates of War Project. 2003. "Formal interstate alliance dataset, 1816-2000." Version 3.03. http://www.correlatesofwar.org/COW2%20Data/Alliances/Alliance_v3.03_dyadic.zip (June 12, 2009).

Crescenzi, M. J. C., J. D. Kathman, and S. B. Long. 2007. Reputation, history, and war. *Journal of Peace Research* 44(6):651-667.

Crescenzi, Mark J. C. 2007. Reputation and interstate conflict. *American Journal of Political Science* 51(2):382-396.

Gartzke, Erik. 2007. The capitalist peace. *American Journal of Political Science* 51(1):166-191.

Salehyan, Idean. 2008. The externalities of civil strife: Refugees as a source of international conflict. *American Journal of Political Science* 52(4):787-801.