Ensuring Peace: Foreign-Imposed Regime Change and Postwar Peace Duration, 1914–2001

Nigel Lo, Barry Hashimoto and Dan Reiter

International Organization / Volume 62 / Issue 04 / October 2008, pp 717 - 736
DOI: 10.1017/S0020818308080259, Published online: 02 October 2008

Ensuring Peace: Foreign-Imposed Regime Change and Postwar Peace Duration, 1914–2001

Nigel Lo, Barry Hashimoto, and Dan Reiter

Abstract This research note develops a new explanation of postwar peace duration: periods of peace following wars last longer when the war ends in foreign-imposed regime change. This study tests this hypothesis on a new data set (an expansion of Fortna's (2004) data) of all periods of peace following interstate war cease-fires, over the period 1914–2001. It also tests for other possible factors affecting postwar peace duration, including international institutions, the revelation of information during war, third-party intervention during war, postwar changes in the balance of power, regime type, past conflict history, and others. The article finds strong support for the central hypothesis that peace lasts longer following wars that end in foreign-imposed regime change. This pacifying effect diminishes over time when a puppet is imposed, but not when a democracy is imposed. There are other results, including that the strength of a cease-fire agreement has almost no impact on peace duration.

What determines the duration of peace following interstate wars? Fortna proposed that the existence and nature of international cease-fire institutions determined postwar peace duration.[1] Werner and Yuen posited that the information environment during and after war determined postwar peace duration.[2] In this research note, we develop and test a third explanation. The breakdown of postwar peace means that one of the belligerents has violated its commitment to war-ending peace settlement. However, when one state suffers foreign-imposed regime change (FIRC) at war's end, the de facto removal of that state's foreign policy sovereignty sub-

Earlier versions of this article were presented at the 2007 annual meeting of the American Political Science Association, Emory University, the University of Pennsylvania, and the Massachusetts Institute of Technology. Thanks to attendees of these presentations for their comments. Thanks to Page Fortna, Suzanne Werner, and Amy Yuen for making their data sets available and answering our many questions. For research assistance and comments, thanks to Rochdi Alloui, Don Beaudette, Jan Box-Steffensmeier, Josh Bridwell, Philipp Fuerst, Eric Reinhardt, Michiyo Sakai, and Brandon Valeriano. This research was supported in part by the University Research Committee of Emory University. All opinions and remaining errors are exclusively our own. The order of the authors is random.

1. See Fortna 2003 and 2004.
2. See Werner 1999; and Werner and Yuen 2005.

International Organization 62, Fall 2008, pp. 717–36
 doi:10.1017/S0020818308080259

stantially reduces the chances that it will violate its war-ending commitment to peace, making postwar peace duration last longer. We test our new theory on an expanded version of Fortna's interstate cease-fire data set, examining the determination of peace duration after all interstate cease-fires from 1914–2001. We find strong support for our hypotheses that FIRC extends postwar peace duration, and that this effect diminishes over time when the victor installs a puppet regime, but it does not diminish over time when the victor installs a democracy. We find mixed support for the information hypotheses. We find almost no support for the institutions hypotheses, the sole exception being some evidence that some types of confidence-building measures can increase postwar peace duration.

Postwar Peace Duration

Werner's theory of postwar peace duration proposes that a settlement ending a war divides the disputed good in accord with the belligerents' understanding of the balance of power at war's end. Postwar shifts in the balance of power encourage the growing state to launch war to get more of the good, thereby shortening postwar peace duration. Werner found empirical support for her hypothesis, and that international institutions had no effect on postconflict peace duration.

Fortna presented a new theory that focused on international institutions rather than the balance of power. Drawing on institutionalism's proposition that international institutions facilitate cooperation, Fortna postulated that stronger international cease-fire agreements keep the peace longer between ex-belligerents. Fortna built a data set of forty-eight cease-fires reached by belligerents fighting interstate wars from 1946–98, using the Correlates of War (COW) data set and including intrawar cease-fires that Werner ignored. Event history analysis indicated that stronger cease-fire agreements increased the duration of postwar peace, and that changes in the balance of power had no effect.

Responding to Fortna, Werner and Yuen describe a new theoretical explanation for the endurance of postwar peace, proposing that the information environment (and especially the agreement between the two sides as to the balance of power) determines the duration of postwar peace. Werner and Yuen built on the bargaining model of war framework to argue that the information provided by wartime fighting determines postwar peace duration. First, they argued that when a war provides a consistent pattern of information, such as when one side wins all the battles, the two sides will have greater agreement as to the balance of power, which in turn makes the postwar peace more stable. Second, they argued that if a war is interrupted because of third party pressure to end the war, then a relatively smaller amount of information was revealed through fighting, and the postwar peace duration ought to be shorter.

Using the Fortna cease-fire data set, Werner and Yuen were the first to demonstrate that Fortna's results on the institutionalist agreement strength variable are insignificant with the use of a Cox event history model, but significant with the

use of a log-logistic model. In the Werner and Yuen model, both the battle consistency and interrupted war variables were statistically significant, and their inclusion made the change in the relative capabilities variable statistically significant and Fortna's agreement strength variable insignificant. Using Fortna's data, Long found that agreement strength does not affect the likelihood that two former belligerents will experience militarized interstate disputes (MIDs) after a cease-fire.[3]

Foreign-Imposed Regime Change and Postwar Peace Duration

War ends when belligerents agree to stop fighting. Postwar peace breaks down when one side breaks its commitment to this peace agreement, meaning that the breakdown of a peace agreement is an example of noncompliance with an international agreement. Foreign-imposed regime change renders compliance with war-ending agreements far more likely. Two sets of (interrelated) mechanisms help FIRCs ensure the peace. First, the architect of an FIRC can change the foreign policy preferences of the target country by executing, imprisoning, or exiling militarist leaders and their supporters, breaking up pro-war or pro-empire industrial cartels, revamping hypernationalist educational curricula, keeping hypernationalist/militarist statements and publications out of the public sphere, and/or empowering or importing leaders with more compliant and/or peaceful foreign policy preferences. The Soviet Union eliminated the non-Communist Hungarian leaderships in 1945 and 1956, replacing them with more pliable Communists each time. After World War II, the Allies removed the Nazi German and Imperial Japanese regimes root and branch, reshaping their governments and societies away from autocratic, militant, and imperial preferences and toward democratic, peaceful, and liberal economic preferences.[4] To further ensure compliance with war-ending settlements, the new target regime's policy preferences may allow for the permanent placement of the victor's troops within the target's borders.

Second, FIRC may transform the political institutions of the target. FIRC targets may be forced to hardwire pacifism into their constitutions or laws, for example, such as renouncing war, banning the deployment of troops outside the national borders, and/or accepting limits on the size and/or weaponry of the armed forces. For example, Italy, Germany, and Japan after World War II, Afghanistan after the Taliban era, and Iraq after the fall of Saddam Hussein all had pacific provisions of various kinds written into their postwar constitutions. Furthermore, some FIRC targets may be forced to adopt democratic institutions. Some political leaders (and scholars) believe that democratic states are more pacific for a variety of reasons: their elected leaders are constrained from starting wars by war-averse publics,

3. Long 2007.
4. Berger 1998.

greater transparency enables democracies to resolve crises short of war, and democracies embrace nonviolent norms of conflict-resolution.[5] Put differently, democratization may be an institutional means of more permanently empowering actors with pacific preferences over actors with militarist preferences.

We propose that FIRC extends postwar peace duration in the following hypothesis:

H1: Peace following interstate war lasts longer when the war ends with one state suffering foreign-imposed regime change.

The pacific effects of FIRC should diminish over time when a puppet is installed, as eventual leadership change in both the victor (puppet-master) and defeated (puppet) state increase the likelihood that the defeated state will be allowed to pursue an independent foreign policy. For example, in 1956 the Soviet Union imposed regime change on Hungary, but Soviet motivation and ability to control Hungarian foreign policy decreased with Soviet General Secretary Mikhail Gorbachev's late-1980s abandonment of the Brezhnev Doctrine, and collapsed with the 1989 fall of the Berlin Wall. However, the effects of imposed democratization (or democracy-FIRC) should be more stable over time, as such regimes' institutions have been, in theory, permanently transformed to be more pacific. Further, a democratic FIRC is likely to have been imposed by a democratic victor, and democracies tend to create more stable and mutually acceptable postwar orders through reliance on multilateralism and the exercise of restraint.[6] This leads to the following hypothesis:

H2: The pacifying effect of FIRCs establishing puppet regimes diminishes over time.

Research Design

Our data set expands the temporal range of the Fortna 1946–98 cease-fire data set backward and forward in time, increasing its range to 1914–2001. Doing so offers several advantages. First, it allows exploring whether the Werner and Yuen null

5. Reiter and Stam 2006. Military occupation is neither co-equal with foreign-imposed regime change nor a necessary condition for the duration of peace, as sometimes occupation has not followed a successful FIRC (the 1956 FIRC on Hungary and the 1978 FIRC on Uganda), sometimes an FIRC with occupation has experienced the breakdown of peace (Germany's 1940 FIRC of France), and sometimes there is a breakdown of peace with an occupation but without an FIRC (French occupation of German territories after World War I). Edelstein's (2004) occupation data has some exclusions that make its application here problematic. Specifically, he does not include any World War II occupations (such as the German occupation of France), nor does he code the deployment of Soviet armed forces after World War II in Eastern European countries as occupations.

6. Ikenberry 2000.

finding for institutions in the 1948–98 data set using a Cox model reflects a genuinely null relationship, or is an artifact of a small sample size. Our expanded data set nearly quadruples the number of cease-fires (to 188 cease-fires with fifty-six resumptions of war (failures)), and increases the number of cease-fire dyad-years by more than seven-fold (from Fortna's 876 cease-fire dyad-years to our more than 6,300 cease-fire dyad years).

Second, expanding the data range will improve our ability to test our principle hypothesis: that cease-fires culminating in FIRC enjoy longer periods of postwar peace duration. In Fortna's data set, only three cease-fires, the 1956 Russo-Hungarian cease-fire, the 1975 North Vietnam-South Vietnam cease-fire, and the 1979 Uganda-Tanzania cease-fire, accompany FIRC. Of these, the Vietnam cease-fire only lasts a single day in the data set because South Vietnam experiences state death when the war ends. Going backward and forward in time will provide a wider array of FIRC cases, improving our ability to assess FIRCs' effects on postwar peace.

Third, expanding the data set backward in time allows the inclusion of major sets of cease-fires, most notably the agreements that ended World Wars I and II. Much of international relations theory is built around explaining why the Versailles agreements failed to make World War I literally the "war to end all wars," and why the post–World War II settlements met such great success at curing German, Japanese, and Italian militarism. Expanding the data allows more rigorous exploration of these historical puzzles.

Our central question is: What factors determine the duration of peace following interstate cease-fires? Following Fortna, our data set has an event history structure, such that each subject is the duration of peace following a cease-fire ending the fighting in an interstate war. Our data set of interstate wars is based on the COW interstate wars data set (3.0), which contains all military conflicts between recognized members of the interstate system inflicting at least 1,000 battle casualties, up to 1997.[7] We also included five post-1997 wars: the Ethiopia-Eritrea war, the Kargil war, the Kosovo war, the Congo war, and the Afghanistan war.

Note that each subject is a cease-fire-dyad, and the dependent variable is the length of time, in days, until that dyad "fails" and experiences a new war. Some dyads never experience failure during the observation period, and are "right-censored" when one member of the dyad dies, or when the temporal end of the data set is reached (2001). Following Fortna, we excluded a very small number of trivial cease-fires that ended fighting only briefly, and included some intrawar cease-fires that were not recognized by COW.

Since the unit of analysis is the dyad, multilateral wars must be disaggregated into dyads, a major issue for World Wars I and II in particular. A single, multilateral war may have multiple cease-fires, as individual dyads may end their fighting by agreement even though the rest of the war rages on. Although Nazi

7. Sarkees 2000.

Germany did not surrender to the Allies until May 1945, it reached individual cease-fires before then with countries such as Poland, Norway, and the Netherlands. All of these individual cease-fires are included as separate subjects. As long as the two signatories of the cease-fire are still members of the interstate system (on state death, see below), each decides year to year whether or not to violate the terms of the cease-fire agreement and re-attack its neighbor, meaning that all of these individual cease-fires within the world wars should be included as separate subjects.

One could argue that the inclusion of all these individual cease-fires biases the analysis, either because the effect may be to inflate artificially the true size of the data set, and/or because the failure of cease-fire subjects may not be independent. Several cease-fires may fail simultaneously if, for example, a coalition from a previous war attacks a former adversary. We include all of these cease-fires because they represent sovereign decisions, and doing so follows Fortna's procedure. Empirically, many individual cease-fires within larger wars do not move together in lock step and need to be recorded and examined separately. A belligerent may sign a separate peace agreement, opting out of larger coalition wars while its allies fight on, as did Romania in 1916, Russia in 1918, France in 1940, and Finland in 1944. Importantly, different peace agreements signed by various members of a coalition may have consequentially different histories. Belligerents may switch sides within a war. France fought with the United Kingdom until 1940, against the United Kingdom in 1941, and then with the United Kingdom again in 1944. Hungary and Italy also switched sides during World War II. Multilateral cease-fires may break down only partially. The 1918 cease-fire between the Allies and Germany broke down in 1939, but only between the United Kingdom/France and Germany, as the United States did not join the war until 1941. Italy and Japan joined Germany in World War II rather than fight with the United Kingdom and France, their World War I allies. We account for the possibility of spatial dependence among cease-fire durations by using robust standard errors, clustering cases together into dyads from the same war, as Fortna, Werner, and Yuen do. Our results are robust across different clustering schemes.

A possible complicating feature is the treatment of minor participants and cease-fires involving several signatories. Fortna generally excluded participants whose committed forces were less than 10 percent than those of the largest coalition member.[8] Though this is a straightforward rule when applied to single front, multilateral wars such as Korea, Vietnam, and the 1991 Persian Gulf War, it becomes more complicated when applied to multicampaign wars such as World Wars I and II. Though Norway's troop contribution was less than 10 percent of that of the American or Soviet troop contribution for the entire war, the 1940 Germany-Norway campaign did end with a separate cease-fire, and Norway

8. Fortna 2004, 46, n. 14.

did make a greater than 10 percent contribution on the Allied side during the Germany-Norway campaign itself. In discerning whether or not an ally made a sufficiently large contribution, we applied the 10 percent rule on a cease-fire-by-cease-fire basis, leading us for example to include a Netherlands-Germany 1940 cease-fire dyad, but not a Netherlands-Japan 1945 cease-fire dyad.

Another salient issue is the treatment of state death (defined as a state's exit from the COW list of state members), which occurs in the context of war if a state is conquered and its national government no longer enjoys international recognition.[9] As an econometric issue, event history analysis uses discontinuous intervals of risk to handle state death. During a subject's discontinuous interval a cease-fire failure is impossible and the data set contains a gap.[10] Sometimes, state death occurs at the same time as a cease-fire, if the cease-fire means the conquest and annexation of the defeated state. In such instances, we follow Fortna's South Vietnam precedent of coding postwar peace duration to last for a day before being right-censored. We also include a number of cease-fire dyads in which Germany "killed" its target during World War II, but then the target was "revived" in the latter years of war. Does the cease-fire fail when the country is revived, as the war is still ongoing? We generally followed COW rules on war participation in determining whether or not a cease-fire ended upon a state's revival. So, for example, we coded the 1940 Germany-France cease-fire as failing in 1944 when France re-enters World War II, according to COW, but we did not code failures for other countries killed by Germany that later regained sovereignty (such as Belgium and Norway), because they are not coded by COW as re-entering the war.

A related concern is the postwar years of intra-world war cease-fires, especially those in which one country defeats another country, establishes a cease-fire, but loses the world war. An example is the Germany-Yugoslavia 1941 dyad, in which Germany wins the 1941 campaign, but loses World War II. For many of these cease-fires we saw no reason to censor the duration of peace at the end of the world war. For example, we coded the duration of the 1941 Germany/Yugoslavia dyad past 1945.

That being said, there are no uncontroversial research design solutions to this issue of post–world war years for intra-World War II cease-fires. We defend our approach in the context of three sets of issues. First, one might argue that the intra-world war cease-fire dyad ought simply to be censored out of the data set once the world war ends. This is probably too severe and atheoretical a solution, unnecessarily discarding too much data. We should note that taking such a step eliminates all cases in which an FIRC cease-fire ended in war, meaning that if one takes this approach FIRC predicts nonfailure perfectly, strengthening support for hypothesis (1) (though it is impossible to get parameter estimates because predic-

9. Fazal 2007.
10. Box-Steffensmeier and Jones 2004, 99.

tion is perfect).[11] Missing data and censored behavior lead us to drop dyad years when at least one member of the dyad is dead.[12]

A second issue concerns coding FIRC. One could argue that once a world war ends, any intrawar regime imposed by the war's loser (such as the regime installed in Yugoslavia in 1941) ought to be recoded to 0, given that the end of the world war brings its overthrow. However, this recoding would erase the two cases of FIRC failure, as the 1999 Kosovo war would be coded as occurring in the context of a non-FIRC for the Italy-Yugoslavia and Germany-Yugoslavia dyads. Furthermore, not allowing FIRC codings to vary biases the results against hypothesis (1), as if an imposed regime is replaced then war ought to be more likely, but that increased likelihood of war would occur with a coding indicating the presence of an FIRC. Lastly, we model this possibility by allowing for the effects of puppet-FIRCs to decrease over time, which allows for a range of factors to nullify those effects, including political evolution in the FIRC target and imposer.

A third and perhaps related issue concerns coding AGREEMENT STRENGTH. One could argue that when a world war ends, intrawar cease-fire mechanisms ought to be considered null and void. We change the AGREEMENT STRENGTH coding when a formal agreement between the belligerents specifically nullifies the earlier agreement. This is not always the case; for example, Russia signed a number of cease-fire agreements with the Central Powers in 1917 and 1918 but did not sign the Versailles Treaty, which ended World War I and might have subsumed these earlier treaties. In the ending agreements for World War II, some minor powers defeated by Germany during the war (such as Yugoslavia) were not signatories.

We followed Fortna's approach and coding rules in coding the institutions (cease-fire agreement) variables for our expanded data set. Using primary source and archival materials, we acquired the texts of almost all cease-fire and germane follow-up agreements over the 1914–2001 period that were not included in her original 1946–98 data set.[13] Overall, we coded the contents of more than 200 agreements,

11. We reran the analysis with a dummy variable for post–World War years of intrawar cease-fires. That dummy was insignificant, and its inclusion did not change the other results.

12. For a list of state deaths, see Fazal 2007, chap. 2. A related issue is what might be called "state fission," when a state breaks apart after war into separate, COW-recognized constituent components, as when post–World War II Germany became East Germany and West Germany. We generally assumed that all postwar components of a former belligerent were party to a war-ending settlement. We followed Fortna, however, in not including a post-1971 cease-fire between Bangladesh and either India or Pakistan. There is also the rarer instance of state fusion, when states combine to form a new state, as when East and West Germany reunified in 1990 to form Germany. For this case, we coded all cease-fire dyads involving East and West Germany as being right-censored in 1990. New cease-fire dyads involving Germany appear in 1990, and the new Germany dyads experience discontinuous intervals in the pre-1990 years.

13. For sources and data, see ⟨http://polisci.emory.edu/assets/reiter/datasets.htm⟩. Accessed 27 June 2008. Following Fortna, we did not include agreements not directly tied to the cease-fire, such as the Nuclear Non-Proliferation Treaty. We were unable to acquire the texts of the 1914 Tsingtau Treaty ending fighting between Germany and Japan and a 1942 follow-up agreement to the 1939 cease-fire ending the Nomonhan War between Japan and the Soviet Union. We coded these agreements based on descriptions in secondary sources.

and used 136 of these agreements to code our variables. Some materials were in French, German, and Japanese, which we had translated to English prior to coding. We read each agreement, and following Fortna, coded several variables, including FORMALISM, WITHDRAWAL OF FORCES, DEMILITARIZED ZONES (DMZ), ARMS CONTROL, PEACEKEEPING, EXTERNAL INVOLVEMENT, INTERNAL CONTROL, PARAGRAPH COUNT, CONFIDENCE BUILDING MEASURES, DISPUTE RESOLUTION, and an INDEX combining all these components into a single measure.[14]

We also collected new data on the Werner and Yuen BATTLE CONSISTENCY and INTERRUPTED WAR variables. Werner and Yuen coded these variables for all wars from 1946–97, and we coded the variable for the remaining wars in the 1914–2001 time period.[15] The BATTLE CONSISTENCY variable ranges from 0 to 1 in which higher values of the variable indicate one side more clearly dominating the war. Specifically, the variable is coded as the length of time from the beginning of the victor's last successful offensive to the end of the war, divided by the total length of the war. So, if the length of the war is 100 days, and the victor's last successful offensive begins on day 81, then the BATTLE CONSISTENCY variable is coded as .2. INTERRUPTED WAR is coded 1 if third-party pressure helps end the war and if the final settlement is signed after the cease-fire, otherwise coded 0.

To code foreign-imposed regime change, we used two different measures.[16] The ARCHIGOS data set (version 2.5) includes information on all national leaders for all states from 1875–2004.[17] We coded a state as having experienced FIRC if within a year of war's end a new leader comes to power, and its coding for the ARCHIGOS Exit variable is 4, indicating the leader was directly deposed by another state.[18] Werner collected data on FIRC from 1816–1980, coding an FIRC as having taken place following an interstate war if there was a change in leadership after war's end and within a year of war's end, if the victor helped determine the character of the new regime, if the new regime reflected foreign interference and not domestic forces for change, and if the new leadership constituted a genuine change in regime.[19] We updated her data set to 2001. There is general agreement between ARCHIGOS and Werner's data set.[20]

14. Fortna 2004.

15. We also discovered (and corrected) a data entry error for the battle consistency variable in the original Werner and Yuen data set for the post–Six Day War peace duration dyads (personal communication with Amy Yuen, 11 June 2007). When this error was corrected, the results in Werner and Yuen 2005 changed, such that the BATTLE CONSISTENCY variable for the 1946–97 time period was no longer significant, but Fortna's AGREEMENT STRENGTH variable became significant with the log-logistic model.

16. Owen 2002 codes FIRC attempts, not FIRC successes.

17. The data and codebook, developed by Hein Goemans, Kristian Gleditsch, and Giacomo Chiozza, are available at ⟨http://mail.rochester.edu/%7Ehgoemans/data.htm⟩. Accessed 27 June 2008.

18. ARCHIGOS also contains a variable on the manner of a leader's entry into power, ENTRY, coded 2 on a 0 to 2 scale if a leader comes to power through the intervention of another state. We elected not to use ENTRY because it misses some important FIRC cases, including post–World War II Germany and 1979 Uganda.

19. Werner 1996.

20. The correlation = .92 for all cease-fire dyads within our sample.

For both FIRC variables (ARCHIGOS and Werner), we coded the variable by dyad, such that there was only an FIRC if the victor in the dyad played a strong role in determining the loser's new regime. We coded the United States, United Kingdom, and France (but not the Soviet Union or any smaller World War II Allies such as Norway) as imposing an FIRC on West Germany; the Soviet Union (but not the United States, United Kingdom, France, or any smaller allies) as imposing an FIRC on East Germany; the United States (but not China, the Soviet Union, or United Kingdom) as imposing an FIRC on Japan after World War II; and the United Kingdom (but not France) as imposing an FIRC on Turkey in 1918 (Werner but not ARCHIGOS codes Turkey as suffering an FIRC in 1918).[21] This allowed for a more precise coding of the variable and for the reality that a newly imposed regime may be in the orbit of one ally but not another. In effect, it biased the results against our central hypothesis, as the long post–World War II spells of peace between West Germany and the Soviet Union, East Germany and the United States/United Kingdom/France, and Japan and the Soviet Union/China/United Kingdom cannot be attributed to FIRCs in our statistical analysis.

Following Fortna, we also included other independent variables. To measure changes in the balance of power, we used COW material capabilities data (version 3.02).[22] To measure whether the stakes in the war affect the duration of postwar peace, we used the International Crisis Behavior (ICB) version 7 stakes variable (we coded the variable for 1914–18, years not covered by ICB), coding the dummy variable as 1 if the ICB stakes variable coded a state's existence to be at stake.[23] To measure the costs of war, we collected data on battle deaths.[24] Notably, more battle dead may indicate more intense fighting and therefore the revelation of more information during war, so an information-oriented bargaining model of war might predict that wars with higher casualties ought to be associated with longer spells of postwar peace.[25] To measure contiguity, we used the COW Direct Contiguity Data Set (version 3.0).[26] For data on past history of conflict, we used Fortna's formula of MIDs/dyad age (MID data version 3.02).[27] To measure whether the war ended in a tie, we used COW data, though we needed to make several coding decisions of our own for the world wars.

We also included two control variables for regime type, a dummy variable coded 1 if both states in the dyad are democratic, and a second dummy variable coded 1 if only one state in the dyad is democratic. A state is deemed to be democratic if it

21. Turkey 1918 is perhaps Werner's most questionable FIRC coding, given that the Turkish leadership was not changed after the war ended. Our results do not change if we run the analysis with the Werner data, though with Turkey coded as not suffering an FIRC in 1918.

22. Singer, Bremer, and Stuckey 1972.

23. Available at ⟨http://www.cidcm.umd.edu/icb⟩. Accessed 27 June 2008. We coded World War II cease-fires ourselves.

24. Clodfelter 2002.

25. See Werner 1999; and Box-Steffensmeier, Reiter, and Zorn 2003.

26. Stinnett et al. 2002.

27. Ghosn, Palmer, and Bremer 2004.

scores five or higher on the Polity IV −10 TO +10 POLITY variable.[28] This tests the existing hypotheses that jointly democratic dyads are more peaceful than other dyads, and that dyads with one democracy and one nondemocracy are more war-prone than other dyads.

Results

Among our 188 cease-fires, there were fifty-six re-occurrences of war (failures), thirty-three FIRCs according to the ARCHIGOS data, and thirty-seven FIRCs according to the Werner data (all of the ARCHIGOS FIRCs were also Werner FIRCs). As a first cut, FIRCs are less likely to be associated with the breakdown of peace. For the ARCHIGOS data, when there was not an FIRC, the cease-fire eventually broke down into war in 35 percent of the cases (34 percent for the Werner data), but if there was an FIRC, the cease-fire eventually broke down into war in only 6 percent of the cases (11 percent for the Werner data).

The small number of FIRC cease-fires that did fail is not strong anecdotal evidence against hypothesis (1). The only two cases of FIRC cease-fire failure in the ARCHIGOS data set are the 1941 Germany-Yugoslavia and Italy-Yugoslavia cease-fire dyads failing in 1999, when Germany and Italy made marginal contributions to the NATO war effort in Kosovo. Neither of these two cases are strong evidence against hypothesis (1), because they are not instances of an FIRC failing to prevent the defeated state from violating its war-ending commitments, as Germany and Italy imposed the FIRC, and they were in the coalition that initiated war against Yugoslavia in 1999. Werner's data include two other FIRC-failure cases. The first, the 1956 failure of the Soviet Union-Hungary 1945 cease-fire, is only limited evidence against the theory. The cease-fire failed not because the FIRC failed to prevent the target (Hungary) from breaking the war-ending treaty and re-attacking the victor (the Soviet Union), but rather because the Soviet Union intervened to reinstall the Communist regime it had imposed. Notably, the non-Communist regime posed no threat to the Soviet Union. The second is the 1944 French re-entry into World War II following the 1940 France-Germany cease-fire. This is in one sense a counter-example to hypothesis (1), as the FIRC failed to prevent the defeated state (France) from re-attacking the victor (Germany). However, the FIRC failed not because there was an internal revolution within France or because the German-backed Vichy government suddenly changed course and attacked Germany. Rather, France was rescued by third parties, the United States and United Kingdom.

We used event history analysis. We first provided Kaplan-Meier nonparametric survival estimates (see Figure 1).

28. Polity data available at ⟨http://www.cidcm.umd.edu/polity⟩. Accessed 27 June 2008. We also tried cutpoints of six and higher and seven and higher. When we used these alternative cutpoints, STATA produces coefficient estimates but not standard error estimates.

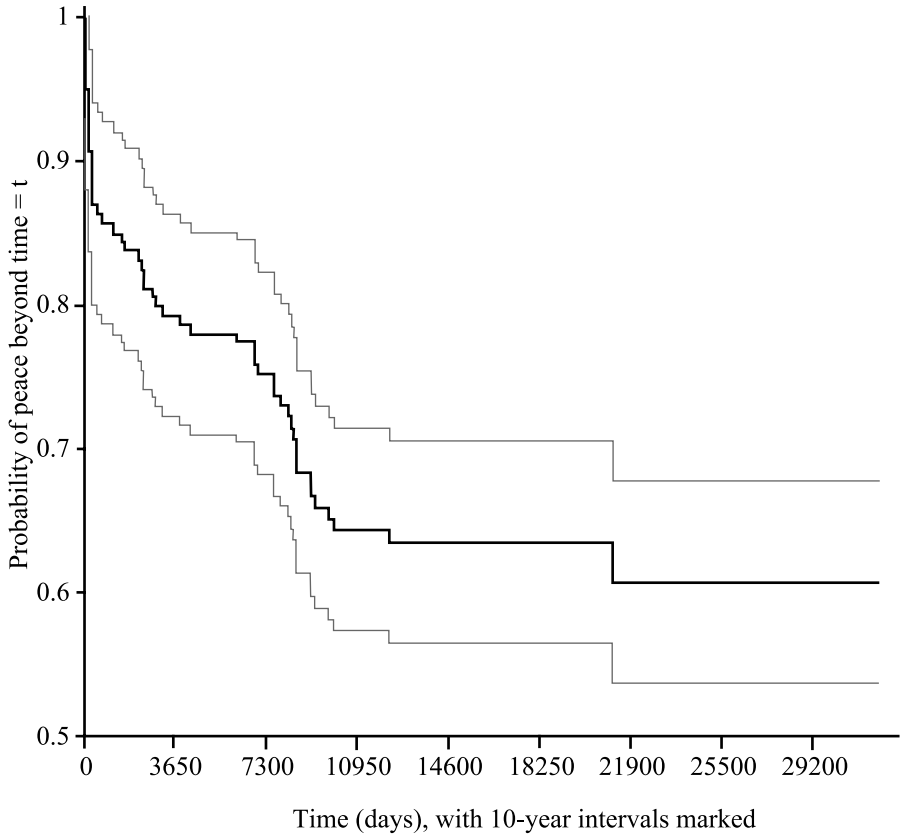

FIGURE 1. *Kaplan-Meier nonparametric survival function confidence intervals from Greenwood standard errors*

We used a semi-parametric Cox event history model. Figure 2 demonstrates that the shape of the model's baseline hazard is not monotonic, making the monotonic Weibull less attractive.

We found nonproportional hazards (NPH) for some of the independent variables. We employed a standard corrective to NPH, including interaction terms of a variable multiplied by the natural log of time elapsed when the diagnostics indicate that the variable has nonproportional effects. We employed the Efron method for handling ties. We used robust standard errors, clustering on peace durations resulting from the same conflicts.

Table 1 displays the results. Coefficients rather than hazard rates are reported. Negative coefficients mean that increases in the independent variable decrease the likelihood of failure and increase the duration of peace, whereas positive coeffi-

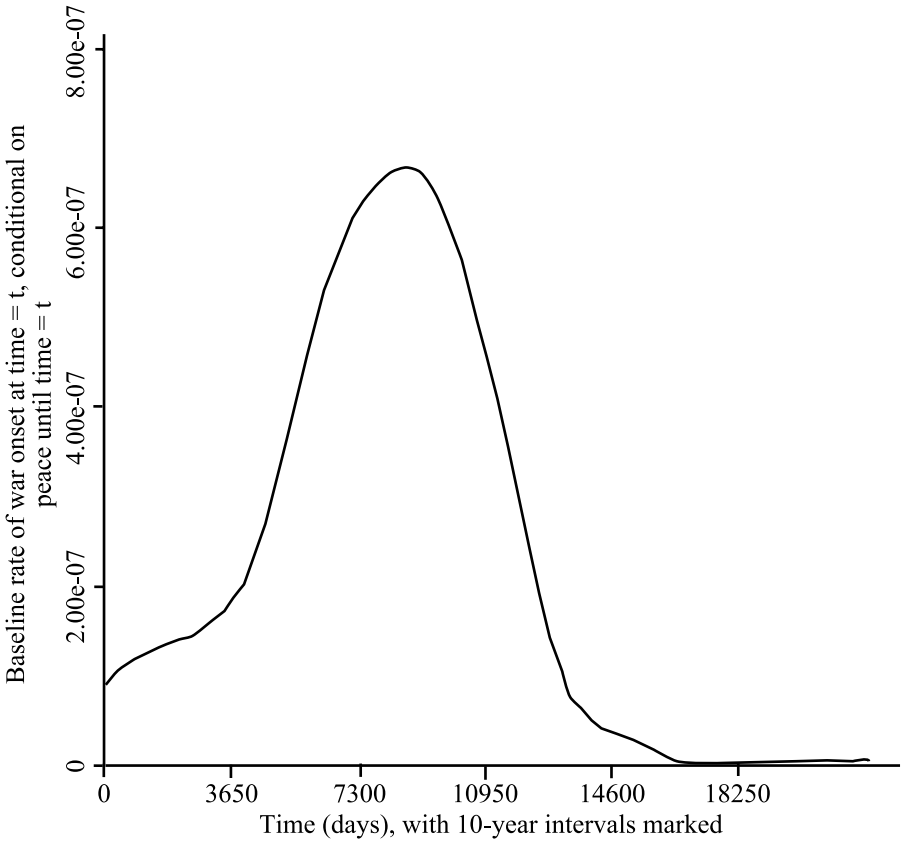

FIGURE 2. *Baseline hazard function from Table 1, Model 1*

cients have the opposite effects. For Model 1, we discovered NPH for variables FIRC, JOINT DEMOCRACY, and CONTIGUITY.

The results indicate strong support for hypothesis (1). If a war ends with an FIRC, then the postcease-fire peace will last significantly longer. Also, as indicated by the positive, statistically significant coefficient of the FIRC-time interaction term, the pacific effects of an FIRC diminish over time. The results are more mixed on the information variables. BATTLE CONSISTENCY is significant, but INTERRUPTED WAR and CHANGE IN CAPABILITIES are not. Among the other variables, only STAKES and WAR DEATHS are significant. JOINT DEMOCRACY is significant, but not in the predicted direction, though the net effect of JOINT DEMOCRACY shifts to the predicted direction when NPH are modeled. AGREEMENT STRENGTH is insignificant.

FIRC has high substantive significance. Setting other independent variables to their means, the chances of the cease-fire breaking down with an FIRC is essentially zero, a tiny fraction of the chances of a cease-fire breaking down if there is

TABLE 1. *Event history analysis of cease-fire duration, 1914–2001*

Variables	*Model 1 (ARCHIGOS data)*	*Model 2 (ARCHIGOS data)*	*Model 3 (Werner data)*	*Model 4 (ARCHIGOS data)*	*Model 5 (ARCHIGOS data)*	*Model 6 (ARCHIGOS data)*	*Model 7 (ARCHIGOS data, log-log model)*
FOREIGN-IMPOSED REGIME CHANGE	−161***	—	−6.18***	−162***	−146***	−153***	807**
	(29.3)		(1.76)	(29.8)	(20.4)	(28.7)	277
FIRC*ln(t)	16.8***	—	.673***	16.8***	15.2***	15.9***	−81.4**
	(3.03)		(.160)	(3.08)	(2.13)	(2.96)	(28.0)
PUPPET-FIRC	—	−161***	—	—	—	—	—
		(29.4)					
PUPPET-FIRC*ln(t)	—	16.8***	—	—	—	—	—
		(3.04)					
CHANGE IN CAPABILITIES	.272	.274	.166	.274	.0330	.280	−.294
	(.376)	(.376)	(.289)	(.374)	(.490)	(.378)	(.270)
BATTLE CONSISTENCY	−.796**	−.809**	−.654*	−.786**	−1.15**	−1.17***	.735
	(.336)	(.342)	(.387)	(.333)	(.465)	(.383)	(.549)
INTERRUPTED WAR	.00661	.00789	.139	.0237	.169	.139	−.0380
	(.377)	(.377)	(.446)	(.410)	(.652)	(.361)	(.400)
AGREEMENT STRENGTH	.0120	.0129	−.0390	—	—	—	.0260
	(.0865)	(.0863)	(.0892)				(.0785)
POST-45 AGREEMENT STRENGTH	—	—	—	−.00665	—	—	—
				(.0796)			
FORMALISM	—	—	—	—	.00236	—	—
					(.418)		
WITHDRAW	—	—	—	—	.505^{+}	—	—
					(.218)		
DMZ	—	—	—	—	−.0828	—	—
					(.250)		
ARMS CONTROL	—	—	—	—	.212	—	—
					(.194)		
PEACEKEEPING	—	—	—	—	.139	—	—
					(.422)		
EXTERNAL INVOLVEMENT	—	—	—	—	−.529	—	—
					(.438)		

INTERNAL CONTROL	—	—	—	—	−.220 (.263)	—	—
PARAGRAPH COUNT	—	—	—	—	−.233 (.486)	—	—
CONFIDENCE BUILDING MSRS.	—	—	—	—	−.455*** (.134)	−.433** (.186)	—
DISPUTE RESOLUTION	—	—	—	—	−.0514 (.334)	—	—
ONE DEMOCRACY	.204 (.434)	.201 (.436)	.275 (.385)	.203 (.434)	.0384 (.435)	.00749 (.392)	.0121 (.428)
JOINT DEMOCRACY	7.42** (3.15)	7.51** (3.14)	−1.98** (.677)	7.43** (3.12)	8.87** (3.71)	8.26** (3.39)	−7.18 (2.96)
JOINT DEMOCRACY*ln(t)	−1.65** (.605)	−1.68** (.600)	—	−1.65** (.598)	−1.98** (.707)	−1.81** (.671)	1.57** (.597)
TIE	.516 (.545)	.514 (.546)	.764 (.542)	.558 (.472)	.748 (.463)	.439 (.522)	−.563 (.515)
WAR DEATHS	−.230* (.101)	−.230* (.101)	−.171* (.0747)	−.228* (.105)	−.222* (.0977)	−.228* (.111)	−.185 (.121)
CONFLICT HISTORY	.339 (.457)	.345 (.460)	−5.40** (2.17)	.346 (.433)	.616 (.499)	.445 (.482)	−.335 (.432)
CONFLICT HISTORY*ln(t)	—	—	.785* (.342)	—	—	—	—
STAKES	.878* (.482)	.874* (.483)	.708* (.421)	.882* (.485)	1.46** (.495)	1.07* (.488)	−.775 (.501)
CONTIGUITY	7.42 (5.77)	7.47 (5.78)	1.35** (.453)	7.39 (5.81)	7.31 (5.51)	7.06 (5.73)	−7.25** (1.89)
CONTIGUITY*ln(t)	−.729 (.644)	−.734 (.644)	—	−.725 (.651)	−.689 (.613)	−.684 (.640)	.751*** (.242)
Constant	—	—	—	—	—	—	9.12*** (1.36)
Log pseudo likelihood	−206.4123	−206.54171	−221.04288	−206.4213	−194.47138	−202.46758	−160.97192

Notes: Robust standard errors are reported. All models are Cox models except for Model 6, which is a logistic model. 186 subjects, 6,368 observations. ***significant at .001 level; **significant at .01 level; *significant at .05 level. All one-tailed tests. $^{+}$significant at .05 level, but sign is not in predicted direction.

no FIRC. Over time, the pacifying effect of a FIRC decreases, but even after thirty years, the chances of an FIRC cease-fire breaking down are still only 1 percent of the chances of a non-FIRC cease-fire breaking down. The average pacifying effect of an FIRC disappears after about forty years.

Recall that hypothesis (2) specifies that the pacifying effect of FIRCs should decline over time when a puppet is installed, but not when democratic institutions are installed. We tested this hypothesis by separating out puppet-FIRCs from democracy-FIRCs. No democracy-FIRCs ended in failure, indicating that the pacification effect of democracy-FIRCs does not decline over time (results not reported because the absence of democracy-FIRC failure precludes the generation of estimates). However, inclusion of a dummy variable coded 1 only if there was a puppet-FIRC (as well as the associated NPH interaction variable) indicates that the pacification effect of puppet-FIRCs does decline over time, as the interaction is significant and positive (see Model 2). This provides support for hypothesis (2).

We carried out a number of specification tests. We reran the analysis with the Werner data (Model 3). Diagnostics indicated NPH only for FIRC and results generally did not change. We controlled for mutual possession of nuclear weapons, alliance (defense pacts) within the dyad, and multilateral war, but we found that none of these variables affected the other results. Only alliance was significant.

Some might argue that FIRCs are correlated with postwar peace duration only because FIRC is an indicator that the target's military has been crushed. Therefore, the real cause of peace is the destruction of the target's military, and, more, generally, an imbalance of military power. We reran the analysis with a measure of the annual preponderance of military expenditures. If the real cause of postwar peace is power imbalance and not FIRC, then this variable should be significant and negative, and perhaps the inclusion of preponderance should make FIRC insignificant. Military expenditure preponderance is more precise than aggregate power preponderance because states such as post–World War II Germany and Japan regained their industrial power but kept small militaries. Yet this variable was not statistically significant (results not reported) and had a low partial correlation with FIRC (.02). A measure of aggregate power preponderance was statistically significant, but incorrectly signed (imbalance making peace breakdown more likely). In short, the results suggest that the political changes of FIRC, not changes in the balance of power, extend postwar peace duration. The results also indicate that FIRC ensures peace through means other than just constraining the size of the defeated state's military.

Some might critique our inclusion of post–world war peace years for intra–world war cease-fires. We reran the analysis, including a dummy variable coded 1 for all such years. Such a new variable was not statistically significant, and its inclusion did not change the other results (results not reported). It was not significant when all FIRC cases are dropped.

The insignificance of the AGREEMENT STRENGTH variable is surprising and intriguing, as it contradicts Fortna's findings. The principle cause of AGREEMENT

STRENGTH's insignificance appears to be the data expansion. Replicating Fortna's Weibull proportional hazards model without the new independent variables used in Yuen and Werner and in this study did not make AGREEMENT STRENGTH statistically significant. To explore the possibility that AGREEMENT STRENGTH mattered only after 1945, we replaced the AGREEMENT STRENGTH variable with an interaction variable of AGREEMENT STRENGTH and a dummy variable coded 1 if the cease-fire occurred after 1945, 0 otherwise (Model 4). The interaction term was insignificant, and the other results were stable. Notably, mean AGREEMENT STRENGTH was larger in the pre-1946 period (4.78) than in the post-1945 period (4.14), though the difference was not statistically significant.

Another possibility is that AGREEMENT STRENGTH matters in the absence of FIRC. We reran the analysis, excluding all FIRC cases (and, of course, the FIRC variables). AGREEMENT STRENGTH remained insignificant. A model with an interaction between AGREEMENT STRENGTH and the post-1945 dummy variable on this reduced non-FIRC data set yielded an insignificant coefficient, as well (results not reported).

We may have found null results because AGREEMENT STRENGTH is an aggregate measure. Recall that AGREEMENT STRENGTH is a scale composed of ten components. Perhaps only some of these components affect peace duration. In Model 5, we dropped the AGREEMENT STRENGTH variable and instead included the individual components of AGREEMENT STRENGTH. The results for FIRC and the other variables were generally stable. Two of the ten components were statistically significant–WITHDRAWAL and CONFIDENCE BUILDING MEASURES (CBMs), though WITHDRAWAL was signed in the wrong direction. We reran ten models, each with only one of the AGREEMENT STRENGTH components. None of the individual AGREEMENT STRENGTH components were statistically significant in the predicted direction, except for the model with CBMs, reported as Model 6. Further analysis indicated that in the three nonzero categories of the CBMs variable, the only category that had a statistically significant effect on peace duration was for "on-site verification or aerial surveillance system established." Neither the "information exchange" or "hot-line" categories had statistically significant effects.

We used a Cox model because of its flexibility. However, the curve in Figure 2 also suggests a log-logistic event history model.[29] We reran Model 1 with a log-logistic event history model (Model 7), and the results did not change.

Even if FIRCs decrease the chances of interstate war reoccurring, they might increase the chances of intrastate war. However, using COW data on civil wars up to 1997, there are no examples of a civil war breaking out within a year of the end of an interstate war that ended with an FIRC, with the exception of civil war breaking out in Uganda in 1980.[30]

29. Box-Steffensmeier and Jones 2004. Diagnostics argued against a shared frailty model with ARCHIGOS data. The result is stable when we control for additional unobservable factors in a conditional gap time model.

30. Civil war data from Sarkees 2000.

A last issue is epiphenomenality. Fortna explored whether the AGREEMENT STRENGTH-peace duration result was spurious, as perhaps only states committed to peace anyway would sign a strong cease-fire agreement. She argued against this speculation, showing that belligerents sign stronger agreements when the breakdown of peace is more likely. We tested for the epiphenomenality of the CBMs finding by exploring whether stronger CBMs provisions are more likely to be signed when the belligerents are more confident that the cease-fire will not be broken. Following Fortna, we ran an ordered logit with each case's initial CBMs (range 0–3) score as the dependent variable, using each cease-fire as a single case.[31] We included all the variables in Model 1 significantly associated with peace duration, BATTLE CONSISTENCY, WAR DEATHS, STAKES, and FIRC, to see if CBMs provisions are stronger when structural conditions make postwar peace more durable. Only FIRC was significant, and it was negatively correlated with CBMs, indicating that CBMs are stronger without an FIRC, when postwar peace is more fragile (results not reported).

Another epiphenomenality concern is that the FIRC-peace duration result is spurious, and that FIRCs are only imposed when peace is likely to endure anyway. We conducted logit regression in which each cease-fire was a single case, the dependent variable was whether or not an FIRC was imposed, and the independent variables were the three factors found in our analysis to affect peace duration, WAR DEATHS, STAKES, and BATTLE CONSISTENCY. The results do not offer clear support for the epiphenomenality critique (results not reported). WAR DEATHS was not significant. STAKES was significant and positive, indicating that when stakes are higher and postwar peace is more likely to break down, FIRC is more likely (the opposite of the epiphenomenality prediction). BATTLE CONSISTENCY was positive and significant, indicating that FIRC is more likely when BATTLE CONSISTENCY is higher. This is evidence at first glance consistent with the epiphenomenality critique. However, the relationship between BATTLE CONSISTENCY and FIRC may be complex, as high stakes may encourage a state to pursue an FIRC, an aim, if achieved, that might require a long string of victorious battles (and hence a high BATTLE CONSISTENCY value). The U.S. campaign in World War II against Japan demonstrates this dynamic. Note also that running Model 1 without the FIRC variables changes the BATTLE CONSISTENCY coefficient slightly, from $-.80$ to $-.88$, indicating that FIRC is not purely a proxy for BATTLE CONSISTENCY.

Conclusion

This research note presents a new argument that FIRC makes postwar peace more durable. The results of the analysis of a new data set on postwar peace duration on all interstate cease-fires from 1914–2001 are striking: FIRCs substantially

31. Ordinary least squares (OLS) produced similar results.

increase postwar peace duration. Furthermore, most dimensions of peace treaties have essentially no effects on postwar peace duration. Of ten aspects of cease-fire agreements explored, only some types of CBMs, on-site verification and aerial surveillance, were statistically significant in the predicted direction. The results offer support for the importance of information factors, especially as the main purpose of CBMs is to provide information.

FIRC is a brutally effective tool for preventing war from re-erupting, especially compared with other tools such as peacekeeping. If a wartime belligerent fears that its adversary may break a war-ending agreement, that belligerent may become motivated to fight to the finish to impose an FIRC, solve a potential compliance problem, and make peace lasting.[32] States may "sink costs" or "tie hands" to increase compliance in areas such as trade, finance, and the environment, but only in the area of ensuring compliance with peace settlements is FIRC a realistic option.

Our understanding of exactly why FIRCs ensure peace is incomplete. We found that the pacifying effect declines over time for puppet-FIRCs but not for democracy-FIRCs, providing some evidence that imposed democratization has a pacifying effect, probably more important than any other possible FIRC mechanisms. Our analysis casts doubt on the possibility that the pacifying effect of FIRC is due solely to the substantial reduction of the size of the target's military, or to the creation of an alliance between the victor and FIRC-target. This suggests that an important reason why FIRCs successfully pacify is that they impose political changes on the target, but future research should explore this further. Future research should also explore other consequences of FIRC, such as whether FIRC makes a state generally more peaceful in its relations with all states, and whether FIRC makes civil wars more or less likely.

References

Berger, Thomas U. 1998. *Cultures of Antimilitarism: National Security in Germany and Japan*. Baltimore: Johns Hopkins University Press.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. Cambridge: Cambridge University Press.

Box-Steffensmeier, Janet M., Dan Reiter, and Christopher Zorn. 2003. Non-Proportional Hazards and Event History Analysis in International Relations. *Journal of Conflict Resolution* 47 (1):33–53.

Clodfelter, Michael. 2002. *Warfare and Armed Conflicts: A Statistical Reference to Casualty and Other Figures, 1500–2000*. 2d ed. Jefferson, N.C.: McFarland and Co.

Edelstein, David M. 2004. Occupational Hazards: Why Military Occupations Succeed or Fail. *International Security* 29 (1):49–91.

Fazal, Tanisha M. 2007. *State Death: The Politics and Geography of Conquest, Occupation, and Annexation*. Princeton, N.J.: Princeton University Press.

Fortna, Virginia Page. 2003. Scraps of Paper? Agreements and the Durability of Peace. *International Organization* 57 (2):337–72.

32. See Wolford, Reiter, and Carrubba 2008; and Reiter forthcoming.

———. 2004. *Peace Time: Cease-Fire Agreements and the Durability of Peace*. Princeton, N.J.: Princeton University Press.

Ghosn, Faten, Glenn Palmer, and Stuart Bremer. 2004. The MID3 Data Set, 1993–2001: Procedures, Coding Rules, and Description. *Conflict Management and Peace Science* 21 (2):133–54.

Ikenberry, G. John. 2000. *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order after Major Wars*. Princeton, N.J.: Princeton University Press.

Long, Stephen B. 2007. Retaliatory Behavior and the Violation of Formal Cease-Fire Agreements, 1948–1998. Unpublished manuscript, Kansas State University, Manhattan.

Owen, John M., IV. 2002. The Foreign Imposition of Domestic Institutions. *International Organization* 56 (2):375–409.

Reiter, Dan. Forthcoming. *The Search for Peace: Information, Commitment, and War Termination*. Princeton: Princeton University Press.

Reiter, Dan, and Allan Stam. 2006. Democracy, Peace, and War. In *Oxford Handbook of Political Economy*, edited by Barry R. Weingast and Donald Wittman, 869–80. London: Oxford University Press.

Sarkees, Meredith Reid. 2000. The Correlates of War Data on War: An Update to 1997. *Conflict Management and Peace Science* 18 (1):123–44.

Singer, J. David, Stuart Bremer, and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820–1965. In *Peace, War, and Numbers*, edited by Bruce Russett, 19–48. Beverly Hills, Calif.: Sage.

Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl, and Charles Gochman. 2002. The Correlates of War Project Direct Contiguity Data, Version 3. *Conflict Management and Peace Science* 19 (2):58–66.

Werner, Suzanne. 1996. Absolute and Limited War: The Possibility of Foreign-Imposed Regime Change. *International Interactions* 22 (1):67–88.

———. 1999. The Precarious Nature of Peace: Resolving the Issues, Enforcing the Settlement, and Renegotiating the Peace. *American Journal of Political Science* 41 (3):912–34.

Werner, Suzanne, and Amy Yuen. 2005. Making and Keeping Peace. *International Organization* 59 (2):261–92.

Wolford, Scott, Dan Reiter, and Clifford Carrubba. 2008. Information, Commitment, and War. Unpublished manuscript, Emory University, Atlanta, Ga.