### Title
The Dyadic Militarized Interstate Disputes (MIDs) Dataset Version 3.0: Logic, Characteristics, and Comparisons to Alternative Datasets




### Journal
Journal of Conflict Resolution, 63(3)

### ISSN
0022-0027



### Authors
Maoz, Zeev
Johnson, Paul L
Kaplan, Jasper
et al.



### Publication Date
2019-03-01

### DOI
10.1177/0022002718784158

Peer reviewed

Journal of Conflict Resolution
1-25

DOI: 10.1177/0022002718784158
journals.sagepub.com/home/jcr
SAGE

# The Dyadic Militarized Interstate Disputes (MIDs) Dataset Version 3.0: Logic, Characteristics, and Comparisons to Alternative Datasets

**Zeev Maoz[1], Paul L. Johnson[2], Jasper Kaplan[1], Fiona Ogunkoya[1], and Aaron P. Shreve[3]**

**Abstract**
We introduce the new, substantially updated, and revised version of the Dyadic Militarized Interstate Disputes (MIDs) dataset. We discuss the underlying logic of constructing dyadic MIDs and demonstrate that these operations generate significant differences between the actual occurrence and properties of MID dyads and those extracted from machine-generated programs such as EUGene, or from the MID participant dataset. We provide some descriptive measures of dyadic MIDs over the period of 1816 to 2010 and compare some of the key dyadic results on the correlates of MIDs using different datasets. We discuss the theoretical and empirical implications of our results.

**Keywords**
Militarized Interstate Disputes, dyadic conflict, dispute outcomes, dispute settlement, EUGene, initiator, target

[1]Department of Political Science, University of California, Davis, Davis, CA, USA
[2]Joint Warfare Analysis Center, United States Strategic Command, Dahlgren, VA, USA
[3]Department of Political Science, Randolph College, Lynchburg VA, USA

**Corresponding Author:**
Zeev Maoz, Department of Political Science, University of California, Davis, Davis, CA 95616, USA.
Email: zmaoz@ucdavis.edu

The Militarized Interstate Disputes (MIDs) dataset is one of the most widely used datasets in international relations.[1] Its first version appeared in the early 1980s (Gochman and Maoz 1984; Maoz 1982). Since then it has been updated repeatedly by different teams of authors as part of the Correlates of War (COW) project (Ghosn, Palmer, and Bremer 2004; Jones, Bremer, and Singer 1996). The most recent version (4.1) of the MID dataset (Palmer et al. 2015) covers the period of 1816 to 2010. A current data collection effort is underway to update the MID dataset to 2015.

A MID is defined as a set of incidents involving the deliberate, overt, government-sanctioned, and government-directed threat, display, or use of force between two or more states. Each pair of incidents belonging to a given dispute is separated by a short temporal interval (Gochman and Maoz 1984, 514). The sources mentioned above provide a detailed discussion of the concept, data collection process, and the coding rules. Many of these aspects are familiar to most conflict scholars, so there is no need to elaborate on them here.

The COW MID datasets contain two complete files covering the entire period of the project (1816–2010), and several files covering portions of the period. The complete files include: (1) the MID(A) file that provides data on the general parameters of each MID—the start and end dates, the number of participants, the highest level of hostility reached, MID reciprocity (whether both sides committed militarized actions), the fatality level, the outcome, and the form of settlement; (2) the MID(B) participant file provides information on each of the states that has participated in each MID, including its entry and exit dates, its level of hostility, whether it was on the initiator's or the target's side, its revisionist/status quo position, its fatality level, and the issues over which it disputed. The partial files include a dyadic MID(D) file covering the period of 1993 to 2001,[2] and an incident file covering the period of 1993 to 2010. The incident file (MID(I)) breaks up each of the MIDs into distinct incidents, each of which involves one or more initiators engaging in a single type of militarized action against one or more targets.[3] Another data collection effort nearing completion extends the incident dataset so that it covers the entire period of 1816 to 2010 (Gibler, Miller, and Little 2016).

A more general dyadic version (Maoz 2005) of the MID dataset has also seen a fair amount of use. This dataset was updated parallel to version 3.0 of the MID dataset, covering the 1816 to 2001 period. However, a large number of studies that focus on the dyadic level of analysis have extracted dyadic MID data either from a computerized algorithm that converts the MID participant file into a dyadic dataset or a more sophisticated conversion engine (the EUGene program: Bennett and Stam 2000). We show that this practice of extracting dyadic MIDs from the MID participant file may lead to multiple anomalies and significant biases. Any study that focuses on dyadic conflict or uses dyadic conflict as the building block for broader frameworks (e.g., network analysis) must rely on the dyadic MID dataset rather than on extracted dyadic MIDs from nondyadic datasets.

This study discusses the logic, coding procedures, and the statistical characteristics of the dyadic MIDs dataset, version 3.0. This dataset builds heavily on the MID dataset. It relies exclusively on the MID coding rules but applies them in a dyadic context. Thus, in the MID 4.2 dataset, 84.6 percent of all MIDs are dyadic, that is, they have only one state on each side of the MID. These "actual" dyadic MIDs—with very few exceptions—are identical in both datasets. However, the remaining 15 percent (377 MIDs in MID 4.2) may account for anywhere between 2 and 144 dyads.[4] Miscoding or misidentifying these dyads may cause quite a few differences in results and substantive inferences.[5]

This article is designed as follows. The next section discusses the basic logic of the dyadic MID dataset and the importance of relying on this dataset in any dyadic analysis of conflict. The third section discusses the coding principles of the dyadic MIDs. The fourth section includes a comparison between machine-extracted dyadic MID data and our hand-coded dataset. In the fifth section, we examine some analytic consequences of relying on different extraction methods of dyadic MIDs in terms of some key variables in MID research. We conclude by discussing some of the shortcomings and limitations of the dataset, as well as the implications of these data for conflict research.

## The Dyadic Logic of MIDs

As Palmer et al. (2015) indicate, the MID dataset is built from the ground up. The basic data collection unit is a militarized interstate incident (MII). An MII is an event where one or more states initiate(s) a specific act involving the threat, display, or use of force against one or more target states. The incident code sheet identifies the start date of the incident, the initiator(s) (but not the target) of the incident, and the type of military action taken.

These incidents are aggregated into MIDs. To be a part of the same MID, any two MIIs must share the same participants, the same location, and one or more issues. These incidents must be separated by a time interval not exceeding six months. If one of these conditions is violated, then these two incidents become separate MIDs.

A multilateral MID consists of a set of incidents that are connected in time, by issues, and location but involve more than two participants. Several participants are grouped into a single MID if there is a substantive link between incidents involving one set of participants and incidents involving another set of participants. This is so whether two states act jointly against a third (e.g., the coalition fighting Iraq in the First Gulf War) or whether one state initiates separate MIIs against several targets (e.g., the submarine attacks of Germany on British, French, and American ships during World War I [WWI]). Finally, dyads may belong to the same MID even if they involve different participants on the same side, but each member of the dyad has joined an ongoing MID between two principal participants (e.g., Italy and Greece in World War II [WWII]). The political context of such incidents suggests that they are part of the same MID. Deciding whether two incidents belong to the same MID requires some judgment even when the dispute is dyadic. When the dispute is multilateral, this task is exponentially more difficult. One must balance coding rules with judgment about whether two incidents involving different states belong to the same MID.

Why aggregate distinct incidents into a single MID? Why aggregate multiple dyadic incidents into a single MID? Ghosn, Palmer, and Bremer (2004, 139-40) offer a brief discussion of this issue, using the Cuban Missile Crisis and the crisis leading to WWII as examples, but they do not explicate the logic of aggregation. Other articles mention the key principles of the aggregation rule but do not discuss the underlying logic. It is therefore important to explicate the general idea of the MID concept. This feeds directly into the logic for disaggregating multilateral MIDs into valid dyadic components.

Briefly, a MID is a continuous conflictual interaction between or among states. The notion of continuous engagements suggests one of two types of strategic interactions. The first is an action–reaction process: state *i* initiates an MII against state *j*, the latter responds by launching another MII against the initiator, the initiator responds by an MII of its own, and so forth. The MID ends when this action–reaction chain ceases—either due to some formal agreement, a decisive victory by one of the disputants in a military clash, or mutual but uncoordinated decisions to stop hostilities. The incidents in this chain are causally connected in that each state responds to the previous action(s) of its adversary. The second type of strategic interaction involves a repeated set of MIIs initiated by one state against another, whether or not the latter responds by launching MIIs of its own. In this version, there may be little or no action–reaction, but we may still see a series of incidents that are linked by actors, issues, time, and space.[6]

In both cases, aggregating multiple MIIs into a single MID makes sense because incidents are not independent of each other. The same logic can be extended to multilateral MIDs. Consider an action–reaction process: state *i* initiates a MII against state *j*. State *k*—perhaps an ally of state *j*—intervenes by threatening to act against state *i* if *i* escalates. State *m*—perhaps an ally of state *i*—mobilizes to help state *i* if state *k* acts on its threat, and so forth. This is a typical model of many crises—including the July–August 1914 crisis that culminated in WWI (Tuchman 1962) or the May to June 1967 crisis that culminated in the Six Day War (Maoz 2006, 80-112). Consider another multilateral process: states *i*, *j*, and *k* form an alliance designed to attack a common enemy state *m*. To justify such an attack, they stage out a scenario that involves one of them issuing some provocation, which causes the target to respond, which then brings the other states into the fight. This was the scenario leading up to the First Balkan War of 1912 (Maoz 1990, 233-39) and to the Sinai War of 1956 (Maoz 2006, esp. 70-74). Here too, the process suggests that the actions of different participants are not independent of each other.

Yet, for a large number of research questions involving the analysis of individual states' conflict involvement or dyadic conflicts, the aggregated incidents at the participant level or the MID level pose several obstacles. For example, does WWII constitute one MID for Germany or does it constitute multiple MIDs (against

Poland, Britain, France, Greece, Yugoslavia, Russia, the United States, and so forth)?

In the First Gulf War (MID #3957 in the MID participant file), Iraq and Jordan are on side A (the side that initiated the MID).[7] States on side B include—among others—Kuwait, the United States, the UK, France, Saudi Arabia, Egypt, and Syria. All the states on side B are listed as having been involved in an interstate war. On side A, Iraq is also listed as having been involved in an interstate war, while Jordan is listed as having a hostility level of 1 (no militarized action). Several problems emerge when we take the participant file and convert it into a dyadic set of disputes: Iraq is correctly designated as the initiator of the overall multilateral MID due to its invasion of Kuwait. However, it did not initiate a MID against any of the other states on side B. The United States' decision to send air and ground forces to the Saudi–Iraqi border initiates the United States–Iraq dyadic MID. Each decision by any other state to deploy forces in the Persian Gulf and Saudi Arabia reflects a choice to initiate a MID against Iraq. The notion that Iraq initiated a MID against any of the states on side B except Kuwait is at odds with the historical record.

Moreover, extracting dyadic MIDs from the participant file would designate Jordan erroneously as an initiator of a dyadic MID against all other states in the list. In reality, Jordan was the target of threats by the United States, the UK, and France but did not commit any MII against other participants. Thus, it cannot be an initiator, nor can it be said to have joined the initiator. The outcome of the war is also problematic. The Gulf War ended in a military victory for the coalition against Iraq. Yet placing Kuwait on the winning side is misleading because Kuwait was under Iraqi occupation throughout the war. Likewise, the war ended in an imposed settlement between the coalition and Iraq, and yet it would be incorrect to classify the settlement as including Jordan. The dyadic MIDs in which Jordan was a target ended in a stalemate.[8]

These examples suggest that if we wish to examine the conflict behavior of individual states—either as participants in large-scale conflicts or as actors selecting particular enemies or responding to challenges originating from specific enemies—we need to consider such actions in a dyadic context. Dyadic analysis—despite some critiques of the dyadic design (Cranmer and Desmarais 2011, 2016)[9]—remains the most common unit of analysis in conflict research. Thus, we need to be careful about how we construct dyadic conflicts: (a) conflict dyads should be valid—that is, the members of the dyad have actually exchanged MIIs with each other—and (b) the variables associated with each dyad should apply specifically to that dyad.

The basic logic guiding the dyadic MID dataset is simple: a dyadic MID represents a specific set of MIIs exchanged between two states. These MIIs may be related to the MIIs exchanged by each of these states against another state or between two other states. However, just as we treat a strictly bilateral MID as a distinct case separate from other bilateral MIDs (even if those are between the same states), it makes sense to treat specific dyads in a MID as distinct—though not necessarily independent—from other dyads in the same MID. The requirement that

MIDs be overt, government-sanctioned, and government-directed (Gochman and Maoz 1984) rests on the assumption that states choose targets deliberately. Selection of specific actions during a dispute is meant to accomplish specific (tactical or strategic) goals but also to signal intent or resolve.

The implication for multilateral MIDs is that every participant who initiates one or more MIIs against one or more other states engages in a deliberate decision process about target selection, choice of specific action (or response to other states' actions), and termination-related decisions. The logic of "who does what to whom, when, and with what consequence" drives the coding of strictly dyadic MIDs. This logic must apply to each dyad in a multilateral MID. Consequently, a MID that involves more than two states requires researching each component dyad separately. This implies answering the following questions:

1. *Is the dyad a valid MID dyad?* Did one state initiate at least one MII targeting the other state? Answering this question is crucial for eliminating *invalid dyads*—that is, dyads consisting of two states on opposite sides of a multilateral MID, each of them interacting with other states in that MID but not with each other.
2. *What are the specific characteristics of the interactions between valid dyad members?* This requires collecting data on all of the MID variables (e.g., start and end dates, levels of hostility, outcome, settlement, issues, and fatalities) for each valid dyad on the basis of the interaction of dyad members with each other.

This logic is based on the realization that data regarding the actions of each individual participant in a MID do not necessarily generalize to the actions of that participant toward all other participants in the MID. The fact that the United States engaged in a war against Iraq in 1991 does not mean that it also engaged in a war against Jordan. Likewise, a characteristic of the MID as a whole—such as the outcome or the settlement—does not necessarily characterize all dyadic interactions in that MID: the outcome of the United States–Iraq dyad is not the same as the outcome of the Iraq–Kuwait dyad. Extracting dyads from the MID dataset or assigning the characteristics of the MID to all component dyads may create multiple coding errors resulting in erroneous inferences in analyses relying on dyadic interactions. We provide more evidence of this below.

The coding strategy we used for the dyadic MID dataset is as follows:

1. We rely fully on the MID incident coding rules and on the coding rules governing the aggregation of incidents into MIDs.[10]
2. For each MID that has more than two participants, we generate all possible dyadic combinations. For example, a MID with three states on side A (initiator) and two states on side B (target) consists of six ($3 \times 2$) possible dyads. These are the candidate dyadic MIDs.

3. For each candidate dyad, we research whether one dyad member initiated one or more MIIs against the other member. If not, the dyad is deleted from the dyadic MIDs dataset. For valid dyads, we proceed in the following manner:
   a. We identify the initiator and target of the dyadic MID. The initiator is that dyad member who launched the first MII against the other. As the example of the First Gulf War demonstrates, the role of some of the participants may switch from their role ascribed by the participant MID file, depending on who and when each of them took action toward the other dyad member.
   b. We define the start and end dates for the dyadic MID. The start date is the date in which the first incident of the dyadic MID occurred. The end date is the end date of the last incident of the MID.[11]
   c. We define the level of hostility for each of the states participating in the dyad as well as the highest level of hostility for the dyadic MID as a whole. This may differ significantly from one dyad to another in a multilateral MID.
   d. We define the issue over which the specific dyad clashed. For example, in the First Gulf war, the Iraq–Kuwait dyad was about territory. The United States–Jordan and the Iraq–Israel (MID #3956) dyads were about policy.
   e. We define, for each dyad, the outcome and settlement specific to that dyad. These may be different across dyads compared with the general MID outcome or settlement (as the example of Jordan's participation in the Gulf War demonstrates).
   f. We identify—if possible—the specific fatality level for each dyad member (and for the dyad as a whole) as part of the specific MIIs making up the dyadic interactions.
4. *Wars:* We inserted a fundamental change in the current version of the dyadic MID dataset with respect to wars. This change, we believe, matches more closely the definition of an interstate war from Singer and Small (1972; Small and Singer 1982) to Sarkees and Wayman (2010) and to alternative war datasets (Reiter, Stam, and Horowitz 2014). The COW project defines an interstate war as a series of sustained battles, between or among the military forces of two or more states, resulting in a total of 1,000 or more battle-related deaths. A state is considered a war participant if either of the two following conditions holds: (a) its troops suffered 100 or more battle-deaths or (b) it deployed 1,000 troops in battle-related activities (Sarkees and Wayman 2010, 39-42). Accordingly, we determined that the start date of a dyadic war must be the date of the first battle that makes up that series of battles. This must be distinguished from declarations of war that are not immediately followed by sustained combat operations. We define a battle in a dyadic sense as a military confrontation

between the armed forces of two states that meets *both* of the following conditions:

i. At least one side deploys 1,000 troops or more in battle-related activities, or
ii. At least one side suffers 100 or more battle-related deaths.

This enables us to avoid lower-level military clashes that are part of a MID but not part of a war in the sense discussed above and has important implications for dyadic MIDs. For example, the start date for Britain and France in WWII is listed as September 3, 1939. However, this is the date on which these states declared war on Germany. The "phony war" period from September 1939 to April to May 1940 involves incidents of fortification, troop movements, mobilizations, and limited naval clashes—all of which are coded as MIIs. However, the participation of Britain and France in WWII starts on April 9, 1940, when their troops clash with German troops in Norway (Clodfelter 2008; Weinberg 2005, 115-18). The same applies to war end dates. We code the end date of a dyadic war as the last day of the last battle in the war. This differs from some of the war end dates in the COW data.

MIDs that escalate into wars may receive different hostility scores for specific years. For example, the MID between the United States and Germany that starts in February 1916 (a German threat to declare war) and involves several German submarine attacks on American ships. WWI for that dyad starts on June 26, 1917 (when the first American force arrives in France). Thus, the record for the Germany–US MID gets a high hostility score of 4 (use of force) for 1916 and a score of 5 for 1917 to 1918.

5. We checked our coding against a modified COW dataset (Reiter, Stam, and Horowitz 2014). We explain and document where we differ from both the original COW dataset and the Reiter, Stam, and Horowitz (2014) data in the Dyadic MID documentation files. Many of these differences are due to the dyadic level of analysis, which induces changes not considered when coding wars without regard to specific dyadic interactions. However, quite a few differences stem from the focus on battles as the start and end points of war rather than on declarations of war for the start date of wars or peace/cease-fire agreements as indications of end dates.
6. Wars are a proper subset of MIDs: all wars are MIDs, but not all MIDs are wars. It is impossible that an event coded as a war in the COW dataset would not be coded as a war-MID (level of hostility 5 in the MID dataset), and vice versa. Each MID coded at level 5 must also be coded as a war in the war dataset. This conception has several important implications for all the variables in the MID and war datasets. For example, it is logically impossible that the start date of a war precede the start date of the corresponding MID. Nor can a war end date follow the end date of the corresponding MID. However, it is possible—and this happens quite often (see the WWII example above)—that the start date of the MID precedes the

start date of the war and the end date of the MID follows the end date of the war. Until now, there had existed a number of discrepancies between the COW MID and the COW datasets. We have resolved all of these discrepancies so that the two datasets are now perfectly compatible.[12]

## MID 4.2 and EUGene Extracted Dyadic MIDs Comparison

Quite a few studies have extracted dyadic MID datasets from the MID participant files. Other studies extracted dyadic MIDs via the EUGene program using the MID dataset (Bennett and Stam 2000).[13] Given that 85.5 percent of the MIDs are strictly dyadic (one state on each side), such a practice may not seem overly problematic. However, in other respects—and depending on the specific research question—extraction of dyadic MIDs from datasets that have not been formed with a dyadic conception in mind may generate significant biases.

To document this argument, we provide a brief comparison of the dyadic MID dataset (hereafter indexed by *d*) to two versions of the MID dataset extracted electronically: one from a computerized conversion of the MID participant and the dispute-level MID files (hereafter indexed by *m*) and another based on extraction of dyadic MIDs from the MID datasets via the EUGene engine (hereafter indexed by *E*).[14]

We start with a simple comparison of case match. The tabular results of this analysis are given in the Online Appendix. A pair of datasets agrees on a dyadic MID if (a) it matches in terms of (undirected) dyad identifier, and (b) it matches in terms of the year during which it was underway.[15] A case is designated an "agreement" case if it (a) appears in both datasets or (b) appears in neither (although it may appear in the third dataset). In general, there exists moderate agreement in case identification between the dyadic MID dataset and the extracted dyadic MIDs from EUGene or the MID 4.2 files. Specifically, the *d*–*E* comparison reveals a 68.7 percent agreement. Likewise, the *d*–*m* comparison suggests a 58.2 percent agreement. The *E*–*m* agreement rate is higher (93.8 percent) but not perfect.[16] The war agreement rates are slightly lower due to some cases identified as war dyads by one dataset but not by another and cases present in one dataset but missing in another. Figure 1 provides a temporal account of these case-identification disagreements.

The top-left panel shows the number of MID dyads underway per year. As we can see from this figure, the trends are highly correlated ($r_{d-m} = .948$, $r_{d-E} = .910$, $r_{m-E} = .949$). However, there are some significant discrepancies. The most visible ones are during the two world wars where both the MID 4.0 and the EUGene extractions considerably inflate the number of dyads. The other instance of significant discrepancy covers the period of 1984 to 1988. This is due to a difference in interpretation regarding the nature of MIDs during the well-known "tanker war" in the Persian Gulf.[17] Another discrepancy is also apparent during the First Gulf War.

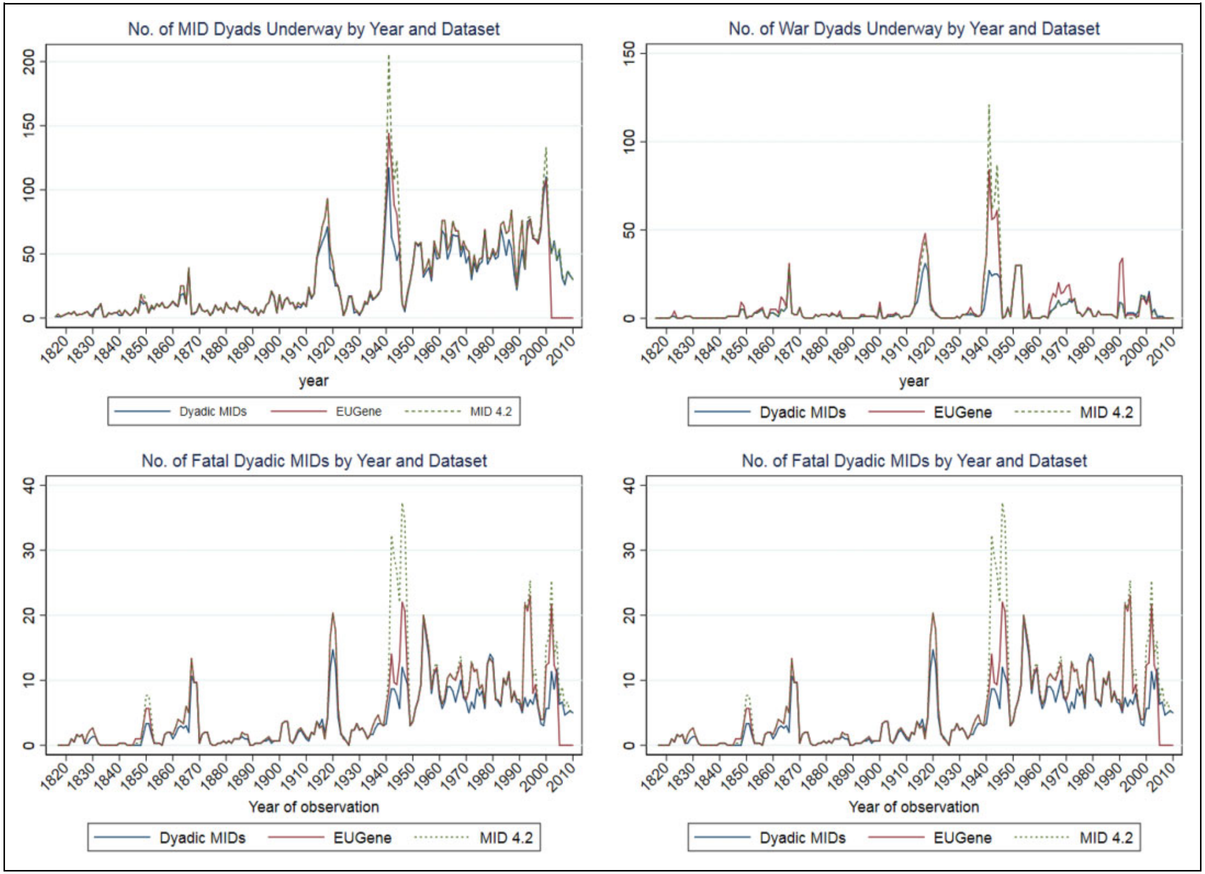


**Figure 1.** Comparison of militarized interstate dispute (MID) frequency and severity across datasets. Data reflect three-year moving averages of the number of MIDs from each dataset. As such, they present a more conservative estimate of the differences among datasets. EUGene data cover the period of 1816–2001 only (based on MID 3.0).

**Table 1.** Correlations of Key Variables across Datasets.

| Variables Compared | All MIDs: Dyadic MID-EU Gene | All MIDs: Dyadic MID-MID 4.2 | All MIDs: Eugene-MID 4.2 | Multilateral MIDs: Dyadic MID-EU Gene | Multilateral MIDs: Dyadic MID-MID 4.2 | Multilateral MIDs: Eugene-MID 4.2 |
|---|---|---|---|---|---|---|
| Role | .632 | .487 | .716 | .462 | .445 | .575 |
| Hostility | .451 | .393 | .627 | .273 | .299 | .442 |
| Escalation | .422 | .353 | .582 | .266 | .314 | .446 |
| Outcome | .307 | .076 | .077 | .194 | .016 | .044 |
| Settlement | .507 | .412 | .708 | .298 | .222 | .636 |
| Fatality | .563 | .419 | .764 | .513 | .570 | .607 |

*Note:* Correlations are tau-b scores. Shaded entries are not statistically significant at $p < 0.05$. MID = militarized interstate dispute.

Note that these comparisons involve "hard" tests: they include both the 85.5 percent strictly bilateral MIDs (which yield the same frequencies in all three datasets) and the multilateral MIDs that may yield different frequencies once they are converted into dyadic components.

Using the hostility scale developed by Maoz (1982, 217-25), we calculated annual hostility scores for each of the dyadic MIDs by dataset. The bottom-left panel of the figure shows the total severity of MIDs by dataset. While the severity scores are still very highly correlated, the correlations drop somewhat ($r_{d-m} = .934$, $r_{d-E} = .912$, $r_{m-E} = .962$).

The top-right part of the figure focuses on interstate wars. The major discrepancies here take place during the two World Wars, but other discrepancies exist in the 1970s and the more recent period (1998–2001). Here, we observe the lowest correlations between the dyadic MID dataset and the extracted dyadic datasets ($r_{d-m} = .830$, $r_{d-E} = .888$, $r_{m-E} = .977$).

Finally, several empirical conflict studies focus on "fatal MIDs-" that is, MIDs involving at least one battle-related death. Here, we find meaningful differences across datasets. These differences are particularly stark during periods dominated by large multilateral wars (e.g., the Crimean War, the two world wars, and the Gulf War), where quite a few invalid dyads, or valid dyads that do not involve fatalities, are prevalent. The correlations between datasets for fatal MIDs reflect these discrepancies ($r_{d-m} = 0.551$, $r_{d-E} = 0.579$, $r_{m-E} = 0.970$).

The high correlations between the annual MID frequencies of the dyadic dataset and the extracted MID datasets may suggest that the resulting population of cases appears to be quite similar. This inference, however, might very well depend on the kind of questions one wishes to address. When we examine specific characteristics of dyadic MIDs, significant differences emerge across datasets. We discuss several such issues.

Table 1 shows the correlations between different MID or dyadic attributes across the three datasets.[18] In general, the correlations between the dyadic MID dataset and the data extracted from EUGene or from the MID 4.2 participant file are low to moderate.[19] When examining only multilateral MIDs (i.e., MIDs that have more than one participant on a given side), these correlations drop—in some cases quite significantly. The general implication here is that while MID frequencies are highly correlated, the particular characteristics of MIDs or MID dyads are quite different depending on how MIDs are coded (in the case of the dyadic MID dataset) or extracted from nondyadic datasets.

The results of Table 1 support our contention that, for most purposes, extracting dyads from the original MID dataset might yield different results in empirical studies of conflict. Several important points emerge from this table. First, the correlations based on event frequencies (e.g., correlations between number of MIDs per year, number of wars per year; those reported in Figure 1) are much higher than the correlations between specific variables that characterize MIDs (those reported in Table 1). This is so because the latter variables are more likely to vary in a dyadic context than the overall MID frequencies. Second, correlations between different extraction methods on variables converted from the general MID file (e.g., outcome, settlement, and fatality level) into a dyadic context, are more likely to receive inaccurate values in dyadic analyses—as mentioned above. Third, correlations between different extraction methods are lower for multilateral MIDs than for bilateral ones, thus confirming our previous observation.

However, to establish the extent to which these differences between datasets affect substantive inferences about conflict behavior, we turn to a comparison of several analyses of MID escalation, outcomes (Maoz 1982, 1983; Partell and Palmer 1999; Sullivan 2007), and settlement patterns (Maoz 1984; Quackenbush and Venteicher 2008). The reasoning behind the selection of these specific comparison lies in the fact that these variables may receive significantly different values when extracted via computerized methods from the MID(A) and MID(B) dataset, compared to actually researched values for these variables in the dyadic MID dataset.

## Method

### *Dependent Variables*

1. MID escalation. This variable is coded as 1 for every dyadic MID that reached the war level and 0 for every MID that did not. This allows us to estimate factors that differentiate between MIDs that escalate to war and those that do not.
2. MID outcome. The MID dataset has a 10-category code for outcomes. We collapse this code to three outcomes: Victory for side A and yield of side B

are coded as a victory for side A. Victory for side B and yield of side A are coded as defeat for side A. The categories of stalemate, compromise, released, and unclear are coded as a draw. The outcome code for join Interstate War is assigned the outcome of the war for the participants (in the EUGene and MID 4.2 extractions).

3. Settlement. Here, we follow the MID 4.2 codes, with the exception of collapsing the "none" and "unclear" categories into a category of "stalemate" (Maoz 1984).

## Independent Variables

As noted, we use a standard set of covariates, which have been used with great regularity in studies of conflict. These include the following (with data sources specified for each):

1. Joint democracy. We use Maoz's (1998) coding of joint democracy, which allows for moving democracy cutoff points over time. This variable correlates (Yule's $Q = .981$; Tau-b = .653) with the more common designation of joint democracy when both states' net democracy scores (democracy–autocracy) are 6 or greater. We use the POLITY IV data (Marshall, Jaggers, and Gurr 2014). For the analysis of MID outcomes, we assign the democracy score to each state in the dyad.
2. Relative capabilities. This measures the extent to which the relative capabilities of dyad members deviate from parity. $\text{RELCAP} = \frac{\text{cinc}_{\text{max}}}{\text{cinc}_{\text{max}} + \text{cinc}_{\text{min}}}$ where $\text{cinc}_{\text{max}}$, $(\text{cinc}_{\text{min}})$ are the highest (lowest) Composite Index of National Capabilities (Correlates of War Project 2008; Singer 1990) score within a dyad.
3. Reputational status. We apply a modified version of Maoz's (2010) coding of reputational status. This variable assumes a value of 0 when both states are minor powers and 1 when one or both states are regional or global powers.
4. Log distance. We use the (logged) Gleditsch and Ward (2001) distance between capitals data.
5. Alliance. The alliance treaties and obligations provisions (ATOP) dataset (Leeds 2005) is supplemented by the COW Alliance dataset over the period of 2004 to 2010 (Gibler 2008). We code a dyad as 1 if it had a formal alliance treaty and 0 otherwise.
6. Log trade. We use the COW Trade dataset (Barbieri, Keshk, and Pollins 2009) and employ the log of total trade value (imports + exports) exchanged between dyad members.
7. Relative international governmental organization (IGO) membership. We employ the COW IGO dataset (Pevehouse, Nordstrom, and Wranke 2004). We use a ratio score of the joint IGO memberships of dyad members to each dyad member's IGO memberships as calculated by Maoz (2010, 42-47).

8. Cumulative number of MIDs/wars. The distribution of MIDs/wars over dyads is extremely left-skewed with a vast majority—over 90 percent—of the dyads having no MID and a tiny minority of dyads accounting for a vast majority of the MIDs (Maoz 2004, 2009; see also Online Appendix).
9. Hostility difference. Using the Maoz (1982) hostility scores, we generate the hostility difference score as: $HD_{ij} = \text{Host}_i - \text{Host}_j$. This score ranges from $-100$ to $+100$.
10. Highest hostility. For the analysis of settlement type, we use the highest hostility score in the MID according to each dataset.

### *Estimation*

The previous analyses reported in Figure 1 and Table 1 above were performed on undirected dyad-year and undirected dyad-MID observations, respectively. Likewise, the estimates of escalation, outcome, and settlements are based on MID-level data. This implies that each undirected MID (since dependent variables are the same for each pair *ij* and *ji* for a given MID) constitutes one observation. Given that start and end dates vary significantly across MIDs, we need to establish a fair comparison between the dyadic MID dataset and the extracted data. Accordingly, we calculated the last year of the MID coded by each of the datasets. We conduct the analyses of these dependent variables at the last year of the MID, according to each dataset's determination of that year.

Note that the *N*s differ significantly across datasets. This is so because MIDs that are identified by one dataset—and therefore are assigned (escalation, outcome, and settlement) scores—are not identified by another dataset. We assume that a user who relies on a given (extracted or actual) dyadic MID dataset estimates these equations using the MIDs identified by that particular dataset. We therefore apply the valid observations from each dataset as the basis for comparison, regardless whether these observations were coded as valid MIDs according to another dataset. We show in the Online Appendix how these datasets differ in terms of the assigned values of escalation, outcome, and settlement.

## 5. Results

Table 2 provides the comparative results of the analysis of MIDs by dataset.

Estimating MID escalation—the probability of a MID escalating to all-out war—results in some meaningful differences between datasets. Democracy has a significant dampening effect on escalation in the dyadic MID dataset, but it does not affect escalation in the EUGene or MID 4.2 datasets. This result is particularly significant in light of the democratic peace debate. The reason for this is primarily the number of wars between democratic dyads during WWII that were nonvalid dyads (e.g., the

**Table 2.** Comparison of Dyadic MID Models of Escalation (Undirected Dyads Only).

| | Dyadic MIDs | EUGene | MID 4.2 |
|---|---|---|---|
| Joint democracy | −2.535* | −0.052 | −0.019 |
| | (1.033) | (0.345) | (0.343) |
| Relative capabilities | 0.149** | −1.047** | −1.025** |
| | (0.025) | (0.124) | (0.129) |
| Log distance | −1.14** | −1.099** | −1.007** |
| | (0.383) | (0.361) | (0.373) |
| Regional–minor | 0.503** | 0.563** | 0.471** |
| | (0.147) | (0.143) | (0.149) |
| Major–minor | 0.254 | 0.977** | 0.985** |
| | (0.154) | (0.149) | (0.153) |
| Major–major | 0.647** | 1.213** | 1.265** |
| | (0.194) | (0.196) | (0.183) |
| Alliance | −0.659** | −0.586** | −0.867** |
| | (0.182) | (0.168) | (0.174) |
| Log trade | −0.221** | −0.291** | −0.295** |
| | (0.042) | (0.04) | (0.037) |
| Joint IGO membership | −0.495* | 0.205 | 0.21 |
| | (0.194) | (0.198) | (0.209) |
| Past MIDs | −0.002 | −0.031** | −0.06** |
| | (0.015) | (0.008) | (0.019) |
| Constant | −1.904** | −0.642 | −0.645 |
| | (0.363) | (0.34) | (0.348) |
| *N* | 4548 | 4014 | 4034 |
| $\chi^2$ | 226.093 | 189.779 | 216.114 |
| Pseudo $R^2$ | 0.071 | 0.067 | 0.078 |

*Note:* Highlighted cells indicate differences in significance or signs between the dyadic MID and the extracted MID datasets. Differences of *N*s are due to a different case identification by different datasets. See Online Appendix for comparisons. (1) The only war between democracies in the dyadic MID dataset was the Spanish–American War of 1898. In the EUGene and MID 4.2 datasets, there were six jointly democratic war dyads. In addition to the Spanish–American War, there were four dyads in WWII (the United States–Finland, the UK–Finland, Australia–Finland, and New Zealand–Finland). Our research shows that none of these were valid dyads, let alone war dyads. (2) We coded escalation only once for any given dyadic MIDs. This does not reflect differences in dates of escalation. For example, WWII is coded as escalation year of 1939 for UK–Germany and France–Germany in the EUGene and MID 4.2 datasets, but the escalation year for these dyads in the dyadic MID dataset is 1940. If we examined duration of escalation, differences would be more pronounced. MID = militarized interstate dispute; WWII = World War II.
*$p < .05$.
**$p < .01$.

United States–Finland, Canada–Finland, UK–Finland, Australia–Finland, and New Zealand–Finland).[20]

According to the EUGene and MID 4.2 datasets, we find that disputes between major and minor powers are more likely to escalate than minor–minor disputes. This is not the case, however, according to the dyadic MID dataset. By contrast, joint IGO membership reduces the probability of escalation according to the dyadic MID dataset, but this factor does not have a statistically significant effect on escalation according to the other two datasets. Finally, past MIDs reduce the probability of escalation according to the EUGene and dyadic MID dataset, but they do have a statistically significant effect on escalation in the dyadic MID dataset.

While these differences in terms of results are minor compared to the similar inferences that can be made about escalation with respect to other critical factors (e.g., alliance, trade, and distance), they are important theoretically. In particular, important aspects of the liberal peace argument (Maoz and Russett 1993; Russett and Oneal 2001) may be supported (via the dyadic MID dataset) or challenged (if using extracted MIDs from other datasets). We now turn to an analysis of MID outcomes and MID settlements. The results of these analyses are given in Table 3.

The results in Table 3 are substantively interesting, and therefore, it is instructive to discuss them more extensively. In the analysis of MID outcomes, we use the "draw" category as the baseline, so contrasts are based on differences between the win-or-lose categories and the draw category.

First, both the dyadic MID and the EUGene data suggest that a democratic initiator (or a joiner on the initiator's side) is significantly more likely to win than to draw. The effect of a democratic initiator on winning is not statistically significant according to the MID 4.2 dataset. Second, the dyadic MID dataset suggests that a democratic target is significantly less likely to win than to draw. However, this effect is not statistically significant according to the other two datasets. Third, a democratic initiator is less likely to lose than to draw according to the dyadic MID and EUGene dataset, but this effect is not statistically significant in the MID 4.2 dataset. Fourth, all three datasets suggest that as the hostility difference between the initiator and the target increases, the former is less likely to win.

Fifth, relative capabilities increase the probability of victory versus draws in the dyadic MID dataset. However, relative capabilities do not have a statistically significant impact on the probability of winning according to the two other datasets. By contrast, relative capabilities do not affect the probability of losing versus drawing. This holds for all three datasets. These two findings corroborate the results of studies that focus on resolve as the key factor that determines MID performance (Maoz 1983; Sullivan 2007, 2012).

Next, according to extracted datasets, alliance reduces the probability of victory as well as the probability of defeat. Hence allies that fight each other are significantly more likely to draw than end their MID in a decisive outcome. This result stems from the extracted (EUGEne and MID 4.2) datasets. However, the dyadic MID dataset does not support the first part of this contention. Specifically, MIDs between allies are not more likely to end in a victory for one of the allies, but they are more likely to end in a draw than in the defeat of one of the allies.

**Table 3.** Comparison of MID Outcomes and Settlement Type by Dataset (Directed Dyads).

| | MID Outcomes | | | | Settlement | | |
|---|---|---|---|---|---|---|---|
| | Dyadic MID | EUGene | MID 4.2 | | Dyadic MID | EUGene | MID 4.2 |
| | Win vs. draw | | | | Agreement vs. Stalemate | | |
| Democratic initiator | 0.385** | 0.436** | 0.209 | Joint democracy | 0.763** | 0.466 | 0.568* |
| | (0.128) | (0.128) | (0.128) | | (0.227) | (0.277) | (0.28) |
| Democratic target | −0.885** | −0.3 | −0.258 | Reputational status | 0.14** | 0.032 | 0.086 |
| | (0.208) | (0.165) | (0.156) | | (0.05) | (0.054) | (0.055) |
| Relative capabilities | 0.982** | 0.361 | 0.117 | Relative capabilities | −0.987** | −0.588 | −0.691 |
| | (0.191) | (0.189) | (0.184) | | (0.329) | (0.361) | (0.37) |
| Alliance | −0.255 | −0.696** | −0.844** | Alliance | −0.035 | −0.135 | −0.14 |
| | (0.152) | (0.162) | (0.158) | | (0.121) | (0.134) | (0.135) |
| Log trade | −0.174** | −0.061 | −0.138** | Log trade | −0.146** | −0.099* | −0.148** |
| | (0.048) | (0.046) | (0.042) | | (0.037) | (0.042) | (0.04) |
| Joint IGO membership | −0.003 | −0.061 | −0.028 | Joint IGO membership | −0.093 | 0.216 | 0.145 |
| | (0.204) | (0.21) | (0.211) | | (0.179) | (0.207) | (0.216) |
| Hostility difference A or B | 0.01** | 0.007** | 0.005** | Maximum hostility | 0.022** | 0.025** | 0.026** |
| | (0.001) | (0.001) | (0.001) | | (0.003) | (0.003) | (0.004) |
| Past MIDs | −0.026 | −0.043** | −0.051** | Past MIDs | −0.012** | −0.006 | −0.008 |
| | (0.018) | (0.009) | (0.018) | | (0.004) | (0.004) | (0.004) |
| Constant | −2.323** | −1.446** | −1.177** | Constant | −2.171** | −2.775** | −2.813** |
| | (0.189) | (0.192) | (0.188) | | (0.366) | (0.414) | (0.441) |
| | Lose vs. draw | | | | Imposed settlement vs. stalemate | | |
| Democratic initiator | −0.272* | −0.552** | 0.125 | Joint democracy | −1.727* | 0.165 | 0.349 |
| | (0.137) | (0.164) | (0.155) | | (0.727) | (0.37) | (0.385) |
| Democratic target | 0.004 | −0.295 | −0.345 | Reputational status | 0.254** | 0.032 | 0.086 |
| | (0.147) | (0.177) | (0.197) | | (0.061) | (0.064) | (0.068) |

*(continued)*

**Table 3.** (continued)

| | MID Outcomes | | | | Settlement | | |
|---|---|---|---|---|---|---|---|
| | Dyadic MID | EUGene | MID 4.2 | | Dyadic MID | EUGene | MID 4.2 |
| Relative capabilities | −0.178 | −0.263 | 0.109 | Relative capabilities | −0.691 | −0.258 | −0.343 |
| | (0.174) | (0.209) | (0.224) | | (0.396) | (0.419) | (0.444) |
| Alliance | −0.572** | −0.595** | −0.606** | Alliance | 0.163 | −0.214 | −0.163 |
| | (0.157) | (0.174) | (0.18) | | (0.147) | (0.165) | (0.171) |
| Log trade | −0.064 | −0.002 | −0.078 | Log trade | −0.174** | −0.019 | −0.025 |
| | (0.046) | (0.052) | (0.051) | | (0.047) | (0.051) | (0.052) |
| Joint IGO membership | −0.288 | 0.282 | 0.05 | Joint IGO membership | −0.235 | 0.507* | 0.555* |
| | (0.2) | (0.254) | (0.267) | | (0.202) | (0.23) | (0.251) |
| Hostility difference A/B | −0.005** | −0.005** | −0.002 | Maximum hostility | 0.057** | 0.1** | 0.129** |
| | (0.001) | (0.002) | (0.002) | | (0.004) | (0.005) | (0.006) |
| Past MIDs | −0.005 | −0.022* | −0.035 | Past MIDs | −0.035** | −0.03** | −0.035** |
| | (0.017) | (0.009) | (0.02) | | (0.007) | (0.006) | (0.008) |
| Constant | −1.221** | −1.444** | −1.778** | Constant | −5.331** | −9.409** | −12.051** |
| | (0.163) | (0.214) | (0.236) | | (0.499) | (0.6) | (0.701) |
| *N* | 2881 | 2446 | 2515 | *N* | 3348 | 2919 | 2859 |
| $\chi^2$ | 241.719 | 178.892 | 114.970 | $\chi^2$ | 456.215 | 747.301 | 926.040 |
| Pseudo $R^2$ | 0.056 | 0.048 | 0.031 | Pseudo $R^2$ | 0.085 | 0.152 | 0.193 |

*Note:* Highlighted cells indicate differences in significance or signs between the dyadic MID and the extracted MID datasets. Differences of *N*s are due to a different case identification by different datasets (see Online Appendix for comparisons). MID = militarized interstate dispute.

Seventh, past MIDs increase the probability of a draw (reduce the probability of either victory or defeat) according to the EUGene dataset. Past MIDs reduce the probability of victory according to the MID 34.2 dataset, but past MIDs have no statistically significant effect on the outcomes of dyadic MIDs, and they do not affect the probability of defeat according to the MID 4.2 dataset. Joint trade reduces the probability of wins relative to draw, but it does not affect the probability of losing versus drawing. The latter part of this result is robust across datasets. However, the former part of the result is based only on the dyadic MID and MID 4.2 datasets. Finally, according to all three datasets, joint IGO membership does not have a statistically significant effect on the type of outcome.

We now turn to determinants of dispute settlement. Here, the stalemate category serves as the baseline. First, jointly democratic dyads are significantly more likely to terminate their MIDs via a negotiated settlement compared to a stalemate. This is based on the results of the dyadic MID and MID 4.2 dataset, but it does not hold for the EUGene dataset. Joint democracy has a significant dampening impact on the probability of imposed settlement according to the dyadic MID dataset, but this also does not hold for the extracted MID datasets.

Reputational status increases both the probability of a negotiated settlement and the probability of an imposed settlement relative to stalemate according to the dyadic MID dataset, but neither effect is statistically significant according to the extracted MID datasets.

Relative capability reduces the probability of a negotiated settlement according to the dyadic MID dataset. However, capabilities do not have a statistically significant effect on settlement type according to the EUGene and MID 4.2 datasets.

Trade levels reduce the probability of a negotiated settlement according to all three datasets. According to the dyadic MID dataset, trade also reduces the probability of or an imposed settlement, but trade does not have a statistically significant effect on an imposed settlement according to the other two datasets.

All in all, we find some meaningful variations in results across datasets when analyzing MID outcomes and settlement. These variations in the results are also due to the fact that the correspondence between datasets is much lower in the case of MID outcome and MID settlement variables compared to the MID occurrence variables.

In summary, a higher-resolution comparison of the dyadic MID dataset with the extracted dyadic MID datasets suggests significant differences in terms of the distribution of key variables such as hostility, escalation, outcome, and settlement. These distributional differences yield some meaningful differences in inferences when we apply various models to estimate these characteristics of MIDs. The general point of this exercise is not to show that one dataset is "better" or yields more "accurate" results than another. On the contrary, the results of these comparative analyses are generally quite similar across datasets. However, differences do exist, and from some theoretical and empirical perspectives, these differences may yield significantly different inferences about the factors that affect such things as initiation, escalation, outcomes, and settlements.

## Discussion

The dyadic MID dataset builds exclusively on the MID datasets. It relies on a fundamental premise: decisions to launch militarized actions are important in that policy-makers take such interactions very seriously. When a state representative issues a specific threat, or when the state uses limited force against another state, the potential for escalation is part of the strategic calculus of the state's leader. This is why actions have been coded in the MID dataset. The significance of the dyadic MID—and where it extends this logic—is that such decisions are more complex in the case of multilateral MIDs. A state joining an existing MID or a state that acts in concert with other states is not only a participant in a MID in general. It must decide which target to select and whether to launch similar actions against all targets. It must decide when to enter against a specific target, when to end the action against a specific target, and how to end the MID. These decisions may differ from one state to another even though they are part of the same side in a MID.

This has important implications. Even though multilateral MIDs constitute only 15 percent of all MIDs, they account for nearly 50 percent of the dyads that participated in MIDs. The multi-actor nature of these MIDs is related to the level of hostility (multilateral MIDs are significantly more likely to involve high levels of hostility), to duration (multilateral MIDs last almost twice as many days as bilateral ones), to outcome (multilateral MIDs are far more likely to end in a win-or-lose outcome than in a draw), and to settlement type (multilateral MIDs are more likely to end in an imposed settlement than bilateral MIDs).

Models of system-level or monadic conflicts are also sensitive to the way we code and aggregate MIDs. As noted, a MID is first and foremost a dyadic affair: a state A commits one or more MIIs toward a state B, and state B may or may not respond to the actions of A. At the monadic level, it makes a difference whether a state was a participant in a dispute (as the MID participant file stipulates) against a single enemy or against several enemies. At the systemic level, it makes a difference whether—in a given year—there were five MIDs, each of them bilateral, or whether—in another year—there were also five MIDs, but some of the MIDs were multilateral. For example, the correlation between the number of dyadic MIDs and the number of MIDs over the period of 1816 to 2010 is 0.778, which is high but not overwhelming.

The implication is that, regardless of whether one studies monadic, dyadic, or systemic conflict using the MID datasets, it is important to (a) spell out the assumptions underlying the selection of the dependent variables even when they are extracted from the same data source, (b) understand the differences between a careful coding of individual dyads and an algorithm-based extraction of dyads from participant or MID-based datasets, and (c) realize that the selection of a specific approach for extracting measures of conflict sometimes has important implications for inference.

What we aimed to do in this project was to provide a logical justification for a careful approach to coding dyadic MIDs, not as a substitute, but rather as a supplement to the existing and widely used MID datasets. The dataset, as all COW datasets, is free, documented, and—just as with any data collection effort—may have problems and inaccuracies. We are open to comments, suggestions, and criticisms and welcome users to share these with us.

## Authors' Note

The Online Appendix and replication materials, including datasets, codebook, and code files are available on the *JCR* replication website as well as on http://www.zmaoz.ucdavis.edu.

## Acknowledgments

We would like to thank our colleagues of the SPINS project at UC Davis, especially to Ireen Litvak-Zur and Jaime Jackson for their support of this project. We also thank Scott Bennett, Glenn Palmer, and the MID 4.2 team, as well as Douglas Gibler, Steven Miller, and Erin Little for discussions about the data collection.

## Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by Minerva Grant #W911NF-15-1-0502 of the Army Research Office, the United States Department of Defense. Neither the ARO nor the DOD is responsible for the views expressed herein.

## Notes

1. The Militarized Interstate Dispute (MID) dataset is by far the most downloaded dataset of all Correlates of War (COW) datasets; it averaged over 1,200 downloads per month over the period of 2015 to 2017. (Data source to be provided on request.) The second-most downloaded dataset—the war dataset—averages about 400 downloads per month over the same period.
2. The dyadic MID(D) and the dyadic MID dataset are identical within the common time frame (1993-2001) they cover jointly. However, MID(D) is a small subset of the dyadic MID dataset. Also MID(D) does not cover most of the variables concerning dispute outcomes, settlement, and dyadic roles.
3. A file with narratives of all MIDs over the period of 1993 to 2010 is also available on the COW website.

4. Strictly dyadic MIDs account for only 47.9 percent of all dyadic MIDs, when automatically extracting from the MID participant dataset. In practice, strictly dyadic MIDs account for 51.5 percent of all dyadic MIDs based on the current dyadic MID dataset. Hence, when using a dyadic design, the fact that 85 percent of all MIDs in the MID(B) participant file are strictly dyadic grossly misrepresents the weight of dyadic MIDs in the overall MID population.
5. Note that the current version of the MID dataset (version 4.2—published simultaneously with the dyadic MID dataset) reflects multiple changes—including dropping of several hundred MIDs and merging of a few dozen others—due to a series of cleaning operations following by a review of the MID dataset by a group at the University of Alabama (Gibler, Miller, and Little 2016) and following a workshop at UC Davis in winter 2014.
6. Note that the incident dataset (MID) that covers the period of 1993 to 2010 is not very helpful in this regard as it contains only information about the state that initiated the MII constituting each incident in a MID, but it does not provide any information about the target(s) of that incident. This means that in a multilateral MID involving more than one state on each side of the MID, we do not have an easy way of determining the actual dyad in that incident (i.e., if the dates of participation of several states on each side overlap). So that extraction of dyadic MIDs from the incident dataset is quite problematic.
7. But see the cautionary remarks about the meaning of sides A and B in the MID dataset (Palmer et al. 2015).
8. One of the key problems in this sense involves attempts to combine the MID participant file (MID(B)) with the general MID file (MID(A)). If one wanted to examine the factors that affect the outcome of a MID in a dyadic context (e.g., who won), combining the MID outcome from the MID(A) file with the MID participant file may yield multiple flawed inferences. For example, if one were to combine the MID participant file with the general (MID(A)) file for the Poland–Germany dyad, one would infer that Poland defeated Germany in WWII. (This is, incidentally, the coding of this dyad in the *War Trap*; Bueno de Mesquita 1981).
9. This critique notwithstanding, the dyad remains the foundational building block of networks; without dyadic data, empirical network analysis is impossible (Maoz 2010).
10. The MID codebook and explanation of the coding rules for incidents and MIDs are given at http://correlatesofwar.org/data-sets/MIDs.
11. As is the case with the MID coding rules, some incidents (e.g., occupation of territory) end six months after they started if not followed by another MII.
12. A separate version of the dyadic war dataset is posted on the COW website (http://www.correlatesofwar.org).
13. Over the period of 2006 to 2017, we found forty-three published papers that extract dyadic MIDs either from the MID participant dataset or from the MID version of EUGene. A list of these papers is provided in the Online Appendix.
14. In the replication package, we provide Stata do files that document the extraction process.
15. This is a "lenient" matching strategy that maximizes the potential agreement between datasets. More "restrictive" matching requires directed dyad id (which may change if we move states from side A to side B in a multilateral MID), identical dispute numbers

(which may change in the dyadic MID dataset compared to the other two datasets), and/or matching of both start and end dates. These may also differ in a dyadic context and thus increase disagreement rates between the dyadic MID and the EUGene or MID 4.2 datasets.

16. Since the EUGene dataset ends in 2001, the comparisons with that dataset are restricted to the period of 1816 to 2001. The comparison between the dyadic MID dataset and the MID 4.2 dataset covers the entire period of 1816 to 2010.
17. This is explained in the documentation provided in the replication materials. However, briefly, the dyadic MID dataset does not consider most of the seizures by Iraq and Iran of ships sailing under different flags of African, Asian, or Latin American nations as valid MIDs. Most of these MIDs are either aggregated or deleted altogether. By contrast, the participant MID dataset considers each such seizure to constitute a dyadic MID between the initiator (typically Iran or Iraq) and the country whose flag is presented on the seized ship if a formal protest was issued by that country against the initiator. See also Gibler, Miller, and Little's (2016) criticism of the coding rules underlying seizures.
18. The actual contingency tables are given in the Online Appendix.
19. Even if we focus only on dyads that are deemed valid in any pair of dataset, correlations are only slightly larger than in Table 1.
20. The dyadic MID dataset identifies one jointly democratic war dyad—the Spanish American War of 1898, which is also identified by the other two datasets.

## Supplemental Material

Supplementary material is available for this article online.

## References

Barbieri, Katherine, Omar M. G. Keshk, and Brian Pollins. 2009. "Trading Data: Evaluating Our Assumptions and Coding Rules." *Conflict Management and Peace Science* 26: 471-91.

Bennett, D. Scott, and Alan Stam III. 2000. "Eugene: A Conceptual Manual." *International Interactions* 26:179-204.

Bueno de Mesquita, Bruce. 1981. *The War Trap*. New Haven, CT: Yale University Press.

Clodfelter, Michael. 2008. *Warfare and Armed Conflicts: A Statistical Encyclopedia of Casualty and Other Figures, 1494–2007*. Jefferson, NC: MacFarland.

Correlates of War Project 2008. 2008. *National Material Capabilities V4.0*. Correlates of War Project 2008. Accessed June 14, 2018. http://www.correlatesofwar.org/.

Cranmer, Skyler J., and Bruce A. Desmarais. 2011. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19:66-86.

Cranmer, Skyler J., and Bruce A. Desmarais. 2016. "A Critique of Dyadic Design." *International Studies Quarterly* 60 (2): 355-62.

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.

Gibler, Douglas M. 2008. *International Military Alliances, 1648–2008*. Washington, DC: CQ Press.

Gibler, Douglas M., Steven V. Miller, and Erin K. Little. 2016. “An Analysis of the Militarized Interstate Disputes Dataset, 1816–2001.” *International Studies Quarterly* 60 (4): 719-30.

Gleditsch, Kristian S., and Michael D. Ward. 2001. “Measuring Space: A Minimum Distance Database.” *Journal of Peace Research* 38 (4): 749-68.

Gochman, Charles S., and Zeev Maoz. 1984. “Militarized Interstate Disputes, 1816–1976: Procedures, Patterns, Insights.” *Journal of Conflict Resolution* 29:585-615.

Jones, Daniel M., Stuart Bremer, and J. David Singer. 1996. “Militarized Disputes, 1816–1992: Rationale Coding Rules and Empirical Patterns.” *Conflict Management and Peace Science* 15:163-212.

Leeds, Brett Ashley. 2005. “Alliance Treaty and Obligations Provisions (Atop) Codebook.” Rice University. Accessed June 14, 2018. http://atop.rice.edu/home.

Maoz, Zeev. 1982. *Paths to Conflict: Interstate Dispute Initiation, 1816–1976*. Boulder, CO: Westview Press.

Maoz, Zeev. 1983. “Resolve, Capabilities, and the Outcomes of Interstate Disputes, 1816–1976.” *Journal of Conflict Resolution* 27:195-229.

Maoz, Zeev. 1984. “Peace by Empire? Conflict Outcomes and International Stability, 1816–1976.” *Journal of Peace Research* 21 (3): 227-41.

Maoz, Zeev. 1990. *Paradoxes of War: On the Art of National Self-entrapment*. Boston, MA: Unwin Hyman.

Maoz, Zeev. 1998. “Realist and Cultural Critiques of the Democratic Peace: A Theoretical and Empirical Reassessment.” *International Interactions* 24:3-89.

Maoz, Zeev. 2004. “Pacifism and Fightaholism in International Politics: A Structural History of National and Dyadic Conflict, 1816–1992.” *International Studies Review* 6:107-33.

Maoz, Zeev. 2005. “Dyadic MID Dataset, 2.0.” Accessed June 14, 2018. http://maoz.ucdavis.edu/datasets.html.

Maoz, Zeev. 2006. *Defending the Holy Land: A Critical Analysis of Israel’s Security and Foreign Policy*. Ann Arbor: University of Michigan Press.

Maoz, Zeev. 2009. “Primed to Fight: The Can/Must Syndrome and the Conflict Proneness of Nations.” *Conflict Management and Peace Science* 26:1-26.

Maoz, Zeev. 2010. *Networks of Nations: The Evolution, Structure, and Impact of International Networks, 1816–2001*. New York: Cambridge University Press.

Maoz, Zeev, and Bruce Russett. 1993. “Normative and Structural Causes of Democratic Peace, 1946–1986.” *American Political Science Review* 87:624-38.

Marshall, Monty G., Jaggers Keith, and Robert Gurr Ted. 2014. Polity IV Aannual Time-Series, 1800-2013. *Center for International Development and Conflict Management at the University of Maryland College Park*. Accessed June 14, 2018. http://www.systemicpeace.org/polity/polity4.htm.

Palmer, Glenn, Vito D’Orazio, Michael Kenwick, and Matthew Lane. 2015. “The MID4 Dataset, 2002–2010: Procedures, Coding Rules and Description.” *Conflict Management and Peace Science* 32:222-42.

Partell, Peter J., and Glenn Palmer. 1999. "Audience Costs and Interstate Crises: An Empirical Assessment of Fearon's Model of Dispute Outcomes." *International Studies Quarterly* 43: 389-405.

Pevehouse, Jon, Timothy Nordstrom, and Kevin Wranke. 2004. "The Correlates of War 2: International Governmental Organizations Data, Version 2." *Conflict Management and Peace Science* 21:101-19.

Quackenbush, Stephen L., and Jerome F. Venteicher. 2008. "Settlements, Outcomes, and the Recurrence of Conflict." *Journal of Peace Research* 45:723-42.

Reiter, Dan, Allan C. Stam, and Michael C. Horowitz. 2014. "A Revised Look at Interstate Wars, 1816–2007." *Journal of Conflict Resolution* 60 (5): 956-76.

Russett, Bruce, and John Oneal. 2001. *Triangulating Peace: Democracy, Interdependence, and International Organization*. New York: W. W. Norton.

Sarkees, Meredith, and Frank W. Wayman. 2010. *Resort to War: Civil and International Wars, 1816–2007*. Washington, DC: CQ Press.

Singer, J. David. 1990. "Reconstructing the Correlates of War Data Set on Material Capabilities of States, 1816–1985." In *The Correlates of War*, edited by J. D. Singer and P. F. Diehl. Ann Arbor: University of Michigan Press, pp. 53-72.

Singer, David J., and Melvin Small. 1972. *The Wages of War, 1815–1965*. New York: John Wiley.

Small, Melvin, and David J. Singer. 1982. *Resort to Arms: Civil and International War*. Beverly Hills, CA: Sage.

Sullivan, Patricia L. 2007. "War Aims and War Outcomes Why Powerful States Lose Limited Wars." *Journal of Conflict Resolution* 51:496-524.

Sullivan, Patricia L. 2012. *Who Wins? Predicting Strategic Success and Failure in Armed Conflict*. New York: Oxford University Press.

Tuchman, Barbara. 1962. *The Guns of August*. New York: Random House.

Weinberg, Gerhard L. 2005. *A World at War*, 2nd ed. New York: Cambridge University Press.