Decisions, Decisions: Re-examining the Trade-Conflict Nexus

Christopher Schwarz
University of Wisconsin-Milwaukee

Nicholas R. Davis
University of Wisconsin-Milwaukee

March 17, 2016

Abstract

Despite trade being recognized as an important factor in the contemporary study of interstate conflict, no consensus has been reached on the causal direction of this effect. We enter this debate by examining the ways in which the various research designs employed by authors interact to condition their reported results. Towards this end, we review two large ongoing and unresolved debates in the literature: spatiotemporal sample selection and measurement of key concepts. We then conduct a meta-analysis of the existing literature to generate testable hypotheses, showing how the findings of authors are conditioned by the decisions they make in designing their research. We then empirically re-evaluate the effect of trade on conflict by performing a series of 48 logistic regressions, specified by iterating over various spatiotemporal domains and operationalizations. Our findings suggest that the support for the liberal perspective may be overstated.

Prepared for the Annual Meeting of the International Studies Association in Atlanta, March 16th–19th 2016. Panel TB20: The Initiation, Dynamics, and Termination of International Conflict.

Introduction and Motivation

Despite trade being recognized as an important factor in the contemporary study of interstate conflict, there is still a debate regarding whether economic interdependence pacifies, aggravates, or has no effect on interstate relations. This debate is by no means new, and such questions have been of interest to scholars for centuries. For example, Immanuel Kant envisioned free trade between nations as a key factor in the creation of a peaceful world. On the other hand, Rouseau argued that the state of war originates from the mutual dependence of naturally unequal states. Interdependence, he argued, breeds competition which inevitably causes discord and conflict. Contemporary scholarship still grapples with these opposing views, and indeed, a lack of consensus in the literature persists.

The contemporary debate has evolved, and now generally revolves around how the decisions researchers make affect their results, the conclusions they draw, and the development of the literature as a whole. Whereas recent research has considered design decisions regarding "the role of geographic proximity, the role of national size, the handling of the missing and zero trade data, and the conceptualization of what constitutes dyadic conflict" (Keshk, Reuveny, and Pollins 2010, 6), a number of other theoretical and modeling issues have not been resolved in a satisfactory manner. In particular, we identify two broad decisions as particularly important if the literature is to be internally consistent, productive, and contribute to our cumulative knowledge regarding the effect of interstate commerce on conflict: spatiotemporal scope and the appropriate operational definition of economic interdependence.

Indeed, given the heterogeneous decisions researchers make in their applied research, it seems reasonable to ask: if authors are analyzing different groups of states, different time periods, and different measures, in what way are individual pieces of scholarship in dialog with each other and the broader literature? Indeed, these decisions are not made in isolation, and while each of these decisions has been informally debated individually, the way in which combinations of these decisions affect the findings has not been systematically examined. This is an important lacuna in the extant literature. As noted by Reuveny, Pollins, and

Keshk (2010), the weight of evidence in favor of the liberal commercial peace may not be as heavy as generally thought given that they use "similar statistical models, indicators, and data." More concerning is the fact that a number of findings generally included in the literature which are frequently cited as support for the liberal perspective (primarily earlier works such as Oneal and Russett 1997, Oneal at al. 1998, Oneal and Russett 1999) make research design decisions that that are called into question by later developments in the literature. That is to say, our prior beliefs as a field regarding the quantity of evidence supporting the liberal peace have not been properly updated in light of developments in the field regarding the validity of our research designs. These developments and the debates around which they revolve are discussed in subsequent sections of this work.

In a similar spirit to Keshk, Reuveny and Pollins (2010), we conduct this study to highlight a number of theoretical and empirical questions in the literature that have not been sufficiently addressed. There are a number of plausible alternative research choices that scholars may make and have made which may influence their eventual findings, and this work is no exception. The driving objective of this research is not to dismiss previous research but to investigate the extent to which their conclusions are robust to changes in alternative schema.

In the interest of promoting transparency and enhanced dialogue in the conflict studies literature, we explore how research design choices condition results. We begin with an overview of the literature on trade and conflict, outlining the broad differences in research design. We then specify several hypotheses about the way design choices condition results and a research design to test them. We then provide results that demonstrate the conditionality of design choices that explains why the extant literature appears to routinely find conflicting findings.

The rest of the paper proceeds as follows. First, we discuss a number of debates in the literature regarding the issues of measurement, spatiotemporal sample selection, and missing data. Second, we conduct a limited meta-analysis of the literature to generate

testable hypotheses regarding the effect of research decisions on results. In the next sections we present our research design and findings, followed by a discussion of the way in which the conclusion one draws is dependent on the design that they employ. Finally, we conclude by suggesting future research paths and improvements that could be made to the design employed by this work with an eye towards more fully vetting the ways in which design drives results.

Decisions, Decisions

Operational Definitions of Interdependence

Despite the relative consensus regarding the conceptual validity of 'economic interdependence,' there has been a vibrant debate regarding how to operationalize the concept. Gartzke and Li (2003a), for example, find eight distinct measures that have been used to relate trade to conflict. This debate initially arose to engage with the contradictory findings presented by Barbieri (1996) and Oneal and Russett (1997).[1] While Barbieri had found that the "overall effect of the model produces a higher probability of conflict when interdependence rises" (1996, 40), Oneal and Russett found that "[h]igher levels of economically important trade ... are associated with lower incidences of militarized interstate disputes and war" (1997, 288). Furthermore, referring to Oneal et al. (1996), they state that "the discrepancy in Barbieri's and Oneal et al.'s results is due to differences in the measurement of interdependence" (Oneal and Russett 1997, 272).

This debate was joined by Gartzke and Li, who sought to explain the divergent findings. They state that "the difference between Barbieri's and Oneal and Russett's approaches to concept operationalization centers on the difference between the importance of a particular bilateral trade relationship to a country's overall trade and the importance of the bilateral

[1]We will refer to Oneal and Russett and the liberal position simply as "OR" throughout this analysis, particularly where space is a concern.

trade relationship to a country's total economy" (2003a, 555-6). Gartzke and Li argue that the measure used by Barbieri reflects "the degree to which a state is disconnected from world trade, [whereas] Oneal and Russett's measure reflects aspects of both trade concentration and dyadic economic openness" (556), and show that trade 'openness' and 'dependence' have a pacifying effect on interstate violence whereas trade 'share' has an aggravating effect.

A number of challenges were published in response to Gartzke and Li (2003a). For example, Barbieri and Peters (2003) argue that Gartzke and Li draw erroneous conclusions from their mathematics, and that dyadic interdependence is positively associated with conflict when truly dyadic measures are used. On the other hand Oneal (2003) agrees with Gartzke and Li, and argues that the substantive impact of dependence is also strong. In reply, Gartzke and Li (2003b) argue that the results of Barbieri and Peters stem from omitted variable bias, and show no statistically significant relationship between their measure of interdependence and conflict when distance and major power status are controlled for. More recently, Hegre argues that the Barbieri measure is not a measure of trade disconnectedness, and that "the results of Barbieri (1996) and Oneal and Russett (1997) cannot be attributed to variable construction" (2005, 223), and that the pacifying effect of trade openness is likely spurious. Gartzke and Li (2005) reply to this critique and show that the effect of trade openness is robust to controlling for the economic size of both states in the dyad.

Most recently, Xiang, Xu, and Keteku (2007) contribute to the debate by arguing that Barbieri's (1996) results are driven by omitting national power from consideration. In particular, they argue that Barbieri's trade measures act as a proxy for power leading to positive bias in the coefficients. When power is controlled for, the authors show that even Barbieri's preferred measure has a strong pacifying effect. This finding is puzzling in light of the above outlined measurement debate; it seems to indicate that there is not yet agreement on the degree to which authors' conclusions are driven by their operational definition. This puzzle is enhanced when one notices the degree of heterogeneity between authors on a number of other decisions made in their contributions to the debate.

Perhaps the most clear example of how the individual contributions to the above debate talk past each other is in regards to the time-frame studies. Constraining our attention to only those articles that contributed to the debate over measurement, six distinct time periods were examined in isolation. In particular, Barbieri's (1996) analysis analyzes data spanning 1970-1938, Oneal and Russett (1997) analyze data covering the 1950-1985 period, Xiang et al. (2007) analyze 1870-1992, Gartzke and Li (2003a, 2005) analyze 1950-1992, both Barbieri and Peters (2003) and Gartzke and Li (2003b) analyze the years 1949-1992, while Oneal (2003) looks at the 1886-1992 period. In this way, whether and in what way a researcher's operational definition affects their results is simply unknown due to authors adopting different designs with choices on more than one dimension simultaneously; the unique impact of measurement decisions is unidentified in the current literature.[2] We take up the discussion of spatiotemporal selection in the next section.

The Problem of Space, Time, and Missing Data

As noted by Barbieri and Schneider, "[c]onfusion arises over scholars' tendency to focus their attention on different spatial and temporal domains," and it is possible that "[t]he differences in findings that arise from alternative research strategies may, in fact, highlight the variations in the trade-conflict relationship that exist under alternative conditions" (1999, 394). We assert, alternatively, that the results presented by various authors are conditioned by the spatiotemporal domain that they study and that a direct analogy may be drawn between this effect and the recent vibrant debate in the field regarding missing trade data. A simple thought experiment is sufficient to illustrate this point.

Suppose a researcher is able to construct an exhaustive list of possible space and time units that comprise the unit of analysis of their intended research, such as dyad-years. Let

[2]This is conceptually identical to a conditional effect, as outlined below. Because the results are conditioned by different time periods and spatial units, the partial derivative with respect to the measure included is a function of not only the measure but also the tacit interaction between the measure and the spatiotemporal domain of the analysis. By not holding all else to be constant, the unique impact of changing the measure is not known.

the researcher then assign a dummy variable to each unit, taking the value of one if it is to be included in the analysis and zero otherwise. If this dummy variable is assigned randomly, then the estimates reported by the researcher in their subsequent analysis would have no problems, provided that they still have more observations than covariates. This then avoids the problem of micronumerosity (Gujarati 2003, ch. 10). If, however, there is a systematic assignment into the population to be analyzed, then estimates will be biased away from the true population value (Allison 2002).

In practice, the assignment of units into the analyzed sample is not random. Rather, researchers consciously choose a discrete subset of the population of cases based upon time and space conditions. This is in large part unavoidable, and is driven in part by incomplete data. For instance, broad data on regime type does not exist before the early 1800s. Similarly, information on GDP is sparse before the 1950s, and trade data has only been systematically reported throughout the world since the end of the Cold War. Indeed, early analyses which attempted to extend their data back to 1870 were able to only collect data on some 16 countries for which GDP and trade data were available, twelve of which were countries of central and western Europe (Oneal and Russett 1997). Insofar as these 'empty cells' are not missing at random, the results will be biased from the true population value. Whereas the problem of missing trade data has recently been discussed (Gleditsch 2002, Barbieri et al. 2009, Gleditsch 2010, Barbieri and Keshk 2011, Boehmer et al. 2011), the broader issue of spatiotemporal selection has largely gone unaddressed.

Along the time dimension, this lacuna is especially troubling in light of potential issues of non-stationarity when working with panel data. In most studies with a time-series component, it is assumed that the mean, variance, and autocorrelation structure of the data does not change over time; i.e. the joint probability distribution does not change when shifted in time. In the trade-conflict literature, this issue has only been addressed tangentially through the debate on if the liberal peace was an artifact of the cold war. In the words of Oneal and Russett (1999, 214), "dyadic patterns of peace and conflict during the cold war

era undoubtedly were substantially conditioned by the particular bipolar, democratic and capitalist versus autocratic and state socialist ideologies and institutions of the times." This relates to the broader issue of temporal selection in general; if there is reason to believe that that the effect of economic interdependence changes as some function of time or 'era,' then pooling may be inappropriate. More in line with our research interest, this implies that the conditional effect of temporal selection is likely non-zero. In the words of Beck and Katz (2001, 494) "the assumption of complete homogeneity of data, across both units and time, is usually suspect."

Whereas temporal selection has gone without extensive debate in the literature, a large discussion has revolved around the use of 'politically relevant dyads' or other restrictions on the spatial domain analyzed. For example, Barbieri (1996, 43) asserts that "the composition of dyads in a given sample may have a more dramatic impact on the empirical findings than variations in measurement" and that "the decision to focus exclusively on 'politically relevant dyads' may be one source of difference" between results.

What are politically relevant dyads, and why might researchers use them as opposed to the complete spatial domain?[3] Relevant dyads are defined at "pairs of states directly or indirectly contiguous and/or in which at least one of the states is a major power" (Lemke and Reed 2001, 127). In general, there are two justifications leveraged for examining only politically relevant dyads. On the one hand, such a selection criteria drastically reduced the number of observations, and thus the demands facing the researcher for collecting information across cases. In the words of Hegre (2000, 10 fn 4), "[a] dataset with 500,000 cases is much more manageable than one with 2,000,000." The second common justification is that analysis should be restricted to relevant dyads because "only in this relatively small subset of dyads is there a possibility for irreconcilable conflicts of interest to arise and create a substantial risk of war" (Weede 1976, 396). In a similar spirit, Oneal and Russett (1997, 273) justify analyzing only relevant dyads because doing so "excludes dyads that, in the great majority of

[3]For a broader theoretical and empirical discussion, see Lemke and Reed (2001), Benson (2005) and Bennett (2006).

cases, had no reasonable opportunity to engage in armed conflict." In another paper the same authors state that because politically relevant dyads account for 83 percent of militarized disputes and such a subset reduces "the number of observations from over 386,000 to 38,039, which is 10% of the total" (1999, 425); that it is a desirable restriction.

A sharp contrast can be drawn between this logic and the responses to Green et al.'s (2001) argument for the inclusion of fixed effects. While we agree that fixed effects are inappropriate given the alternative methods to model heterogeneity in event history data, the change in opinion is worth noting. For example, Oneal and Russett (2001) argue

> It is simply impossible to think that the 97,150 annual observations of the experiences of the 2,751 dyads that managed to live in peace (84 percent of our total number of cases) tells us nothing about the causes of war ... [a]fter all, epidemiologists studying the causes of cancer do not normally limit their analyses to those who have already had the disease. They look at individuals who have never had cancer. (482)

Likewise, Oneal and Russett were joined by Beck and Katz (2001), who levy a similar critique:

> ... most dyads never conflict; in fact, over 93 percent of Green, Kim, and Yoon's dyads — 2,887 out of 3,078 — never do... this means that over 90 percent of the dyads have no impact of the statistical estimates. Thus, a data set that contained only the 7 percent of the dyads that conflict would yield identical estimates to the full dataset (including dyads that never conflict). (488)[4]

The authors then go on to assert that "[i]t is absurd to exclude over 90 percent of the cases from the analysis ... and then conclude that some independent variable like democracy has the opposite effect of what every sensible study has shown" (Beck and Katz 2001, 490). This is puzzling in light of the fact that these 'sensible' studies imposed restrictions on their sample with a very similar spirit — to reduce the cases of the analysis to those that have an *a priori* higher likelihood of conflict. While not exactly the same, the use of fixed effects and

[4]In the dataset we construct below, roughly 87% of onsets occur within politically relevant dyads which themselves comprise approximately 14% of the total dyad-years.

politically relevant dyads have extremely similar effects on the sample that is then analyzed. In both cases, the statistical estimates are conditional on some function of being more likely to fight.

To what extent do these research design decisions influence the results? The next section reviews the way in which design choices in the literature are associated with the direction and existence of the trade-conflict relationship. It also serves as the foundation for our prior beliefs regarding how choices may influence results, from which we generate testable hypotheses.

Where the Literature Stands

A variety of views, both historical and contemporary, have been expressed regarding whether trade would have a pacifying, aggravating, or null effect on interstate conflict. The works of authors such as Kant, Lenin, Morgenthau, Waltz, Montisquieu, Rouseau, and Mill are among the best known within the contemporary scholarship, and are generally thought of as foundational within their various theoretical paradigms. Notably, the views of Kant (and to a lesser degree Montisquieu and Mill) are at the forefront as they provide the foundation for classical liberalism and posit a pacifying effect of trade on commerce. Prominent examples of this view include those authored by Bruce Russett and John Oneal and their co-authors (e.g. 1997, 1998, 1999, 2001, 2010). Highlighted most comprehensively in the works of Barbieri (1996, 2002), the neo-Marxist perspective draws upon the works of Lenin and argues that economic interdependence breeds conflict, much in the same way as Rouseau, and Waltz's neo-realism. Finally, classical realists such as Morgenthau argue that trade is irrelevant to conflict decisions. Recent work such as Gelpi and Greico (2008) come to this conclusion as well.

What does the empirical literature find regarding the effect of trade on conflict? To answer this question, we conduct a systematic overview of the literature. The articles for

this review were identified based upon a number of criteria using Google Scholar. The first criteria used to filter through the vast literature was whether they cited Polachek (1980), Oneal and Russett (1997) or Barbieri (1996). The work by Polachek was selected due to it being widely cited (689 as of this writing) foundational in establishing the 'opportunity cost' logic engrained in subsequent liberal arguments. The works of were selected for being both widely cited (1057 citations and 621 respectively), as well as being central to the operationalization debate outlined in the previous section. In the second step we filtered the citation results by journal. In particular, we retained only those articles that were published in seven high-profile journals, primarily those known within the international relations subfield. The journals used for this filter were (in no particular order) Journal of Conflict Resolution, Journal of Peace Research, International Studies Quarterly, American Journal of Political Science, International Organization, American Political Science Review, and World Politics. After applying this filter, we were left with 146 unique articles to examine.

Third, we decided to consider only those articles that use militarized interstate disputes (MIDs) as their dependent variable and some sort of trade measure as an independent variable to ensure comparability between articles. This left us with 48 unique articles as our sample of the literature published from 1996 to 2015. From these articles, we collected a number of pieces of information. First, we identified the regression(s) which contained the author's 'main results.' If this was not clearly identified by the author, we judged the extent to which their findings were consistent across the models presented. This process yielded 51 regression results, and comprise the main unit of analysis for this section. Second, with these reported results in hand, we identified a number of decisions made by the authors. In particular we identified (1) the operational definition used to represent economic interdependence; (2) the time period over which the analysis was conducted; (3) whether the sample was constrained, such as by including only politically relevant dyads; (4) the number of observations; (5) the number of unique dyads, if reported; and finally (6) whether a pacifying,

aggravating, or null effect was found at the 95% confidence level or higher.

It is important to note that we did not limit our analysis to those articles whose driving goal was to contribute to the debate regarding the effect of trade on conflict. A number of the articles included in our analysis utilize trade as a control variable, yet we view their findings as being just as important as those articles which explicitly attempt to make a contribution to that debate. This comes from our opinion that a result is a result regardless of whether you were looking for it or not; that is to say, our beliefs regarding the effect of trade on conflict should be informed by all analyses in which it is involved, not only those which seek to find a certain relationship.

Where does the literature stand on the effect of trade on conflict? While the majority of findings support the pacifistic effect of trade on conflict, this makes up only 51% of the studies examined, implying that the support for the liberal proposition may have been overstated by proponents. While the support for a null effect (39%) and aggravating effect (10%) are notably smaller individually, a larger 'not pacifying' category would contain 49% of the studies examined. This leads us to our first and broadest hypothesis:

> **Hypothesis 1** *The probability of a pacifying result is greater than null or aggravating; not controlling for measure of spatiotemporal domain.*

To identify the way in which design decisions appear to influence reported results, we may begin to break down the observed effect along each 'decision dimension' one by one. We begin our analysis by examining the temporal dimension. Within the literature reviewed, there are 26 distinct temporal ranges analyzed in the strict sense, 18 of which are unique in that no other study uses that temporal selection. For simplification and to address the issue of how temporal context may condition the results, we find two periods to be both theoretically relevant and extant in the literature. First, as noted above, the only treatment of temporal selection in the literature that we found was questioning whether the liberal result was a function of the cold war period. While a direct way of addressing this critique would have been to examine only the pre or post cold war period in their analyses, authors

Table 1: Decisions and Results in the Literature

	All Dyads, 57%		Relevant Dyads, 43%	
	OR Measure (72%)	¬ OR Measure (28%)	OR Measure (82%)	¬ OR Measure (18%)
Effective Dyads	3.3, 4688.207, 24974.607 (Min, Mean, Max)	210.897, 3744.608, 10312.875 (Min, Mean, Max)	526.765, 3278.897, 12292.373 (Min, Mean, Max)	0.818, 329.589, 691 (Min, Mean, Max)
Pacifying	Oneal and Russett (2001), 1886-1992 Jungblut and Stoll (2002), 1950-1978 Gartzke and Li (2003*b*), 1950-1992 Oneal (2003), 1886-1992 Oneal, Russett and Berbaum (2003), 1885-1992 Hegre (2004), 1951-1992 Hafner-Burton and Montgomery (2006), 1885-1992 Maoz et al. (2006), 1816-2000 Gelpi and Grieco (2008), 1950-1992 Rauchhaus (2009), 1885-2000 Goenner (2010), 1962-2000 Nordhaus, Oneal and Russett (2012), 1950-2000	Gartzke and Li (2003*b*), 1950-1992 Oneal (2003), 1886-1992 Xiang, Xu and Keteku (2007), 1870-1992	Oneal et al. (1996), 1951-1985 Oneal and Russet (1997), 1950-1985 Beck and Jackman (1998), 1951-1985 Hegre (2000), 1950-1992 Mousseau (2000), 1950-1992 Gartzke, Li and Boehmer (2001), 1951-1985 Krustev (2006), 1950-1992 Maoz (2009), 1870-2001 Adam and Sekeris (2015), 1975-2001	Aydin (2010), 1816-1999 Murshed and Mamoon (2010), 1950-2005
No effect	Oneal and Russett (1999), 1950-1992 Green, Kim and Yoon (2001), 1951-1992 Gartzke and Li (2003*a*), 1950-1992 Goenner (2004), 1950-1992 Gartzke (2007), 1950-1992 Dorussen and Ward (2010), 1948-2000 Peterson and Drury (2011), 1914-2000	Peterson (2013), 1984-2000 Chatagnier and Kavaklı (2015), 1962-2000	Beck, Katz and Tucker (1998), 1951-1985 Gartzke (1998), 1951-1985 Russett, Oneal and Davis (1998), 1950-1985 Gartzke (2000), 1951-1985 Reed (2000), 1950-1985 McDonald (2004), 1960-1992 Lupu and Traag (2012), 1960-2000 Kim (2013), 1950-2001 Peterson (2015), 1961-2000	Mansfield and Pevehouse (2000), 1950-1985 Gartzke and Li (2003*c*), 1950-1992
Aggravating	Davies (2002), 1950-1982 Peterson (2011), 1885-2000	Barbieri (1996), 1870-1938 Barbieri and Peters (2003), 1949-1992 Gartzke and Li (2003*b*), 1950-1992		

generally just extended their analysis to include more years in their sample. Indeed, the only study that strictly examined the pre-cold war period was Barbieri (1996). All other studies examined either (1) begin after WWII and include the cold war period or (2) begin before WWII and include the cold war period. Thus, by only excluding Barbieri (1996) we can think of these two broad temporal ranges.

Such a division leaves 50 total results, 12 of which begin before WWII and 38 of which begin after WWII. Of those that began before WWII, 83% of studies found a pacifying effect while the remaining 17% is split evenly between aggravating and null effects. Interestingly, for those studies that examine the post-WWII era, 50% find a null effect, 42% find a pacifying effect, and the remaining 8% find an aggravating effect. While inconclusive on it's own, this implies that the worries of spuriousness were overstated given that studies that examine only the post-WWII era are less likely to find a pacifying effect. This leads to the following general expectation:

> **Hypothesis 2** *The probability of finding a pacifying effect is greater when the pre-WWII period is included, relative to only examining the post-WWII period; not controlling for temporal or measurement decisions.*

Next, we examine the spatial dimension. Keeping only those 50 results examined above, There is a large decree of heterogeneity among the spatial domain employed by authors. In particular, 56% of the studies examined use all dyads in their analysis whereas 44% use some sort of spatial restriction, typically some sort of political relevance criteria. When examining politically relevant dyads, there results are evenly split between finding a null effect and a pacifying effect with 50% finding each outcome. For those that examine all dyads, roughly 54% find a pacifying effect, 32% find a null effect, and 14% find an aggravating effect. As such, we expect to find that:

> **Hypothesis 3** *The probability of finding a pacifying effect is greater when all dyads are examined as opposed to politically relevant dyads; not controlling for spatial or measurement decisions.*

Third, we examine the effect of measurement. For convenience, we combine all measures that are not the bilateral-trade-weighted-by-GDP measure popularized by Oneal and Russett. Even combining these measures into one category, they represent only 20% of the sample gathered. These seems to support the assertion made by Reuveny, Pollins, and Keshk (2011) that the majority of evidence for the liberal perspective is "based on RO or RO-like papers that employ almost similar statistical models, indicators, and data." Of the studies that did not use the Oneal and Russett measure, 20% report an aggravating effect while the remaining 80% are evenly split between reporting null and pacifying effects. Those that use the Oneal and Russett measure, however, 55% report a pacifying effect, 40% report no effect, and the remaining 5% report an aggravating effect.

> **Hypothesis 4** *The probability of finding a pacifying effect is greater when the Oneal and Russett measure is used relative to other measures; not controlling for spatiotemporal decisions.*

Having described the effect of these individual decisions, we can now start to think about how, quite literally, these decisions interact to shape results. For simplicity, we begin with pairwise decisions, first examining the spatiotemporal dimensions. Among those results that examined politically relevant dyads in the post-WWII era, 55% find no effect and 45% find a pacifying effect. For those that instead also examine the pre-WWII era, 100% of the studies find a pacifying effect. For those that examine all dyads in the post-WWII era, 38% find a pacifying effect, 44% find a null effect, and 18% find an aggravating effect. When the pre-WWII era is included, however, 80% find a pacifying effect whereas the remaining 20% is evenly split between null and aggravating effects. To summarize our expectations derived from the spatiotemporal dimensions:

> **Hypothesis 5** *Among politically relevant dyads, the probability of finding a pacifying effect is greater when the pre-WWII period is included relative to the post-WWII period; not controlling for measurement.*

> **Hypothesis 6** *Among all dyads, the probability of finding a pacifying effect is*

greater when the pre-WWII period is included relative to the post-WWII period; not controlling for measurement.

Next, we examine how space and measurement decisions interact. Interestingly, among politically relevant dyads there is no difference between effects based on measurement, regardless of the measure employed, 50% of the results fall into each of the null and pacifying effect categories. When all dyads are examined, however, results are evenly split between each outcome when the Oneal and Russet measure is not used. When the Oneal and Russett measure is used, however, 59% of the findings report a pacifying effect, 32% report no effect, and 9% report an aggravating effect. As such, we expect the following:

Hypothesis 7 *Among politically relevant dyads, there is no difference in the probability of finding a pacifying or null effect based upon the measure employed; not controlling for temporal decisions.*

Hypothesis 8 *Among all dyads, there is a greater probability of finding a pacifying result when the Oneal and Russet measure is employed relative to other measures; not controlling for temporal decisions.*

Turning to the the final pairwise combination, we analyze how temporal and measurement decisions condition the reported results. First, those studies that include in their analysis the pre-WWII era unanimously find a pacifying effect. In the same time period, those that use the Oneal and Russett measure report a pacifying effect 78% of the time while the remaining 22% are evenly split between finding an aggravating or no effect. Those studies that examine only the post-WWII era and utilize the Oneal and Russett measure, however, report a null or pacifying effect at an equal rate (48%) and find an an aggravating effect 3% of the time. When the Oneal and Russet measure is not utilized, however, researchers have found a pacifying effect 14% of the time, no effect 57% of the time, and an aggravating effect 29% of the time. We thus expect:

Hypothesis 9 *When the pre-WWII era is included in the analysis, there is a greater probability of finding a pacifying effect when the Oneal and Russett measure is not used; not controlling for spatial decisions.*

Hypothesis 10 *When only the post-WWII era is analyzed, there is a greater probability of finding a pacifying result when the Oneal and Russet measure is employed relative to other measures; not controlling for spatial decisions.*

Having examined the effects of each pairwise and individual decision, we now turn to examining each triplet of decisions on the effect of an individual spatial, temporal, and measurement decision conditional on each other. Studies that examined politically relevant dyads and include the pre-WWII era unanimously find a pacifying effect regardless of the measure employed. When limited to the post-WWII era, however, a null effect is more common, 53% when the Oneal and Russett measure is used and 67% of the time when it is not.

Switching to studies that analyze all dyads, when the pre-WWII period is included, a pacifying effect is most likely regardless of measure: 75% find a pacifying effect when the Oneal and Russett measure is employed (the remaining 25% is split evenly between pacifying and aggravating) whereas 100% find a pacifying effect when the Oneal and Russett measure is not used. Finally, when scholars examine only the post-WWII period, a pacifying effect is only expected when the Oneal and Russet measure is employed (50% pacifying, 43% null, 7% aggravating) since when their measure is not employed, the results are evenly split between aggravating and null effects. These expectations are summarized in the following hypotheses:

Hypothesis 11 *Among politically relevant dyads, when the pre-WWII period is included, a pacifying effect is expected regardless of measure.*

Hypothesis 12 *Among politically relevant dyads, when the analysis is limited to the post-WWII period, a null effect is expected regardless of measure.*

Hypothesis 13 *Among all dyads, when the analysis includes the pre-WWII period, a pacifying effect is expected regardless of measure.*

Hypothesis 14 *Among all dyads, when the analysis examines only the post-WWII period, a pacifying effect is more likely to be found when the Oneal and Russett measure is used.*

We now turn to empirically vetting the above expectations by iterating along research decisions, holding else constant.

The Choices We Made

Replicability of research design and flexibility to test the many design choices outlined in the review above is paramount. In developing our own research design, we attempted to incorporate as many elements of the extant literature as possible. In order to engage the extant research on its own terms, we include multiple trade measures, the widest temporal domain possible, and both all dyads and politically relevant dyads in our analysis. Therefore, to test the above hypotheses, we created a nondirected dyad dataset built on basepairs generated by EUGene 3.204 covering the time period 1870–2007.[5] We then combined data on conflict, trade, GDP, and several controls to best replicate the data utilized in the literature reviewed above. The variables employed in our analysis are outlined in the below.

Sample Restrictions

We examine the spatial dimension by including all dyads as well as a subset of only politically relevant dyads in the analysis below. Consistent with the literature, we identify politically relevant dyads as those observations in which the members were contiguous states or one member of the dyad was a major power, using the Direct Contiguity data (Stinnett et al. 2002, Gochman 1991) and the list of major powers from Correlates of War. This distinction necessarily constricts the spatial domain covered by the analysis to those dyads most likely to experience conflict. We also specify six temporal domains based on the literature. Besides the "complete" data 1870–2007, we isolate the 1870–1938 time period as a replication of Barbieri (1996) and the 1950–1992 temporal domain to replicate the bulk of the OR and liberal outcome studies. Additionally, we examine the periods 1950–2000 and 1950–2007 to

[5]Due to the data limitations of the various sources we utilize we could not cover the entire 1816–2010 range as originally intended.

replicate the temporal domains of those studies which benefited from the MID data updates. Finally, although no reviewed studies employ this temporal domain, we include the post Cold War period 1992–2007 as a robustness check.

Measurement

Summary statistics calculated on all dyad observations and politically dyad relevant observations 1870–2007 appear in Table 3, located in the appendix. The dataset includes 404,390 dyad-year observations and 62,902 politically relevant dyad-year observations after missing data was listwise deleted. The dependent variable, `onset`, records all interstate conflicts from the Correlates of War Militarized Interstate Dispute dataset (Palmer et al. 2015). We used the MID 4.01B dataset, which contains one record for each participant in a MID from 1816–2010. We recast the participant-level data as dyadic to merge onto the base dyad set from EUGene.[6] This variable ranges from zero to one, and is quite rare; 0.2 percent of all dyads experience MIDs while 0.9 percent of politically relevant dyads experience MIDs.

The literature operationalizes trade in several different ways. We calculate four separate measures used most often to evaluate how this design choice affects outcomes. We begin with the Barbieri and Keshk bilateral trade data, which covers the time period 1870–2007.[7] Our first measure, `flow` is simply the sum of trade in each dyad-year observation. We also specify `share`, using the method employed by Barbieri and colleagues. For each observation `share` is the sum of dyad_{ij} trade divided by total trade_i bounded between zero and one. The third trade measure, `openness`, is trade_i divided by GDP_i. Our final trade measure is `dependence`, which is the Oneal and Russet measure: trade flow_{ij} divided by GDP_i.[8] In the

[6] We understand that there may have been slight modifications to the MID data from previous versions (i.e. 3.10) employed by the literature we are engaging.

[7] This choice should not be seen as prejudiced against the liberal position. Typically alternatives do not cover as wide of a temporal domain (Gleditsch 2002). Since the Barbieri data are simply trade flows from state one to state two within a dyad, it is unlikely that the data is biased towards the neoMarxist position due to calculation or transformation. However, the way that Gleditsch deals with missing data suggests that we include a trade measure robustness check in future iterations of this paper.

[8] Values of the openness and dependence measures can exceed one due to a mismatch of units between Maddison GDP and trade.

entire sample all four trade measures have small means and standard deviations relative to the maximal values, suggesting that trade has low variation but there are some outliers at the upper bound. This is also true in the politically relevant sample.

Following the literature, we also specify several control variables. First, `joint democracy` is a dummy variable equal to one when both states' POLITY scores are at minimum +7. Among all dyads, about 15 percent of observations experience joint democracy, compared to about 22 percent in the politically relevant sample as a result of several major powers being democracies. We also include a measure of `power`, taking the absolute value of the difference of the Composite Index of National Capability index scores in the dyad (CINC version 4.0). The mean of this measure is higher in the relevant dyad sample than among all dyads. We also include `contiguity`, which is a dummy variable for contiguous dyads using the maximally inclusive measure from the Direct Contiguity data, and `distance` between capitals. Predictably, contiguity is much more frequent in the politically relevant set. Finally, we include a `peace years` measure, which is the sum of dyad-years since the last MID ended. Due to uncertainty about MIDs prior to 1816, we assume that peace years begin in that year, accounting for the high maximum value for this measure. For years experiencing a MID, this value is zero. There is a statistically significant difference between the mean value of this measure among all dyads and among politically relevant dyads.

Method

Following the majority of literature reviewed above, we employ logistic regression to test our hypotheses. Since we are dealing with hypotheses about the impact of one or more design choices on the effect of trade on conflict, it is necessary to estimate models in an iterative process. Since we have two spatial dimensions (all dyads, politically relevant dyads), four measures of trade (share, flow, openness, dependence) and six temporal domains including periods before and after World War II, we specify 48 logistic regression models. Each regression model contains one measure of trade and the control variables outlined above. The

data used in each model is a subset of the overall listwise-deleted data frame using spatial and temporal restrictions.

Results

The results of our 48 logistic regression iterations appear in the appendix. Due to space concerns and readability, we have outlined our findings in Table 2 below. We enter the logistic regression model results in the table as their temporal domain restriction and trade measure, located under the column that corresponds to spatial dimension choice and row that corresponds to the outcome of interest. Overall, the posterior distribution of results is rather shocking given that our prior expectations almost uniformly supported the OR (liberal) position. In twelve of our fourteen hypotheses we expected a confirmation of the liberal position that trade reduces conflict. All but one (H12) of our fourteen hypotheses fail to find supporting evidence. In all models, we resisted the practice in the literature to use robust standard errors, which should bias our results towards the liberals,[9] yet still found no support for the pacifying effect of trade on conflict. We have demarcated the results which had the hypothesized sign but failed to gain statistical significance at $p < 0.05$.

Our findings can be discussed in terms of our three types of hypotheses. First, the general hypothesis H1 and the the unidimensional design choice hypotheses (H2, H3, and H4) were the *a priori* most likely to find support. Since we have no findings that support the liberal position in any results, H1 is not supported. Neither is H2, which examines the pre-World War II temporal constraint, nor H3 that represents the broad spatial domain. H4 examines the role of the liberal trade measure, `dependence`, also with a null finding.

The second group of hypotheses involve bi-dimensional choices. These introduce spatial differences while holding temporal domain constant (H5 and H6) and then explore spatial differences and trade measure differences without specifying temporal domain. As stated

[9]Including robust standard errors would bias the results towards the classical realist finding (null effect) as passing the threshold for statistical significance is diminished.

Table 2: Summary of Findings

	All Dyads		Relevant Dyads	
	OR Measure	¬ OR Measure	OR Measure	¬ OR Measure
Pacifying	–	–	–	–
No effect	1870–2007 dependence	flow, openness	1870–2007 dependence	flow, openness
	1870–1938 dependence†	share, flow, open.	1870–1938 dependence†	share, flow, open.
	1950–1992 dependence†	share, flow, open.	1950–1992 dependence††	share, flow, open.
	1950–2000 dependence†	share, flow, open.	1950–2000 dependence†	share, flow, open.
	1950–2007 dependence	share, flow, open.	1950–2007 dependence†	share, flow, open.
	1992–2007 dependence	share, openness	1992–2007 dependence	share, openness
Aggravating	–	1870–2007 share	–	1870–2007 share
	–	1992–2007 flow	–	1992–2007 flow

Note: Results from 48 logistic regression models, each with the same controls (see above). Entries represent finding supporting that outcome, given the temporal domain and measure listed.

† Signifies a pacifying sign (negative) despite a lack of statistical significance.

†† Signifies a pacifying sign (negative) statistically significant at $p < 0.10$.

above, none are supported. Perhaps the most interesting finding here is that for H7 there was a prior expectation of an equal probability of pacifying and null outcomes. While this is not born out by the results (again, no liberal outcomes supported) the posterior probability of the classical realist position is quite a bit higher than the prior. H9 and H10 explore choices in temporal constraint and measure choice while allowing the spatial dimension to vary. The results for the test of H10, located in Table 6 (model 8) come the closest to a finding supporting the liberal position, but at a lower level of statistical significance.

The third and final group of hypotheses explores how tri-dimensional choices affect outcomes. Again, the prior expectation of support for the liberal position is not realized. These hypotheses (H11, H12, H13, and H14) all select on the spatial, temporal, and measurement dimensions. H12 is the only hypothesis that is supported, as our prior expectation of null findings across each of the trade measures in the post-World War II politically relevant dyad sample is born out in the model results, located in Table 6 models 5 through 8. Notably, the replication of the OR design (H14) has no support. Interestingly, the neoMarxist position finds limited support (not hypothesized) in samples without temporal or spatial constraints for two of the non-OR trade measures, `share` and `flow`.

Discussion

Despite trade being recognized as an important factor in the contemporary study of interstate conflict, there is still a debate over whether economic interdependence pacifies, aggravates, or has no effect on interstate relations. Contemporary scholarship has suffered from a persistent lack of consensus in the literature, despite the debate evolving to revolve around how the decisions researchers make affect their results. Even as the literature acknowledges that design choices matter, researchers continue to make heterogeneous choices without clearly stated expectations as to how these choices have constrained their findings. We sought to empirically asses the way authors analyze different groups of states, different time periods,

and different measures as they seek to contribute to the overall dialogue on trade and conflict. Rather than approach this literature with a normative bias, we reviewed the major contributions to the literature over the past two decades and began our analysis with a prior expectation that the liberal position would be the most likely outcome supported despite manipulation of design choices. The results presented above reveal that the posterior distribution of results almost entirely lies within the classical realist position (null effect).

As noted by Reuveny, Pollins, and Keshk (2010), the weight of evidence in favor of the liberal commercial peace uses similar statistical models, indicators, and data. As stated above, a number of findings in the literature which are frequently cited as support for the liberal perspective (primarily early OR studies) make research design decisions that that appear not to be advisable. Our prior beliefs have now been updated regarding the validity of their research designs, and their findings surrounding the liberal capitalist peace are overstated. The role of spatial and temporal choices particularly condition results. We hope to eventually identify the temporal or spatial conditions under which the liberal position is supported, but in light of our systematic analysis, including good-faith efforts to replicate typical OR design choices, we can say with a degree of confidence those conditions have not been obtained.

We acknowledge that missingness (particularly, in the trade data) might be playing a role in the inability to successfully replicate OR's results with their own design. Our method of listwise deletion was a strategic choice to eliminate a fourth dimension of variation in design choices. For an initial foray into the question of how design choices affect outcomes we felt that it was necessary to limit the dimensions covered. It is unlikely that employing listwise deletion on only the 1950–1992 temporal slice would make changes to the composition of the sample, but in conjunction with alternative trade data (i.e. Gleditsch) and GDP data, the sample could be significantly altered. This analysis is not intended to be the final word on either research design choices nor on the debate over the effect of trade on conflict. Rather, we intend to employ several robustness checks to explore additional design choices. Besides

examining alternative trade flow data and GDP, in future iterations of this paper we will look at how control variable construction matches the design choices of the literature.

References

Adam, Antonis and Petros G. Sekeris. 2015. "Self-Containment: Achieving Peace in Anarchic Settings." *Journal of Conflict Resolution* pp. 1–31.
URL: *http://jcr.sagepub.com/content/early/2015/01/13/0022002714564428.abstract*

Allison, Paul D. 2002. "Missing data: Quantitative applications in the social sciences." *British Journal of Mathematical and Statistical Psychology* 55(1):193–196.

Aydin, Aysegul. 2010. "The deterrent effects of economic integration." *Journal of Peace Research* .

Barbieri, Katherine. 1996. "Economic interdependence: A path to peace or a source of interstate conflict?" *Journal of Peace Research* 33(1):29–49.

Barbieri, Katherine. 2002. *The Liberal Illusion: Does Trade Promote Peace?* University of Michigan Press.

Barbieri, Katherine and Gerald Schneider. 1999. "Globalization and peace: Assessing new directions in the study of trade and conflict." *Journal of Peace Research* 36(4):387–404.

Barbieri, Katherine, Omar Keshk and Brian Pollins. 2008. "Correlates of war project trade data set codebook." *Codebook Version* 2.

Barbieri, Katherine and Omar MG Keshk. 2011. "Too Many Assumptions, Not Enough Data." *Conflict Management and Peace Science* 28(2):168.

Barbieri, Katherine, Omar MG Keshk and Brian M Pollins. 2009. "Trading data evaluating our assumptions and coding rules." *Conflict Management and Peace Science* 26(5):471–491.

Barbieri, Katherine and Richard Alan Peters. 2003. "Measure for Mis-measure: A Response to Gartzke & Li*." *Journal of Peace Research* 40(6):713–719.

Beck, Nathaniel and Jonathan N Katz. 2001. "Throwing out the baby with the bath water: A comment on Green, Kim, and Yoon." *International Organization* 55(02):487–495.

Beck, Nathaniel, Jonathan N Katz and Richard Tucker. 1998. "Taking time seriously: Time-series-cross-section analysis with a binary dependent variable." *American Journal of Political Science* 42(4):1260–1288.

Beck, Nathaniel and Simon Jackman. 1998. "Beyond linearity by default: Generalized additive models." *American Journal of Political Science* pp. 596–627.

Bennett, D Scott. 2006. "Exploring Operationalizations of Political Relevance." *Conflict Management and Peace Science* 23(3):245–261.

Benson, Michelle A. 2005. "The relevance of politically relevant dyads in the study of interdependence and dyadic disputes." *Conflict Management and Peace Science* 22(2):113–133.

Boehmer, Charles R, Bernadette ME Jungblut and Richard J Stoll. 2011. "Tradeoffs in Trade Data Do Our Assumptions Affect Our Results?" *Conflict management and peace science* 28(2):145–167.

Chatagnier, J Tyson and Kerim Can Kavaklı. 2015. "From Economic Competition to Military Combat Export Similarity and International Conflict." *Journal of Conflict Resolution* p. 0022002715613565.

Davies, Graeme AM. 2002. "Domestic Strife and the Initiation of International Conflicts A Directed Dyad Analysis, 1950-1982." *Journal of Conflict Resolution* 46(5):672–692.

Dorussen, Han and Hugh Ward. 2010. "Trade networks and the Kantian peace." *Journal of Peace Research* 47(1):29–42.

Gartzke, Erik. 1998. "Kant we all just get along? Opportunity, willingness, and the origins of the democratic peace." *American Journal of Political Science* pp. 1–27.

Gartzke, Erik. 2000. "Preferences and the democratic peace." *International Studies Quarterly* 44(2):191–212.

Gartzke, Erik. 2007. "The capitalist peace." *American journal of political science* 51(1):166–191.

Gartzke, Erik and Quan Li. 2003*a*. "Alls well that ends well: A reply to Oneal, Barbieri & Peters." *Journal of Peace Research* 40(6):727–732.

Gartzke, Erik and Quan Li. 2003*b*. "Measure for measure: Concept operationalization and the trade interdependence-conflict debate." *Journal of Peace Research* 40(5):553–571.

Gartzke, Erik and Quan Li. 2003*c*. "War, peace, and the invisible hand: Positive political externalities of economic globalization." *International Studies Quarterly* 47(4):561–586.

Gartzke, Erik and Quan Li. 2005. "Mistaken identity: A reply to Hegre." *Conflict Management and Peace Science* 22(3):225–233.

Gartzke, Erik, Quan Li and Charles Boehmer. 2001. "Investing in the peace: Economic interdependence and international conflict." *International organization* 55(02):391–438.

Gelpi, Christopher F and Joseph M Grieco. 2008. "Democracy, interdependence, and the sources of the liberal peace." *Journal of Peace Research* 45(1):17–36.

Gleditsch, Kristian Skrede. 2002. "Expanded trade and GDP data." *Journal of Conflict Resolution* 46(5):712–724.

Gleditsch, Kristian Skrede. 2010. "On Ignoring Missing Data and the Robustness of Trade and Conflict Results A Reply to Barbieri, Keshk, and Pollins." *Conflict Management and Peace Science* 27(2):153–157.

Gochman, Charles S. 1991. "Interstate metrics: Conceptualizing, operationalizing, and measuring the geographic proximity of states since the Congress of Vienna." *International Interactions* 17(1):93–112.

Goenner, Cullen F. 2004. "Uncertainty of the liberal peace." *Journal of Peace Research* 41(5):589–605.

Goenner, Cullen F. 2010. "From toys to warships: Interdependence and the effects of disaggregated trade on militarized disputes." *Journal of Peace Research* 47(5):547–559.

Green, Donald P, Soo Yeon Kim and David H Yoon. 2001. "Dirty pool." *International Organization* 55(02):441–468.

Gujarati, Damodar N. 2003. *Basic Econometrics*. 4th ed. New York: McGraw-Hill.

Hafner-Burton, Emilie M and Alexander H Montgomery. 2006. "Power positions international organizations, social networks, and conflict." *Journal of Conflict Resolution* 50(1):3–27.

Hegre, Håvard. 2000. "Development and the liberal peace: What does it take to be a trading state?" *Journal of Peace Research* 37(1):5–30.

Hegre, Håvard. 2004. "Size asymmetry, trade, and militarized conflict." *Journal of Conflict Resolution* 48(3):403–429.

Hegre, Håvard, John R Oneal and Bruce Russett. 2010. "Trade does promote peace: New simultaneous estimates of the reciprocal effects of trade and conflict." *Journal of Peace Research* 47(6):763–774.

Jungblut, Bernadette ME and Richard J Stoll. 2002. "The Liberal Peace and Conflictive Interactions: The Onset of Militarized Interstate Disputes, 1950-78." *Journal of Peace Research* 39(5):527–546.

Keshk, Omar MG, Rafael Reuveny and Brian M Pollins. 2010. "Trade and conflict: Proximity, country size, and measures." *Conflict Management and Peace Science* 27(1):3–27.

Kim, Nam Kyu. 2013. "Testing Two Explanations of the Liberal Peace The Opportunity Cost and Signaling Arguments." *Journal of Conflict Resolution* p. 0022002713484280.

Krustev, Valentin L. 2006. "Interdependence and the duration of militarized conflict." *Journal of Peace Research* 43(3):243–260.

Lemke, Douglas and William Reed. 2001. "The relevance of politically relevant dyads." *Journal of Conflict Resolution* 45(1):126–144.

Lupu, Yonatan and Vincent A Traag. 2012. "Trading communities, the networked structure of international relations, and the Kantian peace." *Journal of Conflict Resolution* p. 0022002712453708.

Mansfield, Edward D and Jon C Pevehouse. 2000. "Trade blocs, trade flows, and international conflict." *International organization* 54(04):775–808.

Maoz, Zeev. 2009. "The effects of strategic and economic interdependence on international conflict across levels of analysis." *American Journal of Political Science* 53(1):223–240.

Maoz, Zeev, Ranan D Kuperman, Lesley Terris and Ilan Talmud. 2006. "Structural equivalence and international conflict a social networks analysis." *Journal of Conflict Resolution* 50(5):664–689.

Marshall, Monty G., Ted Robert Gurr and Keith Jaggers. 2010. *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010*. Polity IV Project.

McDonald, Patrick J. 2004. "Peace through trade or free trade?" *Journal of conflict resolution* 48(4):547–572.

Mousseau, Michael. 2000. "Market prosperity, democratic consolidation, and democratic peace." *Journal of Conflict Resolution* 44(4):472–507.

Murshed, Syed Mansoob and Dawood Mamoon. 2010. "Not loving thy neighbour as thyself: Trade, democracy and military expenditure explanations underlying IndiaPakistan rivalry." *Journal of Peace Research* 47(4):463–476.

Nordhaus, William, John R Oneal and Bruce Russett. 2012. "The effects of the international security environment on national military expenditures: A multicountry study." *International Organization* 66(03):491–513.

Oneal, John R. 2003. "Measuring interdependence and its pacific benefits: A reply to Gartzke & Li." *Journal of Peace Research* 40(6):721–725.

Oneal, John R and Bruce M Russet. 1997. "The classical liberals were right: Democracy, interdependence, and conflict, 1950–1985." *International Studies Quarterly* 41(2):267–294.

Oneal, John R and Bruce Russett. 1999. "Assessing the liberal peace with alternative specifications: Trade still reduces conflict." *Journal of Peace Research* 36(4):423–442.

Oneal, John R and Bruce Russett. 2001. "Clear and clean: The fixed effects of the liberal peace." *International Organization* 55(02):469–485.

Oneal, John R, Bruce Russett and Michael L Berbaum. 2003. "Causes of peace: Democracy, interdependence, and international organizations, 1885–1992." *International Studies Quarterly* 47(3):371–393.

Oneal, John R, Frances H Oneal, Zeev Maoz and Bruce Russett. 1996. "The liberal peace: Interdependence, democracy, and international conflict, 1950-85." *Journal of Peace Research* 33(1):11–28.

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(2):222–242.

Peterson, Timothy M. 2011. "Third-party trade, political similarity, and dyadic conflict." *Journal of Peace Research* 48(2):185–200.

Peterson, Timothy M. 2013. "Dyadic trade, exit costs, and conflict." *Journal of Conflict Resolution* p. 0022002713478794.

Peterson, Timothy M. 2015. "Insiders versus Outsiders Preferential Trade Agreements, Trade Distortions, and Militarized Conflict." *Journal of Conflict Resolution* 59(4):698–727.

Peterson, Timothy M and A Cooper Drury. 2011. "Sanctioning Violence The Effect of Third-Party Economic Coercion on Militarized Conflict." *Journal of Conflict Resolution* 55(4):580–605.

Polachek, Solomon William. 1980. "Conflict and trade." *Journal of Conflict resolution* 24(1):55–78.

Rauchhaus, Robert. 2009. "Evaluating the Nuclear Peace Hypothesis A Quantitative Approach." *Journal of Conflict Resolution* 53(2):258–277.

Reed, William. 2000. "A unified statistical model of conflict onset and escalation." *American Journal of Political Science* pp. 84–93.

Reuveny, Rafael, Brian M. Pollins and Omar MG Keshk. 2010. "The perils of doctrine: a reply to Hegre, Oneal, and Russett." *Journal of peace research* 47(6):1–5.

Russett, Bruce, John R Oneal and David R Davis. 1998. "The third leg of the Kantian tripod for peace: International organizations and militarized disputes, 1950–85." *International Organization* 52(03):441–467.

Singer, J David. 1988. "Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985." *International Interactions* 14(2):115–132.

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 Mangagement and Peace Science* 19(2):58–66.

Weede, Erich. 1976. "Overwhelming preponderance as a pacifying condition among contiguous Asian dyads, 1950–1969." *Journal of Conflict Resolution* 20(3):395–411.

Xiang, Jun, Xiaohong Xu and George Keteku. 2007. "Power The Missing Link in the Trade Conflict Relationship." *Journal of Conflict Resolution* 51(4):646–663.

Appendix

Descriptive statistics are found in Table 3.

Regression results for all spatial and measure choices, 1870–2007, are found in Table 4.
Regression results for all spatial and measure choices, 1870–1938, are found in Table 5.
Regression results for all spatial and measure choices, 1950–1992, are found in Table 6.
Regression results for all spatial and measure choices, 1950–2000, are found in Table 7.
Regression results for all spatial and measure choices, 1950–2007, are found in Table 8.
Regression results for all spatial and measure choices, 1992–2007, are found in Table 9.

Methodological Notes

When running models with all dyads (several temporal domains) we encountered perfect separation. The apparent cause is the peace years variable perfectly predicting zeros; the ideal robustness check is to re-run all models with a Firth correction to penalize the likelihood calculation when there is perfect separation. We suspect that the results will not change, since we do not experience the typical symptom of inflated coefficients or standard errors in these models.

Table 3: Descriptive Statistics

	All Dyads					Politically Relevant Dyads				
Statistic	N	Mean	St. Dev.	Min	Max	N	Mean	St. Dev.	Min	Max
MID Onset	404,390	0.002	0.040	0	1	62,902	0.009	0.094	0	1
Trade (share)	404,390	0.020	0.070	0	1.000	62,902	0.060	0.122	0	1
Trade (flow)	404,390	419.790	5,765.660	0	544,015.000	62,902	2,199.661	14,386.800	0	544,015.000
Trade (openness)	404,390	0.001	0.008	0	1.184	62,902	0.003	0.014	0	1.016
Trade (dependence)	404,390	0.002	0.015	0	1.576	62,902	0.006	0.028	0	1.576
Joint Democracy	404,390	0.151	0.358	0	1	62,902	0.223	0.416	0	1
Power (difference)	404,390	0.014	0.033	0	0.373	62,902	0.057	0.064	0	0.373
Contiguity	404,390	0.057	0.231	0	1	62,902	0.364	0.481	0	1
Distance	404,390	4,460.717	2,735.704	5	12,347	62,902	3,232.248	2,787.067	5	11,989
Peace years	404,390	37.499	33.475	0	277	62,902	48.317	43.292	0	277

Note: Observations are non-directed dyad-year. Data ranges from 1870 to 2007. Missingness eliminated by listwise deletion.

Table 4: Temporal Domain: 1870–2007

	Dependent Variable: MID onset							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trade (share)	0.015*				0.023*			
	(0.008)				(0.007)			
Trade (flow)		0.030				0.010		
		(0.021)				(0.022)		
Trade (openness)			0.037				0.022	
			(0.020)				(0.028)	
Trade (dependence)				0.026				0.004
				(0.029)				(0.033)
Joint democracy	−0.210*	−0.193*	−0.199*	−0.196*	−0.202*	−0.173*	−0.178*	−0.173*
	(0.046)	(0.044)	(0.045)	(0.045)	(0.048)	(0.046)	(0.047)	(0.047)
Power (difference)	0.339*	0.347*	0.348*	0.348*	0.149*	0.165*	0.165*	0.165*
	(0.027)	(0.026)	(0.026)	(0.026)	(0.033)	(0.032)	(0.032)	(0.032)
Contiguity	0.633*	0.628*	0.632*	0.633*	0.346*	0.350*	0.351*	0.352*
	(0.028)	(0.028)	(0.028)	(0.028)	(0.038)	(0.038)	(0.038)	(0.038)
Distance	−0.551*	−0.554*	−0.555*	−0.554*	−0.366*	−0.380*	−0.379*	−0.380*
	(0.091)	(0.091)	(0.091)	(0.091)	(0.116)	(0.116)	(0.116)	(0.116)
Peace years [1]	−271.222*	−275.712*	−276.385*	−276.245*	−148.276*	−150.901*	−151.434*	−151.140*
	(28.808)	(28.697)	(28.726)	(28.728)	(21.044)	(21.010)	(21.009)	(21.029)
Peace years [2]	254.344*	255.957*	253.386*	253.273*	58.715	59.049	58.140	58.500
	(45.035)	(44.954)	(45.017)	(44.988)	(34.881)	(34.895)	(34.866)	(34.910)
Peace years [3]	−431.495*	−429.539*	−433.062*	−432.665*	−204.502*	−203.487*	−204.475*	−203.890*
	(46.441)	(46.403)	(46.465)	(46.471)	(27.810)	(27.843)	(27.838)	(27.886)
Constant	−8.054*	−8.060*	−8.061*	−8.060*	−6.551*	−6.591*	−6.592*	−6.589*
	(0.091)	(0.091)	(0.091)	(0.091)	(0.131)	(0.131)	(0.131)	(0.131)
N	404,390	404,390	404,390	404,390	62,902	62,902	62,902	62,902
Log Likelihood	-3,479.499	-3,480.222	-3,480.104	-3,480.824	-2,680.044	-2,683.666	-2,683.496	-2,683.748
AIC	6,976.999	6,978.444	6,978.207	6,979.649	5,378.087	5,385.331	5,384.991	5,385.495

All coefficients are standardized; standard errors in parentheses.

* indicates significance at $p < 0.05$

Table 5: Temporal Domain: 1870–1938

	Dependent Variable: MID onset							
	All Dyads				Relevant Dyads			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trade (share)	1.607				4.294			
	(9.173)				(8.769)			
Trade (flow)		0.038				0.020		
		(0.050)				(0.051)		
Trade (openness)			−1.922				−1.939	
			(1.655)				(1.731)	
Trade (dependence)				−2.014				−2.020
				(1.691)				(1.760)
Joint democracy	−0.557*	−0.552*	−0.537*	−0.536*	−0.496*	−0.475*	−0.460	−0.458
	(0.241)	(0.236)	(0.236)	(0.236)	(0.243)	(0.237)	(0.237)	(0.237)
Power (difference)	0.030	0.039	0.041	0.040	−0.165	−0.147	−0.141	−0.141
	(0.085)	(0.079)	(0.077)	(0.077)	(0.106)	(0.101)	(0.099)	(0.099)
Contiguity	0.246*	0.241*	0.269*	0.272*	0.093	0.098	0.121	0.123
	(0.079)	(0.078)	(0.077)	(0.078)	(0.087)	(0.086)	(0.087)	(0.087)
Distance	−0.378	−0.378	−0.372	−0.363	−0.402	−0.401	−0.380	−0.367
	(0.240)	(0.241)	(0.242)	(0.242)	(0.264)	(0.264)	(0.266)	(0.266)
Peace years [1]	−70.396*	−70.335*	−66.255*	−66.603*	−55.023*	−54.138*	−51.301*	−51.516*
	(29.389)	(29.088)	(29.384)	(29.448)	(23.853)	(23.670)	(23.823)	(23.849)
Peace years [2]	−25.969	−23.752	−27.103	−27.393	−18.827	−17.229	−19.128	−19.054
	(37.389)	(37.092)	(37.388)	(37.447)	(32.242)	(32.066)	(32.218)	(32.218)
Peace years [3]	−73.054*	−71.531*	−74.998*	−75.101*	−63.846*	−63.194*	−65.587*	−65.559*
	(28.267)	(28.196)	(28.282)	(28.297)	(24.076)	(24.053)	(24.062)	(24.046)
Constant	−6.521*	−6.664*	−6.759*	−6.793*	−5.624*	−5.919*	−6.010*	−6.040*
	(0.650)	(0.311)	(0.327)	(0.338)	(0.641)	(0.367)	(0.372)	(0.382)
N	14,647	14,647	14,647	14,647	9,333	9,333	9,333	9,333
Log Likelihood	-317.471	-317.216	-316.534	-316.527	-278.158	-278.194	-277.403	-277.395
AIC	652.942	652.432	651.068	651.055	574.317	574.388	572.806	572.791

All coefficients are standardized; standard errors in parentheses.

* indicates significance at $p < 0.05$

Table 6: Temporal Domain: 1950–1992

	Dependent Variable: MID onset							
	All Dyads				Relevant Dyads			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trade (share)	0.022				0.037			
	(0.022)				(0.023)			
Trade (flow)		−0.014				−0.039		
		(0.035)				(0.038)		
Trade (openness)			−0.024				−0.285	
			(0.091)				(0.181)	
Trade (dependence)				−0.091				−0.294
				(0.109)				(0.177)
Joint democracy	−0.275*	−0.246*	−0.244*	−0.234*	−0.240*	−0.193*	−0.164*	−0.162*
	(0.078)	(0.071)	(0.073)	(0.073)	(0.080)	(0.073)	(0.074)	(0.074)
Power (difference)	0.473*	0.477*	0.477*	0.476*	0.276*	0.286*	0.287*	0.286*
	(0.033)	(0.033)	(0.033)	(0.033)	(0.044)	(0.043)	(0.043)	(0.043)
Contiguity	0.574*	0.576*	0.575*	0.576*	0.248*	0.260*	0.258*	0.258*
	(0.035)	(0.035)	(0.035)	(0.035)	(0.054)	(0.054)	(0.054)	(0.054)
Distance	−0.800*	−0.806*	−0.806*	−0.809*	−0.815*	−0.824*	−0.838*	−0.842*
	(0.122)	(0.122)	(0.122)	(0.122)	(0.171)	(0.171)	(0.171)	(0.171)
Peace years [1]	−189.655*	−193.201*	−193.113*	−193.075*	−143.508*	−147.290*	−148.038*	−147.587*
	(29.763)	(29.584)	(29.603)	(29.634)	(28.824)	(28.743)	(29.039)	(29.040)
Peace years [2]	145.866*	143.274*	142.594*	141.267*	−41.210	−43.985	−46.449	−46.144
	(47.151)	(47.111)	(47.222)	(47.289)	(41.042)	(41.049)	(41.513)	(41.508)
Peace years [3]	−450.700*	−452.483*	−452.685*	−451.792*	−218.210*	−219.495*	−218.757*	−218.309*
	(46.244)	(46.259)	(46.297)	(46.353)	(29.351)	(29.386)	(29.613)	(29.612)
Constant	−8.172*	−8.174*	−8.175*	−8.176*	−6.827*	−6.859*	−6.872*	−6.877*
	(0.124)	(0.125)	(0.125)	(0.125)	(0.196)	(0.195)	(0.196)	(0.196)
N	234,505	234,505	234,505	234,505	31,420	31,420	31,420	31,420
Log Likelihood	-2,005.206	-2,005.551	-2,005.596	-2,005.164	-1,466.704	-1,467.313	-1,465.923	-1,465.609
AIC	4,028.412	4,029.102	4,029.193	4,028.328	2,951.408	2,952.625	2,949.845	2,949.217

All coefficients are standardized; standard errors in parentheses.

* indicates significance at $p < 0.05$

Table 7: Temporal Domain: 1950–2000

	Dependent Variable: MID onset							
	All Dyads				Relevant Dyads			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trade (share)	0.004				0.012			
	(0.011)				(0.011)			
Trade (flow)		0.023				0.003		
		(0.027)				(0.029)		
Trade (openness)			0.018				−0.036	
			(0.045)				(0.058)	
Trade (dependence)				−0.023				−0.069
				(0.052)				(0.064)
Joint democracy	−0.139*	−0.135*	−0.137*	−0.126*	−0.099	−0.082	−0.074	−0.067
	(0.053)	(0.051)	(0.052)	(0.052)	(0.054)	(0.051)	(0.053)	(0.053)
Power (difference)	0.453*	0.456*	0.456*	0.455*	0.282*	0.292*	0.292*	0.292*
	(0.033)	(0.032)	(0.032)	(0.032)	(0.041)	(0.040)	(0.040)	(0.040)
Contiguity	0.636*	0.632*	0.635*	0.637*	0.359*	0.362*	0.363*	0.363*
	(0.032)	(0.033)	(0.032)	(0.032)	(0.049)	(0.049)	(0.049)	(0.049)
Distance	−0.728*	−0.731*	−0.729*	−0.732*	−0.659*	−0.672*	−0.676*	−0.682*
	(0.114)	(0.114)	(0.114)	(0.114)	(0.158)	(0.157)	(0.158)	(0.158)
Peace years [1]	−286.695*	−287.742*	−288.025*	−287.966*	−217.675*	−219.850*	−220.030*	−219.913*
	(37.184)	(37.070)	(37.035)	(37.023)	(38.810)	(38.767)	(38.765)	(38.765)
Peace years [2]	75.780	74.232	75.245	74.674	−116.735*	−118.284*	−118.471*	−118.488*
	(61.307)	(61.403)	(61.313)	(61.268)	(55.497)	(55.538)	(55.517)	(55.510)
Peace years [3]	−576.954*	−578.489*	−577.878*	−576.678*	−296.872*	−297.580*	−297.111*	−296.433*
	(57.715)	(57.814)	(57.762)	(57.710)	(37.587)	(37.628)	(37.610)	(37.605)
Constant	−8.317*	−8.321*	−8.320*	−8.320*	−7.110*	−7.142*	−7.143*	−7.144*
	(0.116)	(0.116)	(0.115)	(0.116)	(0.182)	(0.181)	(0.181)	(0.181)
N	314,494	314,494	314,494	314,494	43,178	43,178	43,178	43,178
Log Likelihood	-2,498.857	-2,498.587	-2,498.859	-2,498.821	-1,889.403	-1,889.969	-1,889.754	-1,889.236
AIC	5,015.714	5,015.174	5,015.717	5,015.643	3,796.805	3,797.938	3,797.508	3,796.472

All coefficients are standardized; standard errors in parentheses.

* indicates significance at $p < 0.05$

Table 8: Temporal Domain: 1950–2007

	Dependent Variable: MID onset							
	All Dyads				Relevant Dyads			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trade (share)	0.005				0.014			
	(0.008)				(0.008)			
Trade (flow)		0.035				0.014		
		(0.024)				(0.025)		
Trade (openness)			0.027				0.010	
			(0.025)				(0.032)	
Trade (dependence)				0.007				−0.011
				(0.033)				(0.037)
Joint democracy	−0.169*	−0.166*	−0.170*	−0.163*	−0.161*	−0.141*	−0.142*	−0.136*
	(0.048)	(0.046)	(0.046)	(0.047)	(0.050)	(0.048)	(0.048)	(0.048)
Power (difference)	0.436*	0.440*	0.439*	0.440*	0.247*	0.263*	0.263*	0.263*
	(0.030)	(0.029)	(0.029)	(0.029)	(0.037)	(0.036)	(0.036)	(0.036)
Contiguity	0.658*	0.651*	0.657*	0.658*	0.376*	0.378*	0.380*	0.381*
	(0.029)	(0.030)	(0.030)	(0.030)	(0.043)	(0.044)	(0.043)	(0.043)
Distance	−0.622*	−0.627*	−0.624*	−0.625*	−0.481*	−0.503*	−0.500*	−0.503*
	(0.100)	(0.100)	(0.100)	(0.100)	(0.136)	(0.136)	(0.136)	(0.136)
Peace years [1]	−268.939*	−269.437*	−270.991*	−270.729*	−140.655*	−142.739*	−143.066*	−142.717*
	(29.143)	(29.027)	(29.004)	(29.006)	(21.135)	(21.125)	(21.101)	(21.136)
Peace years [2]	239.021*	237.785*	238.667*	238.452*	53.404	52.612	52.574	52.941
	(47.040)	(47.192)	(47.057)	(47.033)	(34.869)	(34.995)	(34.928)	(35.018)
Peace years [3]	−451.154*	−451.861*	−452.326*	−451.524*	−198.928*	−198.560*	−198.843*	−198.071*
	(48.498)	(48.627)	(48.532)	(48.542)	(27.960)	(28.046)	(28.025)	(28.102)
Constant	−8.193*	−8.197*	−8.198*	−8.197*	−6.791*	−6.833*	−6.832*	−6.831*
	(0.099)	(0.099)	(0.099)	(0.099)	(0.149)	(0.148)	(0.148)	(0.148)
N	386,792	386,792	386,792	386,792	53,038	53,038	53,038	53,038
Log Likelihood	-3,093.006	-3,092.110	-3,092.706	-3,093.138	-2,357.765	-2,359.028	-2,359.124	-2,359.127
AIC	6,204.012	6,202.219	6,203.412	6,204.277	4,733.530	4,736.056	4,736.248	4,736.254

All coefficients are standardized; standard errors in parentheses.

* indicates significance at $p < 0.05$

Table 9: Temporal Domain: 1992–2007

	Dependent Variable: MID onset							
	All Dyads				Relevant Dyads			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trade (share)	0.004				0.014			
	(0.009)				(0.009)			
Trade (flow)		0.104*				0.077*		
		(0.035)				(0.036)		
Trade (openness)			0.020				0.016	
			(0.028)				(0.030)	
Trade (dependence)				0.007				0.0004
				(0.037)				(0.037)
Joint democracy	−0.087	−0.090	−0.087	−0.082	−0.122	−0.104	−0.102	−0.096
	(0.064)	(0.060)	(0.061)	(0.062)	(0.067)	(0.063)	(0.064)	(0.065)
Power (difference)	0.293*	0.296*	0.301*	0.302*	0.106	0.144*	0.144*	0.145*
	(0.063)	(0.058)	(0.058)	(0.058)	(0.071)	(0.065)	(0.065)	(0.065)
Contiguity	0.856*	0.828*	0.852*	0.854*	0.612*	0.594*	0.613*	0.614*
	(0.056)	(0.057)	(0.056)	(0.056)	(0.075)	(0.075)	(0.075)	(0.075)
Distance	−0.226	−0.247	−0.230	−0.232	0.161	0.094	0.120	0.116
	(0.175)	(0.174)	(0.174)	(0.174)	(0.217)	(0.216)	(0.217)	(0.217)
Peace years [1]	−200.085*	−197.849*	−201.357*	−201.188*	−92.801*	−92.847*	−94.320*	−93.923*
	(34.719)	(34.705)	(34.639)	(34.677)	(18.953)	(18.975)	(18.953)	(18.985)
Peace years [2]	224.096*	221.003*	223.336*	223.417*	97.936*	97.261*	97.330*	97.969*
	(30.957)	(30.815)	(30.880)	(30.950)	(20.228)	(20.135)	(20.249)	(20.441)
Peace years [3]	−146.182*	−142.372*	−146.788*	−146.134*	−74.301*	−71.607*	−73.755*	−72.999*
	(34.787)	(34.465)	(34.603)	(34.830)	(19.869)	(19.825)	(19.917)	(20.065)
Constant	−8.483*	−8.475*	−8.488*	−8.487*	−6.927*	−6.995*	−6.993*	−6.989*
	(0.176)	(0.176)	(0.176)	(0.176)	(0.239)	(0.238)	(0.238)	(0.238)
N	161,116	161,116	161,116	161,116	23,009	23,009	23,009	23,009
Log Likelihood	-1,082.825	-1,078.917	-1,082.681	-1,082.876	-890.581	-889.450	-891.443	-891.574
AIC	2,183.650	2,175.835	2,183.361	2,183.753	1,799.162	1,796.900	1,800.886	1,801.148

All coefficients are standardized; standard errors in parentheses.

* indicates significance at $p < 0.05$