A Menace to the Democratic Peace? Dyadic and Systemic Difference*

RESPONSE

SEUNG-WHAN CHOI
University of Illinois at Chicago

Do democracies become less peaceful with one another as the overall number of democratic states in the international system increases? Gartzke and Weisiger (2013, 2014) claim that they do. However, I argue that their evidence stems from a mathematical error in their statistical model. Once I correct that error, their findings no longer hold. In other words, when Gartzke and Weisiger’s model receives proper specification, no dyadic and systemic difference persists, while democracy emerges as a pacifying factor that reduces the likelihood of interstate dispute.

This study replicates two recent studies by Gartzke and Weisiger (2013, 2014) that seek to challenge the findings of democratic-peace proponents. Gartzke and Weisiger argue that as democracies become more prevalent in the international system, the threat posed to them by autocratic states diminishes. Moreover, heterogeneity among democratic states reduces the pacifying effect of democratic dyads. Both studies test this theory by pitting democracy against dyadic and systemic difference in the same model. Gartzke and Weisiger report that the latter two factors eliminate the significance of the former. In their closing remarks, they (2014:142) conclude that “we find no evidence that a more democratic world is a less conflictual one.” If such a finding withstands scrutiny, then Gartzke and Weisiger’s work deals a harsh blow to the democratic-peace hypothesis.

In this response, I re-evaluate the validity of Gartzke and Weisiger’s findings. I focus on their mathematical and statistical analysis rather than the logic of their theoretical argument. I show that Gartzke and Weisiger’s findings result from a statistical artifact. When I correct their model, I find that the difference that they identify between dyadic and systemic analysis disappears. Instead, democracy emerges as a pacifying influence on international conflict. Simply put, democratic-peace theory prevails over rival interpretations.

A Brief Literature Review

Gartzke and Weisiger try to debunk the theory that democracies are less likely to go to war against each other (see Leeds 1999; Mitchell, Gates, and Hegre 1999; Mitchell 2002; Kadera, Crescenzi, and Shannon 2003; Choi and James 2005, 2007; Choi 2010b, 2013; Russett 2010; Colaresi 2012; Hayes 2012; Park 2013). In particular, their studies fall within a long tradition of debating, critiquing, and refining the specific democratic-peace research program developed by Oneal and Russett (1999, 2005), and Russett and Oneal (2001). For example, James, Solberg, and Wolfson (1999) argue that when we treat democracy and peace as parts of an endogenized system, peace drives democratization more than democratization drives peace. In earlier work, Gartzke (2007) proposes a research program centered on a putative “capitalist peace.” He claims that capitalism, rather than democracy, explains the apparent pacifying effects of democracy on interstate conflict (see also Barbieri 1996). Based on pooled panel data for 114 countries from the years 1970 to 2001, Choi (2010a) reports that socioeconomic and political globalization exerts a more important pacifying effect than that of democracy and economic interdependence. Gibler (2012) contends that the democratic-peace phenomenon is only a by-product of territorial peace. In light of these challenges, a number of scholars endeavor to show the vitality of democratic-peace theory (see, for example, Oneal and Russett 2000; Choi 2011; Dafoe 2011; Crescenzi and Kadera 2015).

Mathematical Proofs

Gartzke and Weisiger (2013, 2014) challenge democratic-peace theory. They describe their theory in terms of “dynamic difference” that is separated into differences at the dyadic and systemic levels. In practical terms, the strength of democracy as the basis for collective identity shifts as the number of democracies in the system increases. When democracies become numerous, they no longer see one another as part of a besieged community of confederates.
Gartzke and Weisiger add both dyadic and systemic difference to a standard democratic-peace model in the following fashion:

\[
\text{Interstate Conflict} = \gamma_1 + \gamma_2 \text{Low Democracy} + \gamma_3 \text{Dyadic Difference} + \gamma_4 \text{Systemic Difference} + \gamma_{5 \times k} \text{Control Variables} + \epsilon_1
\]  

(1)

Gartzke and Weisiger’s estimated results, as presented, suggest that the inclusion of the variables “Dyadic Difference” and “Systemic Difference” causes “Low Democracy” to become insignificant. If this finding proved valid and robust, we would be looking at a credible menace to democratic-peace theory. However, their model suffers from important mathematical problems that I correct below. Once corrected, their evidence against democratic-peace theory disappears.

In their studies, the variable Low Democracy takes in the lower score between the two Polity democracy scores in a given dyad; they construct Dyadic Difference by deducting the low democracy score from the high democracy score in a dyad; and they measure Systemic Difference as “annual observations of the standard deviation in (non-missing) polity scores for the system” (Gartzke and Weisiger 2014:138).

Given the fact that Gartzke and Weisiger calculated Dyadic Difference by deducting Low Democracy from High Democracy, the first equation can be expressed as follows:

\[
\text{Interstate Conflict} = \gamma_1 + \gamma_2 \text{Low Democracy} + \gamma_3 (\text{High Democracy} - \text{Low Democracy}) + \gamma_4 \text{Systemic Difference} + \gamma_{5 \times k} \text{Control Variables} + \epsilon_1
\]

(2)

When rearranged, the second equation becomes:

\[
\text{Interstate Conflict} = \gamma_1 + (\gamma_2 - \gamma_3) \text{Low Democracy} + \gamma_3 \text{High Democracy} + \gamma_4 \text{Systemic Difference} + \gamma_{5 \times k} \text{Control Variables} + \epsilon_1
\]

(3)

The model contains a biased coefficient for Low Democracy and a misleading coefficient for Dyadic Difference. When we compare the effect of Low Democracy in Equation 1 to that in Equation 3, the coefficient of the latter is a sum of \((\gamma_2 - \gamma_3)\) instead of \(\gamma_2\) itself; thus, the coefficient on Low Democracy in Equation 3 is incorrectly estimated and thus biased. Accordingly, Gartzke and Weisiger’s biased estimation causes Low Democracy to become statistically insignificant because the estimated coefficient turns out to be \(\gamma_2\) minus \(\gamma_3\), not \(\gamma_2\). More importantly, Dyadic Difference disappears in Equation 3. This means that Gartzke and Weisiger failed to test the effect of Dyadic Difference as claimed; instead, they produced and reported an estimated coefficient for High Democracy. In the section of Empirical Proofs, I provide evidence that the estimated coefficients and standard errors for Dyadic Difference and High Democracy are exactly identical. This implies that what Gartzke and Weisiger call Dyadic Difference is actually High Democracy, as the estimated coefficient is \(\gamma_3\) as deduced in Equation 3.

Although Systemic Difference achieves statistical significance in Gartzke and Weisiger’s Table 2 (2013:178) and Table 3 (2014:141), it does not measure their theoretical framework’s independent variable. As I noted earlier, they measure Systemic Difference as “annual observations of the standard deviation in (non-missing) polity scores for the system” (2014:138). The standard deviation of Systemic Democracy fails to capture whether the international system consists of more or less democratic political systems, but it is rather an indication of how tightly all the polities are clustered around the mean in the data set. For example, an average democracy score of 7 and an average autocracy score of −7 yield the same positive value of standard deviation for dyadic years.

How should we correct the measurement error in Gartzke and Weisiger’s studies such that the standard deviation in (non-missing) polity scores “would be the same whether 90% of the countries were democracies or 90% were autocracies” (Dafoe, Oneal, and Russet 2013:210)? We need an alternative that captures what Gartzke and Weisiger intend to operationalize with respect to Systemic Difference—whether the world is populated with more or fewer democracies. Simply put, a new measure must be able to identify increases and decreases in regime type heterogeneity (difference) at the systemic level.

Because skewness is a measure of the degree of asymmetry of a distribution whose value can be positive or negative (Agresti and Finlay 2009), it provides a good alternative to the standard deviation for assessing how different political systems are distributed each year. When the distribution of polities is skewed to the left, we obtain a negative value that we can interpret as the presence of more autocracies in the international system. When the distribution is right-skewed, we obtain a positive score that indicates the existence of more democracies. In the Empirical Proofs section, I show how dramatically estimated results change when we substitute this more appropriate measure.

Empirical Proofs

I re-examine Gartzke and Weisiger’s Model 14. I do so for two reasons. First, it constitutes the main model that includes all of the key dyadic and systemic variables to test the essence of their theoretical claims. Second, it is also an extended version of a model that appears in their earlier publication (2013:178) on this topic. I report the replicated results in Table 1. To show the bias related to Dyadic Difference in the presence of Dyadic Democracy (low), Model T1 in Table 1 replaces the former with Dyadic Democracy (high).

Comparing the results in Models 14 and T1 with respect to the effects of Dyadic Difference and Dyadic Democracy (high) reveals that both have exactly the same coefficient (0.105) and standard error (0.023). This strongly suggests that, as I argued in the previous section, what Gartzke and Weisiger call “Dyadic Difference” is, in fact, “Dyadic Democracy (high).” This matters a great
deal, because the straightforward interpretation of Dyadic Democracy (high) in Model T1 is that holding Dyadic Democracy (low) constant, the likelihood of a fatal dispute rises as the polity score of the more democratic countries increases.

Dyadic Democracy (low) becomes insignificant when pitted against Dyadic Difference in Gartzke and Weisiger’s Model 14. This is because (as noted earlier) Dyadic Difference causes the estimated coefficient on Dyadic Democracy (low) to become $\tilde{\gamma}_5$ minus $\tilde{\gamma}_5$, not $\tilde{\gamma}_5$. This erroneous result leads Gartzke and Weisiger to highlight the significant effect of Dyadic Difference over the insignificant effect of Dyadic Democracy (low): “difference is most likely an important proximate determinant of conflict” (2013:182, emphasis in original). Yet, note that Dyadic Democracy (low) becomes statistically significant in the hypothesized direction when I replaced Dyadic Difference with Dyadic Democracy (high) in Model T1. The corrected model apparently confirms a dampening effect of Dyadic Democracy (low) on interstate dispute. The comparison reveals that the significance of Dyadic Democracy (low) changes dependent upon the inclusion or exclusion of the wrong measure, Dyadic Difference, in the model; it clearly negates Gartzke and Weisiger’s assertion about the effectiveness of Dyadic Difference.3

Model T2 in Table 1 corrects the bias that emerges from the inclusion of another erroneous measure, Systemic Difference, in the Gartzke and Weisiger studies. While the Gartzke and Weisiger studies incorrectly turn to the standard deviation of Dyadic Democracy (low) to measure Systemic Difference, I employ its skewness to remedy the measurement error. When I implement the remedy in Model T2, I find the significant effects of three systemic-related variables including Systemic Difference, while the significance of Dyadic Democracy (low) and Dyadic Democracy (high) remains intact.5 These results are inconsistent with Gartzke and Weisiger’s (2014:130) claim that they attempt to “separate out the effects of aggregate democracy from regime type difference... [and find that] greater systemic development encourages peace, difference propagates war, and increased

---

3To provide more evidence related to the bias of Dyadic Difference, Appendix I presents a replicated result of Gartzke and Weisiger’s (2013:178) Table 2. Models 2.1, 2.2, 2.3, and 2.4 display the replicated estimates of the Gartzke and Weisiger study, while Models A1 and A2 include Democracy (high) in place of Dyadic Difference. Across models, the coefficient and standard error of Democracy (high) are shown to be identical to those of Dyadic Difference. Also note the emergence of the significant Democracy (low) variable with a correct coefficient sign in Models 2.1, A1 and A2.

5Gartzke and Weisiger’s (2013:178) study also proposes “Proportion of Democracy” as another indicator of Systemic Difference. It measures a ratio of the number of democracies whose Polity score is greater than 6 to the total number of countries per dyadic year. However, the same variable is introduced as an alternative measure of Systemic Democracy in Gartzke and Weisiger’s (2014:141) study—and without any explanations. Because the way that the two different studies by Gartzke and Weisiger introduce the variable of Proportion of Democracy is confusing, I do not discuss its validity as a measure of Systemic Difference.

6Because the new measure, skewness, for Systemic Difference appears to wipe out the significance of Systemic Development—another key variable for the Gartzke and Weisiger research—this study runs another robustness test where skewness is dropped out of Model T2. The results show that Systemic Development fails to achieve significance even in the absence of the new Systemic Difference measure, while Dyadic Democracy (low) remains significant.
systemic democracy has no consistent impact on interstate conflict.”

When a sample size is large, statistical significance does not necessarily indicate a meaningful finding in a practical sense. Accordingly, we should estimate the substantive effects of variables for the purpose of empirical verification. Among the eleven variables in Model T2, Dyadic Democracy (low) is the most theoretically interesting, significant variable so that its substantive effect is calculated by setting the continuous variables in the estimations at their means and the dichotomized variables at 1. I then adjust Dyadic Democracy (low) one at a time to see the change in the predicted probability of a fatal dispute. I find that the likelihood of a fatal dispute reduces by nine percent when Dyadic Democracy (low) changes from the mean to the mean plus one standard deviation, holding all other factors constant. The change to two standard deviations of Dyadic Democracy (low) is related to a decrease of eighteen percent in the likelihood of a fatal dispute. Thus, this substantive effect analysis confirms the pacifying effect of democratic pairs.

**Closing Remarks**

The two recent studies by Gartzke and Weisiger (2013, 2014) report that when scholars take into account dynamic difference in the context of the standard democratic-peace model, the peacefulness of democratic pairs evaporates into the air, while “peace at the system level is linked to development” (2014:142). To some extent, these findings undermine the research programs of the democratic peace. Peace scholars have therefore taken the Gartzke and Weisiger studies as a serious challenge to their findings (e.g., Dafoe et al. 2013). In the spirit of Popper’s (1963:216) philosophy—“science is perhaps the only human activity in which errors are systematically criticized and, in time, corrected”—I took a close look at their studies. I found that Gartzke and Weisiger’s claims lack substantiation because their estimated results derive from incorrect measures of dynamic difference. After correctly specifying Gartzke and Weisiger’s models, I demonstrate that Dyadic and Systemic Difference no longer matters. Indeed, the democratic-peace theory’s expectations concerning the pacifying effects of democracy hold up. Moreover, greater Systemic Development does not lead to a more peaceful world.

My aim is not to foreclose further investigation of the research programs suggested by Gartzke and Weisiger. Rather, the presence of mistakes and errors represents part of the scholarly voyage into the universe of scientific truth. Unintended errors are integral to future fact-finding as they push us to seek better models and more accurate conceptualizations. The history of the natural sciences suggests the wrong concept of geocentrism was the necessary precursor to heliocentrism, before the latter became accepted as scientific truth. Similarly, in our own field, many scholars were less keen to explore alternative international relations theories until witnessing the failure of Kenneth Waltz’s neorealism to predict the end of the Cold War (see Lebow and Risse-Kappen 1996). Accordingly, I hope that more political scientists—including Gartzke and Weisiger—will engage in a constructive exchange of ideas and criticisms through replication projects.

**References**


Appendix 1

Table A1. Low Democracy, High Democracy, and Dyadic Difference: Gartzke and Weisiger (2013), Models 2.1–2.4, Table 2.178

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 2.1</th>
<th>Model 2.2</th>
<th>With Democracy (high)</th>
<th>Replicated Model A1</th>
<th>Model 2.3</th>
<th>With Democracy (high)</th>
<th>Replicated Model A2</th>
<th>Model 2.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democracy (Low)</td>
<td>-0.071 (0.015)**</td>
<td>0.038 (0.017)**</td>
<td>-0.070 (0.016)***</td>
<td>0.040 (0.021)*</td>
<td></td>
<td>-0.232 (0.034)***</td>
<td>0.023 (0.020)</td>
<td></td>
</tr>
<tr>
<td>Democracy (High)</td>
<td><strong>0.109 (0.015)</strong>*</td>
<td><strong>0.111 (0.017)</strong>*</td>
<td></td>
<td></td>
<td></td>
<td><strong>0.255 (0.031)</strong>*</td>
<td><strong>0.255 (0.031)</strong>*</td>
<td></td>
</tr>
<tr>
<td>Dyadic Difference</td>
<td><strong>0.109 (0.015)</strong>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systemic Difference</td>
<td>-0.154 (0.436)</td>
<td>-0.154 (0.436)</td>
<td></td>
<td></td>
<td></td>
<td>1.704 (0.498)***</td>
<td>1.704 (0.498)***</td>
<td></td>
</tr>
<tr>
<td>Dyadic * Systemic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance (ln)</td>
<td>-0.394 (0.017)***</td>
<td>-0.394 (0.017)***</td>
<td>-0.394 (0.017)***</td>
<td>-0.394 (0.017)***</td>
<td></td>
<td>-0.394 (0.016)***</td>
<td>-0.394 (0.016)***</td>
<td></td>
</tr>
<tr>
<td>Contiguity</td>
<td>-0.298 (0.047)**</td>
<td>-0.298 (0.047)**</td>
<td>-0.297 (0.047)***</td>
<td>-0.297 (0.047)***</td>
<td></td>
<td>-0.295 (0.045)***</td>
<td>-0.295 (0.045)***</td>
<td></td>
</tr>
<tr>
<td>Alliance</td>
<td>0.059 (0.111)</td>
<td>0.059 (0.111)</td>
<td>0.062 (0.112)</td>
<td>0.062 (0.112)</td>
<td></td>
<td>0.070 (0.110)</td>
<td>0.070 (0.110)</td>
<td></td>
</tr>
<tr>
<td>Capability Ratio</td>
<td>1.355 (0.371)***</td>
<td>1.355 (0.371)***</td>
<td>1.350 (0.373)***</td>
<td>1.350 (0.373)***</td>
<td></td>
<td>1.357 (0.365)***</td>
<td>1.357 (0.365)***</td>
<td></td>
</tr>
<tr>
<td>Major Power</td>
<td>1.386 (0.151)***</td>
<td>1.386 (0.151)***</td>
<td>1.378 (0.156)***</td>
<td>1.378 (0.156)***</td>
<td></td>
<td>1.404 (0.154)***</td>
<td>1.404 (0.154)***</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.791 (0.351)***</td>
<td>-0.791 (0.351)***</td>
<td>-0.768 (0.360)***</td>
<td>-0.768 (0.360)***</td>
<td></td>
<td>-1.245 (0.562)***</td>
<td>-1.245 (0.562)***</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-10886.9</td>
<td>-10886.9</td>
<td>-10886.66</td>
<td>-10886.66</td>
<td></td>
<td>-10858.81</td>
<td>-10858.81</td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>2117.60</td>
<td>2117.60</td>
<td>2156.35</td>
<td>2156.35</td>
<td></td>
<td>2250.75</td>
<td>2250.75</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>634,684</td>
<td>634,684</td>
<td>634,684</td>
<td>634,684</td>
<td></td>
<td>634,684</td>
<td>634,684</td>
<td></td>
</tr>
</tbody>
</table>

(Note. *p < .05; **p < .01; ***p < .001, two-tailed tests.) The numbers in bold are discussed in footnote 3.)