

Stylized Facts and Comparative Statics in (Social) Science Inquiry

RESPONSE

NEIL NARANG

Department of Political Science, University of California, Santa Barbara

AND

BRAD L. LEVECK

Department of Political Science, University of California, Merced

In a recent article, David Blagden (2019) critiques our research published in the *International Studies Quarterly* (LeVeck and Narang 2017a), in which we draw on the well-known “wisdom of crowds” phenomenon to argue that, because democracies typically include a larger number of decision makers in the foreign policy process, they may produce fewer decision-making errors in situations of crisis bargaining. As a result, bargaining may fail less often. Blagden’s critique focuses on two supposed flaws: first, that “[d]emocracies may have a larger number of more diverse policymakers, of course, but this relationship is not necessary,” and second, that “weighing against the superior ability of large groups to average towards accurate answers, meanwhile, is a substantial drawback of larger groups: the diminishing ability to take and implement decisions” due to additional veto players. In this article, we demonstrate the ways in which we believe Blagden’s critique to be misguided in its approach to social science inquiry. In particular, we argue that much of his critique requires that we reject two hallmarks of scientific inquiry: the use of stylized facts in theory building; and the use of comparative statics to generate testable hypotheses.

In a recent response article, David Blagden (2019) engages with our research published in the *International Studies Quarterly*, in which we propose a new theoretical mechanism for democratic peace that highlights a previously unexplored advantage enjoyed by democracies in crises (LeVeck and Narang 2017a). Specifically, we argue that because democracies typically include a larger number of decision makers in the foreign policy process, they may produce fewer decision-making errors in situations of crisis bargaining and thus bargaining among larger groups of diverse decision makers may fail less often. Using data from experiments in which subjects engage in ultimatum bargaining games, we find strong support for the idea that collective decision making among larger groups of decision makers decreases the likelihood of bargaining failure when compared to the performance of individuals, smaller groups, and even foreign policy experts.

At the same time that Blagden is generous in noting that “crowd wisdom is a powerful insight” and an “important contribution to comparative analyses of strategic effectiveness,” he is forceful in his critique that our “logics are of only limited applicability to democracies’ specific foreign policymaking processes,” and that our findings ultimately “tell us little about the peacefulness or otherwise of democracies,” and thus they “tell us little about democratic peace specifically.” In sum, “the ‘wisdom of crowds’—while a powerful insight—does not in fact do much to support democratic peace theory...it simply does not apply well to the

dynamics and surrounding politics of democratic foreign policy-making.”

We welcome Blagden’s response to our research article. Intellectual critique and debate are essential hallmarks of healthy academic inquiry. For this reason, we applaud the editors of the *International Studies Quarterly* for inviting our response to Blagden’s comments.

In what follows, we highlight the ways in which we believe Blagden’s critique to be misguided in its approach to social science inquiry. In particular, we argue that much of his critique requires that we reject two hallmarks of scientific inquiry: the use of stylized facts in theory building; and the use of comparative statics to test the hypotheses generated by scientific theories.

Stylized Facts and Theories

While Blagden is quick to acknowledge that “[d]emocracies may have a larger number of more diverse policymakers,” he immediately counters that “this relationship is not necessary.” For this reason, Blagden pointedly argues that the conditions under which crowds are known to be collectively “wise,” “may well correlate somewhat with the presence of democracy, but that is neither a necessary nor a sufficient finding for anything of much interest.”

We largely agree with the first part of this argument. That democracies do not *necessarily* have a larger number of more diverse decision makers in all cases is uncontroversial. Indeed, even passive observers of international relations should be able to recruit at least one or two historical cases of crisis bargaining where sides sought to estimate their opponent’s reservation price for war, “in which the politics of democratic foreign policy dramatically [shrunk] the number of *de facto* influential decision-makers in the foreign policy process” (though, oddly, neither of Blagden’s examples

Neil Narang is an associate professor of political science at the University of California, Santa Barbara and faculty affiliate at the Center for International Security and Cooperation at Stanford University. Formerly, he was a Council on Foreign Relations’ international affairs fellow serving as a senior adviser in the Office of the Secretary of Defense for Policy.

Brad L. LeVeck is an assistant professor of political science at the University of California, Merced.

in Brexit and the election of Donald Trump seem to qualify as instances).¹

However, our argument was never that larger decision-making groups were a *necessary* feature of foreign policymaking in democracies. Instead, we build our theory around the stylized fact that there is a correlation between the number of effectively independent decision makers and regime type. We argue that this may partially explain the well-known correlation between regime type and conflict. As we explain below, this is useful even if there are numerous instances where democratic decision making does not include a larger number of diverse decision makers.

A stylized fact is a term used to refer to a simplified presentation of an empirical finding, which—although generally true—may have inconsistencies in the detail. The origin of the term is typically traced back to the economist Nicholas Kaldor in 1961 (Kaldor 1961), who stated six now famous “stylized” facts that economists had collectively learned about economic growth—findings that were so widely accepted to be true that they had moved from research papers to textbooks, since they were no longer thought to be particularly controversial as general empirical statements (Jones and Romer 2010).

Importantly, Kaldor (1961) argued that because “facts as recorded by statisticians, are always subject to numerous snags and qualifications, and for that reason are incapable of being summarized,” theorists “should be free to start off with a stylized view of the facts—i.e., concentrate on broad tendencies, ignoring individual details.” Others have made the even stronger argument that, for theorizing to ever occur, “stylization is indispensable” because theories require some level of generalization and scrutiny of any fact is likely to produce counterexamples within a given setting (Jan Tinbergen 1939, as translated in Don and Verbruggen 2006).

Given these arguments, it is perhaps unsurprising that stylized facts form the basis of countless theories across the social sciences (Hirshman 2016). This includes the field of international relations, where theories of international institutions, trade, and war all appeal to some set of stylized facts—either in the identification of a macro phenomenon to be explained, or in the causal process linking the independent variables to dependent variables.

Of course, stylized facts should still be reasonable and generally true. To this end, we spent many pages of our original article and our appendix carefully detailing multiple theoretical and empirical reasons for why it is *generally accepted* that democracies are characterized by larger groups of independently deciding individuals *at multiple levels of governance* when compared to autocracies (Hyde and Saunders 2020).

Theoretically, we first argued that by holding periodic elections, citizens can express their views on which leader or mix of representatives is best suited to conduct international

¹With respect to Blagden’s general critique that some questions in international relations do not have “discoverably correct answers,” we refer readers to the supplementary information for our original publication, in which we review a distinction in the psychology literature between “intellective tasks,” where there is a clear *ex-post* evaluation criterion (as with guessing the reservation price of an opponent in situations of ultimatum/crisis bargaining), and “judgmental tasks,” where there is no clear *ex-post* evaluation criterion, as in the case of Brexit referenced by Blagden. We make no claim about the ability of crowds to systematically outperform individuals in judgmental tasks. Meanwhile, Blagden’s vague reference to the election of Donald Trump casting “*prima facie* doubt on the argument that democracies...accurately assess others’ signals” would need to be much further elaborated to warrant a meaningful response here (Hafner-Burton, Narang, and Rathbun 2019); however—whatever the exact claim—the election of Donald Trump is clearly not a case of crisis bargaining nor is it obviously an intellective task, and thus it is unclear how it is relevant to the theoretical mechanism outlined in our article.

Table 1. Difference in institutional constraints and veto points between democracies and autocracies

Regime	Mean XCONST	Mean POLCONIII
Democracy	6.31	0.40
Autocracy	0.76	0.01

affairs. Second, citizens in democracies can more efficiently express approval or disapproval for their leader’s policies through public polls, which should lead leaders to anticipate this accountability to constituencies *ex ante* when making foreign policy decisions.² Third, democracies tend to have freer markets with exchanges that can react almost instantly to inform leaders about the expected outcome of a particular policy choice. Fourth, democracies tend to establish different domestic institutions with diverse approaches or perspectives on foreign policy. Finally—even when aggregating beliefs across similar numbers of individuals—participants in autocracies often lack the incentive to tell leaders the truth (Reiter and Stam 2002). As a result, democratic leaders may have access to a larger number of diverse perspectives, even if they hear from a similar number of advisers.

Empirically, we argued that there is little controversy in the academic literature that democracies tend to have a larger group of decision makers involved in the foreign policy process. In our original article, and in Table 1, we demonstrate that the Polity IV index measure—on which the democratic peace phenomenon is based—is primarily driven by the variable XCONST, which reflects the fact that a larger number of independent actors constrains the executive. Meanwhile, the Polity IV index is highly correlated with the Political Constraint Index (POLCONIII), which is used in some robustness checks of democratic peace, and which also measures the raw number of institutional veto players in a polity and their relative independence in terms of preferences and ideological viewpoints. For these reasons, and many others in the article, we felt more comfortable with the stylized fact that democracies are generally characterized by larger groups of independently deciding individuals compared to autocracies in our experimental setup.

Blagden does not take direct issue with many of these points. In fact, as we already noted, he concedes the general correlation between democracy and larger groups of decision makers. So, his determination that this relationship is not of interest because it is not a necessary feature of democracy suggests that he is antagonistic towards the use of such stylized facts in theory building. For all of the reasons given above, we disagree: the “wisdom of crowds” hypothesis may help us understand the origins of democratic peace, even if there are examples where democratic policymaking does not draw on larger or wiser crowds.

However, we should also note that, at the same time that Blagden’s response appears to be antagonistic to the role of stylized facts in theory building, it is also confusing in its contradictions. For example, Blagden seems willing to accept the stylized fact of the democratic peace finding as a phenomenon worthy of explanation, even though the link between democracy and war is not perfect and depends critically on a series of qualifications (“few, if any, clear cases of war,” “between mature democratic states,” etc.). However, Blagden offers an even more startling contradiction in concluding that:

²We disagree with Blagden’s un-strategic view of elected leaders when he argues that expectations of future electoral accountability do nothing to influence leaders’ present incentive to align their judgment with that of their constituents.

Insofar as a genuine inter-democratic peace may exist, therefore, we are back to some combination of intersubjective recognition (e.g., Hayes 2011), norms (e.g., Mitchell 2012), and democracies' superior ability to make binding commitments to one another (e.g., Lipson 2005), along with all the over-determining co-variables (the U.S. alliance network, U.S. power preponderance, extended nuclear deterrence, economic interdependence, regional organizations, etc.).

Oddly, Blagden's critique of our article should apply equally to all of these arguments as well. Perhaps it is possible that democracies are more likely to share intersubjective recognition, liberal norms, and a superior ability to make binding commitments to one another. However, to use Blagden's words, "this relationship is not necessary." It is unclear, then, why Blagden concludes that "the 'wisdom of crowds'—while a powerful insight—does not in fact do much to support democratic peace theory" at the same time that he is willing to accept other theories based on even more stylized facts.

Comparative Static Claims and Theories

In a second criticism, Blagden notes that our experiment only addresses variation in the number of independent decision makers across regime types. It does not address other institutional factors that may covary with regime type and also affect a country's ability to bargain effectively. For example, according to Blagden, larger groups may lead to the "decision-making sclerosis of too many veto players...rather than ever-larger group size delivering ever-greater crowd wisdom."

However, such critiques largely misunderstand the logic behind our experimental design and the inferences we try to draw from it. Given our broader theory, we experimentally tested the claim that larger groups of decision makers characteristic of democracies would produce collective *judgments* that have a lower likelihood of bargaining failure when compared to the performance of individuals, small groups, and even foreign policy experts. In doing so, we were effectively testing a comparative static claim, which asks how an output of a model changes with a parameter of the model, holding all else constant (Ashworth and Bueno de Mesquita 2006). We were not testing the more ambitious claim that all features of democracy point in the direction of wiser collective decision making.

Comparative static claims are central and common across political science research (Cameron and Morton 2002), including analyses of party behavior under conditions of increasing electoral competitiveness, deference to precedence under conditions of changing judicial independence, and the value of concessions officered in situations of international crisis bargaining under changing conditions (Ashworth and Bueno de Mesquita 2006). While Ashworth and Bueno de Mesquita (2006, 214) note that "formal theorists typically base the testable predictions of their models on comparative statics—the analysis of how changes in the parameters of a model affect the model's solution," empirical scholars also engage in a comparative static analysis when they aim to identify the impact of a particular variable on an outcome of interest, *ceteris paribus*.³

Importantly, the "all else constant" part of comparative static claims is not a claim that reality is simpler than it is. Instead, this assumption helps us think through and test

³See LeVeck and Narang (2017b) and Narang and LeVeck (2019) for examples of comparative static analyses from our own work.

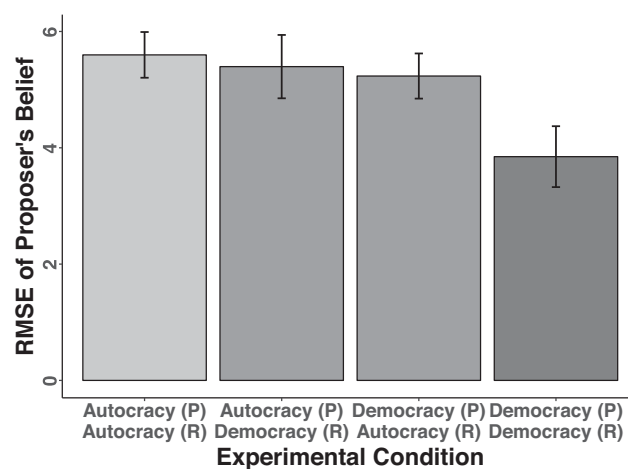


Figure 1. Mean error in proposers' beliefs about responders' minimum acceptable offer

the causal effect of specific variables, even though they may operate as part of a more complex system. All else may not be constant across regime types, and democracy may also correlate with other features that influence decision-making outputs—including a larger number of veto players. However, these possibilities do not impeach the validity of our experimental findings. Given that factors like veto player induced gridlock are somewhat separable from group size (Tsebelis 2000), it was reasonable for our study to temporarily set them aside.⁴ Doing so allowed us to rigorously and clearly test a more basic claim about group size, collective judgments, and ultimatum bargaining outcomes. Were we to simultaneously increase factors like group size and the number of veto players, we would lose the key advantage that experiments afford in terms of isolating specific causal effects.

By holding "all else constant" and measuring the collective judgments of both proposers and responders in our experiment, we were also able to show that the "wisdom of crowds" mechanism can also account for the *dyadic* nature of democratic peace finding, contrary to Blagden's claim that "crowd wisdom...is a monadic effect." In the supplementary information to the original article, and in Figure 1, we demonstrate that the mean error in the proposer's forecasts about a responder's minimum acceptable offer decreases slightly when either the proposer or the responder is characterized by a larger decision-making group (a monadic effect), but that this error is significantly (and nonlinearly) lower when both the proposer and the responder draw on larger decision-making groups (a dyadic effect). This dyadic effect occurs because the responder's reservation price for costly conflict becomes less variable under aggregation. This decreased variance causes the responder's reservation price to be closer to very accurate forecasts about what the responder's reservation price will be in expectation.

⁴Ironically, Blagden's critique on this account would be flawed by his own standards, insofar as he implies gridlock necessarily increases with group size—something that may only be true under much stronger assumptions that are unstated by Blagden. Properly understood, the literature on veto players argues that the possibility for gridlock will turn on many other factors beyond group size, including the heterogeneity in preferences among actors and the specific institutions in which they are imbedded (Tsebelis 2000)—factors we purposely held constant in our experiment.

Conclusion

As we noted at the end of our original article, we see our findings as one step in a larger research program. No single experiment can hope to test every facet of a theory, and our study is no exception. We therefore hope that future research will build on our findings. This includes examining the scope conditions that determine when democracies include more decision makers compared to autocracies. Likewise, other studies might examine whether or when other institutional features outweigh the benefits of larger decision-making groups. However, the need for future research in these areas, and others, is a far cry from Blagden's claim that there are "key flaws" with our theory, or that our findings "tell us little about the peacefulness or otherwise of democracies." Rather than undermine our study, Blagden's arguments do more to highlight the epistemological and methodological differences between our scientific approach and his own.

References

- ASHWORTH, SCOTT, AND ETHAN BUENO DE MESQUITA. 2006. "Monotone Comparative Statics for Models of Politics." *American Journal of Political Science* 50 (1): 214–31.
- BLAGDEN, DAVID. 2019. "Do Democracies Possess the Wisdom of Crowds? Decision Group Size, Regime Type, and Strategic Effectiveness." *International Studies Quarterly* 63 (4): 1192–5.
- CAMERON, CHARLES, AND REBECCA MORTON. 2002. "Formal Theory Meets Data." In *Political Science: The State of the Discipline*, Vol. III, edited by Ira Katznelson and Helen Milner, 784–804. New York: W.W. Norton.
- DON, FRANS JOOS HENDRIK, AND JOHAN PETER VERBRUGGEN. 2006. "Models and Methods for Economic Policy: 60 Years of Evolution At CPB." *Statistica Neerlandica* 60 (2): 145–70.
- HAFNER-BURTON, EMILIE M., NEIL NARANG, AND BRIAN C. RATHBUN. 2019. "Introduction: What Is Populist Nationalism and Why Does It Matter?" *The Journal of Politics* 81 (2): 707–11.
- HAYES, JARROD. 2011. "The Democratic Peace and the New Evolution of an Old Idea." *European Journal of International Relations* 18 (4): 767–91.
- HIRSCHMAN, DANIEL. 2016. "Stylized Facts in the Social Sciences." *Sociological Science* 3: 604–26.
- HYDE, SUSAN D., AND ELIZABETH N. SAUNDERS. 2020. "Recapturing Regime Type in International Relations: Leaders, Institutions, and Agency Space." *International Organization* 1–33. First published online January 14, 2020. <https://doi.org/10.1017/S002081831900036>.
- JONES, CHARLES I., AND PAUL M. ROMER. 2010. "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital." *American Economic Journal: Macroeconomics, American Economic Association* 2 (1): 224–45.
- KALDOR, NICHOLAS. 1961. "Capital Accumulation and Economic Growth." In *The Theory of Capital*, edited by F.A. Lutz and D.C. Hague, 177–222. London: Macmillan.
- LEVECK, BRAD L., AND NEIL NARANG. 2017a. "The Democratic Peace and the Wisdom of Crowds." *International Studies Quarterly* 61 (4): 867–80.
- . 2017b. "How International Reputation Matters: Revisiting Alliance Violations in Context." *International Interactions* 43 (5): 797–821.
- LIPSON, CHARLES. 2005. *Reliable Partners: How Democracies Have Made a Separate Peace*. Princeton, NJ: Princeton University Press.
- MITCHELL, SARA MCLAUGHLIN. 2012. "Norms and the Democratic Peace." In *What Do We Know About War?*, edited by John A. Vasquez, 167–88. Lanham, MD: Rowman & Littlefield.
- NARANG, NEIL, AND BRAD L. LEVECK. 2019. "International Reputation and Alliance Portfolios: How Unreliability Affects the Structure and Composition of Alliance Treaties." *Journal of Peace Research* 56 (3): 379–94.
- REITER, DAN, AND ALLAN C. STAM. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.
- TSEBELIS, GEORGE. 2000. "Veto Players and Institutional Analysis." *Governance* 13 (4): 441–74.