**R&R Memo for manuscript #ISQ-2022-11-0579 entitled "Planes, Trains, and Armored Mobiles: Introducing a Dataset of the Global Distribution of Military Capabilities (rDMC)"**

Please find my memo indicating changes made to address editor and reviewer comments. I am grateful to the editorial team and reviewer for the opportunity to revise and resubmit the manuscript, as well as for the thoughtful and enlightening feedback they have provided. I am very happy with ways the paper has improved and hope that you all feel the same way.

The memo is organized by reviewer comment:

1. Each suggested change is provided first *verbatim and italicized*
2. My response is indented
3. Where appropriate, I have replicated old and new figures/tables and also pasted the new verbatim text from the manuscript/appendix highlighted in gray.

I hope this eases the process of re-reviewing the paper by minimizing back and forth between the memo and the revised manuscript. Thank you all, again.

**Editorial Team review**

*Editorial Team: “Make clear that this is not new data, but a database/dataset that makes currently available data more accessible to a wider community. This point is made by both reviewers.”*

This has been clarified in the abstract, introduction, and section on “scope and data generating process”. The conclusion also now clarifies that future versions of the data will diversify the sources used.

*Editorial Team: “Discuss in greater detail the reliability issues raised by Reviewer 2. Be clear that the currently available data have biases. This is also a shared concern of both reviewers.”*

A new paragraph has been added to the end of the “scope and data generating process” section to discuss this in greater detail. It is replicated in full below under Reviewer 2’s concern.

*Editorial Team: “Reviewer 2 (point 5) asks for an explicit finding that runs counter to our current understanding in the literature. We recognize that this submission is a research note and, therefore, not intended to make a larger empirical contribution per se. However, a demonstration of this argument would increase the impact of the article. It may be enough to provide a face-valid example using the database/dataset. If the author chooses not to include such a demonstration, it would be better to delete the oblique references “calling into question several findings in the literature.””*

I have added a new subsection “6.3 Validation and Application: US and China Polarity” that makes an empirical contribution by showing how rDMC can contribute to our understanding of the US-China military capabilities gap in debates about China’s rise by replicating and extending Brooks and Wohlforth (2016). Given the current length of the paper and the research note format, I felt a replication and validation of descriptive data rather than a full statistical model was more appropriate. Nonetheless, I feel this has improved the paper and increased the potential impact of the article and also provided face validity about the quality and utility of the dataset. The other sections of the paper that called into question several findings in the literature have been rephrased to show how the data could be used to test these theories/hypothesis, rather than this research note doing that testing itself.

*Editorial Team: “Figure 4 is showcasing the variety of categories included in the database/dataset (which other figures also do), but it is overwhelming. The two subplots with “Na” raise questions. Without a specific legend regarding gradation, the reader does not know how widespread any particular weapon system is.”*

Figure 4 and its accompanying in-text references have been removed. The subsection “Technological trends across time” still discusses the research on diffusion and provides an example of how rDMC could be used for research on diffusion and military effectiveness with the original example and figure of aircraft.

*Editorial Team: “Reviewer 2’s comment of changing the year regarding Figure 7 seems reasonable.”*

This change has been made. Figure 6 (which used to Figure 7) now uses 1994 instead of 2004. The accompanying text has been updated to reflect the countries that are most similar to one another in 1994, with a particular emphasis on great powers.

**Reviewer 1**

*Reviewer 1: “This is not a new source of data as all of it is drawn from the existing Military Balance. Thus, the utility of this project is the way it allows researchers to manipulate military capabilities data in a variety of ways to examine change over time and comparisons between states.”*

The abstract, introduction, and data collection section now clarify that the data comes from the existing Military Balance, and what is novel is it is now (1) organized in a structured and machine-readable format (.csv files) and (2) a new ontology of consistent weapons categories has been produced (rDMC long and rDMC wide), making the sample consistent across countries and years.

*Reviewer 1: “The manuscript is publishable as a research note in its current form with the exception of what appears to be a comment from a previous reviewer on page 7 just below table 2.”*

That comment has been removed.

**Reviewer 2**

*Reviewer 2 Important Change: “1. Ultimately, the author(s) are really introducing a new dataset, not new data. This fact is not really apparent until p6. Since we use the word “dataset” somewhat ambiguously in political science, I think, in the interests of candor, it is especially important the author(s) make it clear that all of these data are already available to scholars, just not in a very useful form. The author(s) have done us all an important service in categorizing the data and making it accessible, including by introducing their tech categories and sorting everything into them. I think this contribution should be clear in both the intro and abstract.”*

The abstract, introduction, and data collection section now clarify that the data comes from the existing Military Balance, and what is novel is that it is now (1) organized in a structured and machine-readable format (.csv files) and (2) a new ontology of consistent weapons categories has been produced (rDMC long and rDMC wide), making the sample consistent across countries and years. The words “new” and “novel” are no longer used to describe the dataset

*Reviewer 2 Important Change: “2. To justify the need for this dataset, the author(s) state “most current research uses measures only of the size of state militaries.” This description matches my impression of the literature. However, the author(s) cite almost none of this current research, so far as can tell. This omission is odd, because they are otherwise quite thorough in canvassing existing scholarship. (The author(s) do cite several scholars who have criticized/attempted to improve various measures, e.g. Lebovic (1999).) I imagine the author(s) did not want to “name names” and risk offending anyone, since the gist of their point is that relying on such coarse measures is a mistake. I believe it is vital, though, to cite some of the research they are criticizing.”*

Cites have been added to the section on existing measures of military capabilities. While there are many relevant works using aggregate measures of military spending, these cites were chosen based on (1) an interest in citing both foundational and recent work, (2) using aggregate military spending measures as a dependent variable in some cases, and independent variable in others, (3) honoring ISQ’s commitment to citing works by underrepresented groups, and (4) being high-quality and influential works. The section is reproduced below and new citations listed:

“Despite the importance of disaggregated military capabilities and a recognition of the shortcomings of aggregate measures for explaining concepts of interest, most current research uses measures only of the size of state militaries (Ward and Davis 1992; Sample 2002; Nordhaus, Oneal, and Russett 2012; Cappella Zielinski, Fordham, and Schilde 2017; Alley and Fuhrmann 2021).”

Michael D. Ward and David R. Davis, “Sizing up the Peace Dividend: Economic Growth and Military Spending in the United States, 1948–1996,” American Political Science Review 86, no. 3 (September 1992): 748–55, <https://doi.org/10.2307/1964136>.

Susan G. Sample, “The Outcomes of Military Buildups: Minor States vs. Major Powers,” Journal of Peace Research 39, no. 6 (November 1, 2002): 669–91, <https://doi.org/10.1177/0022343302039006002>.

William Nordhaus, John R. Oneal, and Bruce Russett, “The Effects of the International Security Environment on National Military Expenditures: A Multicountry Study,” International Organization 66, no. 3 (July 2012): 491–513, <https://doi.org/10.1017/S0020818312000173>.

Rosella Cappella Zielinski, Benjamin O. Fordham, and Kaija E. Schilde, “What Goes up, Must Come down? The Asymmetric Effects of Economic Growth and International Threat on Military Spending,” Journal of Peace Research 54, no. 6 (November 1, 2017): 791–805, <https://doi.org/10.1177/0022343317715301>.

Joshua Alley and Matthew Fuhrmann, “Budget Breaker? The Financial Cost of US Military Alliances,” Security Studies 30, no. 5 (October 20, 2021): 661–90, <https://doi.org/10.1080/09636412.2021.2021280>.

*Reviewer 2 Important Change: “3. As I said above, I am persuaded that these data are largely reliable and accurate, and more so than e.g. data from the IMF. Nonetheless, the author(s) too blithely dismiss objections to its completeness. Yes, “it is widely recognized as the best unclassified source...” —but there’s the rub. Obviously, many important capabilities will not be unclassified. There are biases built into this dataset, and it will probably undercount important platforms, especially among the most powerful states. It’s important to acknowledge that limitation. (And no, I’m not persuaded that checking its accuracy against government reports, when available, from countries like New Zealand proves its overall accuracy. New Zealand is not China, nor America for that matter.)”*

I thank the reviewer for this important point. Believing rDMC and the Military Balance to be the existing state of the art does not mean they are free from bias. I have added a paragraph to the “Scope and Data Generating Process” section that describes what some of these biases may be (temporal, cross-national, and platform-specific), and identifies an avenue of future research in identifying these biases with examples from autocratic GDP data and Russian losses in the war in Ukraine. The new paragraph is pasted below in full:

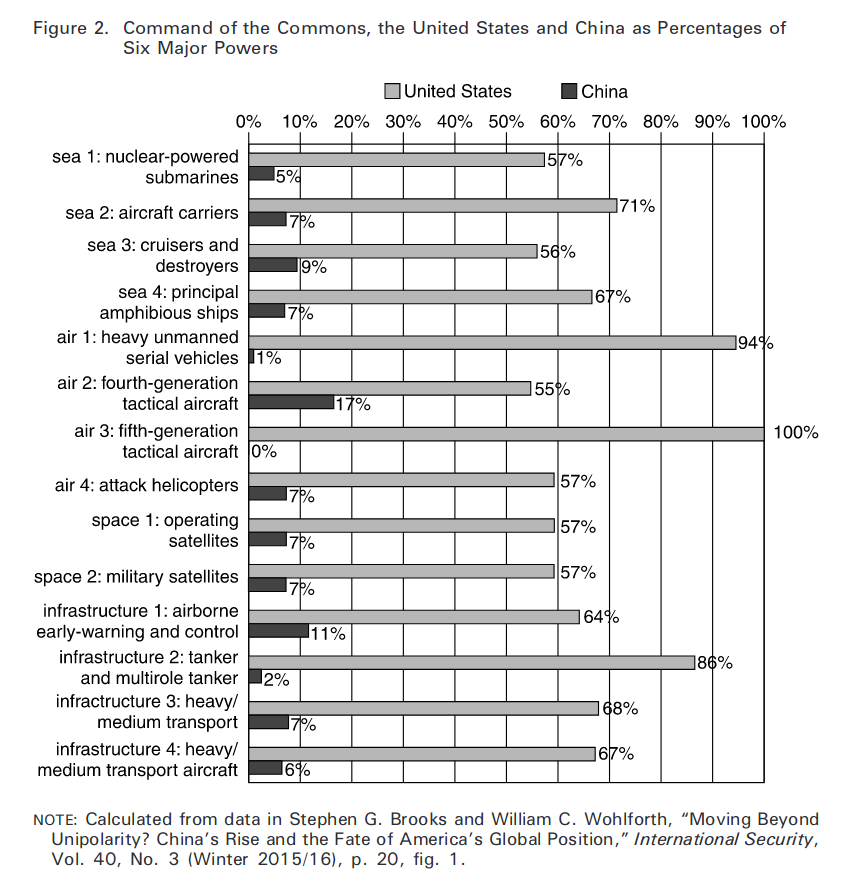
However, being the most comprehensive data on military capabilities and from a highly reputable source does not guarantee accuracy. States themselves are uncertain about the military capabilities of others, and more pronounced uncertainty likely exists in the open source data used here (Kaplow and Gartzke 2021). Earlier years may have less accurate coverage as both private and government intelligence and surveillance efforts were less sophisticated, making it harder to know the capabilities of states that were far away (Lin-Greenberg and Milonopoulos 2021). New Zealand's transparency in capability reports is an anomaly, as many countries – particularly those of interest to security scholars like Russia and China – try to deceive their foes (or even friends) by misrepresenting their capabilities by exaggerated or underplaying their strength in particular platforms (Mastro 2016; Mawdsley 2016). Weapons systems also differ in whether open source intelligence is likely to identify them correctly, as capabilities whose effectiveness relies on secrecy and deception may not be known to the rest of the world until they are used (Green and Long 2020). It is important that scholars acknowledge this limitation and, where possible, take steps to address it. For example, biases in GDP estimates have been identified using night-time-light data gathered from satellites, allowing scholars to quantify the extent to which autocratic regimes exaggerate their economic growth (Martinez 2022). In the future, similar measures could be applied to rDMC by, as has been done with Russia in the invasion of Ukraine, combining reports about military use and loss during combat to pre-war inventory estimates (Zhukov, n.d.).

*Reviewer 2 Important Change: “4. I think the dataset’s modularity will be especially valuable to researchers in different areas of security studies. It should be mentioned in the abstract.”*

Modularity is now mentioned in the abstract.

*Reviewer 2 Important Change: “5. The author(s) hint, and come close to saying, that the data call into question several findings in the literature. They dance around this especially on p12. I think they should explicitly identify a finding, then carefully show how their new dataset should change our minds about it. Of all the revisions I’m suggesting, this will be the most time-intensive.”*

I have added a new subsection “6.3 Validation and Application: US and China Polarity” that makes an empirical contribution by showing how rDMC can contribute to our understanding of the US-China military capabilities gap in debates about China’s rise by replicating and extending Brooks and Wohlforth (2016). For reference, here is the figure in Brooks and Wohlforth that is replicated:



Given the current structure of the paper and the research note format and word limit, I felt a replication and validation of descriptive data rather than a full statistical model was more appropriate. I feel this has improved the paper and increased the potential impact of the article and also provided face validity about the quality and utility of the dataset.

The other sections of the paper that called into question several findings in the literature have been rephrased to show how the data could be used to test these theories/hypothesis, rather than this research note doing that testing itself.

The author is currently working on other papers that use rDMC to change our minds about existing finding in more robust and theoretically-driven way. The goal is this piece is to introduce the dataset, make it publicly available as soon as possible, and identify some ways in which it can be used by other scholars. My hope is that the current structure accomplishes those ends within the smaller word limit that exists for a research note.

*Reviewer 2 Other Comments: “1. I think the title is too cute. It’s clever, but it also seems unserious. I’d leave it up to the author(s), but I think a straightforward title like “Introducing a Dataset of Global, Disaggregated Military Capabilities (rDMC)” would be better.”*

The reviewer’s point is well taken, but I have decided to keep the title as is since it is known colloquially to some relevant scholars as the “Planes, Trains, and Armored Mobiles” dataset.

*Reviewer 2 Other Comments: “2. The author(s) should cite the growing literature on multi-domain warfare, e.g. Perkins (2017). After all, the very premise of such warfare is that different domains are not interchangeable, which is the justification for this dataset.”*

The section on “The case for disaggregating defense” now cites some of the research on multi-domain warfare. A complete excerpt of the new text is below:

More recently, the increasing interest in multi-domain warfare demonstrates that among practitioners and academics alike, there is a recognition that the various tools of warfare - and the domains in which they operate - serve different ends and come with distinct advantages and disadvantages(Tan 2017; Perkins 2017; Kreps and Schneider 2019; Lindsay and Gartzke 2019)

*Reviewer 2 Other Comments: “3. While it is obvious from figure 1, I think it would also be worthwhile to say, on p5, that the data has almost 100% coverage of the great powers.”*

Clarification added in the sentence with missingness percentages.

*Reviewer 2 Other Comments: “4. On p7, the manuscript seems to include a passage from a previous review. I assume the author(s) forgot to delete this.”*

Passage omitted.

*Reviewer 2 Other Comments: “5. Footnote 7 says that standardizing unit-level names will take some time. Why can’t this be automated?”*

As this is the main improvement I want to do for rDMC 2.0, I am currently trying to figure out ways to standardize the names based on existing military sources (NATO names) and existing public sources (wikipedia has pages for many of the units and each page has a standardized unique numeric key). Part of the impetus for including the main battle tanks figure in this manuscript was as a test case in standardizing the unit names, which I was unable to automate and had to do by hand for that figure. Having also now done that for 4th and 5th generation fighter aircraft in the new replication/extension, I am even more confident standardizing the names will take some time (and Ras). A few things have made the automation process challenging, although I still think it will be possible and plan to do it:

1. Names differ across countries – the Soviet MiG-21 has the NATO reporting name Fishbed. When the Soviet Union sold it to China, China renamed it the Chengdu J-7 which NATO calls the Fishcan. Even if we assume China made some changes so the Chengdo-7 is similar, but not identical to/synonymous with the Mig-21, one part of standardizing the names is deciding whether to standardize by giving every unit its NATO reporting name, producing a new dataframe of “equivalent to” that identifies all the other names for the Fishbed-21, or coming up with a new standardized name since NATO does not have names for all Soviet aircraft (nonetheless all equipment in the world).

2. Fuzzy string matching – the Military Balance pdf’s cannot be scraped and were instead processed by hand. Coders were instructed to enter the unit names exactly as they were written in the PDF since not all were experts in military capabilities. As a result, there are entries for “F35”, “F-35”, “F 35”, “F-35 Lightning II”, etc. Many, but not all, of these differences concern delimiters (spaces, hyphens, etc) that can partially be matched automatically, but the use of full names or cases where the NATO name or full name are in parentheses complicates it.

3. Variants – some times, data is provided separately for the F-35A, F-35B, and F-35C. Other times, data is just provided for F-35 (no variant specified). Since the delimiters here are applied inconsistently, it’s hard to know how to draw those out. It’s also unclear where the differences across the variants are significant enough to be of interest to scholars and how that should be handled systematically.

4. Combined variants/units – Related to the point above, other times the data is aggregated across some (but not all) variants of a same unit (F-35A/B) but yet other times counts are provided for an aggregation of multiple units (in 2011, Armenia had 8 T-54/T-55). Since other country-years have counts for the T-54 and T-55 separately, it’s hard to know how to deal with the aggregated counts.

*Reviewer 2 Other Comments: “6. Footnote 10 is a testament to the care the author(s) and their coders have taken in constructing the dataset.”*

Thank you for the kind words.

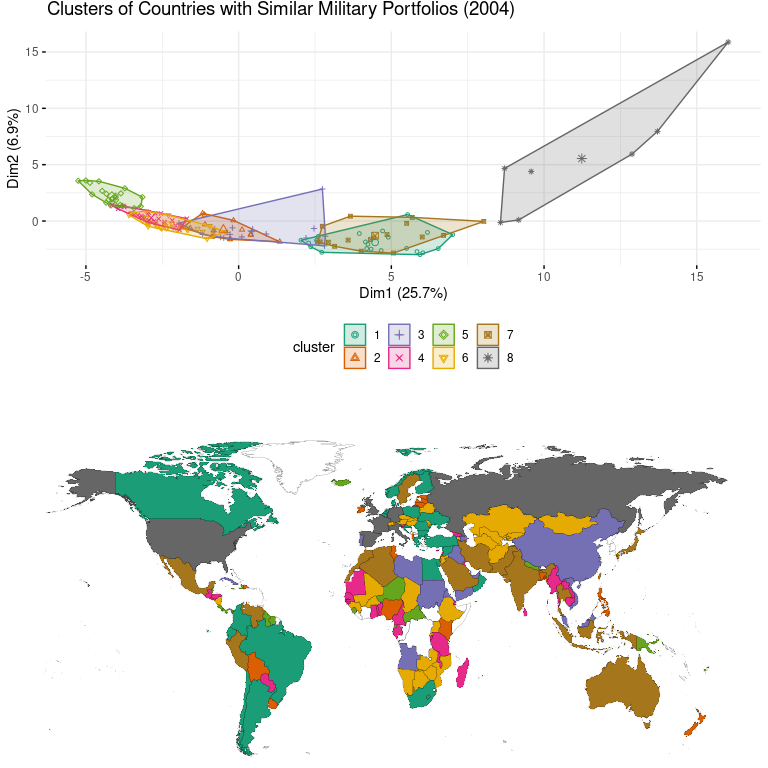
*Reviewer 2 Other Comments: “7. I do not find figure 4 helpful. But perhaps other reviewers might.”*

Figure 4 and its accompanying in-text references have been removed. The subsection “Technological trends across time” still discusses the research on diffusion and provides an example of how rDMC could be used for research on diffusion and military effectiveness with the original example and figure of aircraft.

*Reviewer 2 Other Comments: “8. Figure 7 looks at a single year, 2004, to show how states do/don’t cluster. I assume the author(s) chose 2004 because it’s a decade before their data end. But it’s also a year when the US is at war alongside a variety of countries. Presumably that affects the composition of their arsenals and their similarity. Would it be possible to look at a year where we can get a better feel for peacetime arming? 1994 would allow the author(s) to stay with their same ten-year intervals. I think, when constructing this figure, “boring is better” when choosing a year to depict.”*

The new figure and accompanying text now depict 1994. Both are reproduced below:

There are eight distinct clusters of countries that share significant commonality in the military capabilities they possess. Some great powers are similar to one another (the US, Russia, France, and UK), but others appear more distinct (China and India). In some cases the US is similar to allies like Germany and Italy, but not with others like Poland and Canada. There is similarity across states in some geographic regions like Central Asia, but significant dissimilarity across states in other regions like the Pacific Rim.



*Reviewer 2 Other Comments: “9. Figure 7 seems to suggest that the Great Powers are playing one game, and everyone else is playing another. Contrary to the implication at the top of p16, I’m therefore not surprised that the US looks more like France than Poland, since the former still tries to play in the big leagues, while the latter does not.”*

The countries listed as being similar vs different from the US has been updated given the change in year displayed. Interestingly, China is no longer in the same cluster as the US and Russia, now instead possessing similar weapons (albeit, in different quantities and more advanced variants) as Cuba, Iraq, North Korea, Libya, and Yemen. There may be something interesting to be said (later) about autocratic inclination towards certain weapons, maybe given arms sales patterns, coup risk, recent history of civil conflict, industrialization patterns, etc. Food for thought!