

Editors: **Diane Labrosse** and **Seth Offenbach**

Web and Production Editor: **George Fujii**

Michael Beckley. “The Power of Nations: Measuring What Matters.” *International Security* 43:2 (Fall 2018): 7-44. DOI: https://doi.org/10.1162/isec_a_00328.

Published by ISSF on **19 July 2019**

<https://issforum.org/to/iar122>

Review by **Robert J. Reardon**, North Carolina State University

In this article, Michael Beckley makes an important contribution to how scholars measure state power, arguing that *net* rather than *gross* indicators of a state’s military and economic resources better capture a state’s capabilities. The question of how to measure state power is central to both theoretical and policy debates. In theoretical terms, power lies at the center of any effort to understand how states can influence one another in the international arena. Arguably no outcome in international affairs can be properly understood without attention to the relative power of the actors involved. Likewise, some of the most important policy debates hinge on perceptions of relative power, including those over the rise of China and the strategic posture the United States ought to adopt as a result. Advocates of U.S. retrenchment tend to argue that China is rapidly closing the gap in capabilities with the United States and will soon eclipse U.S. capabilities.¹ Beckley and others who advocate a strategy of deep engagement argue that the United States still enjoys, and will likely continue to enjoy for the foreseeable future, a sizeable advantage over China. It is an ambitious objective to take on such a core question, and Beckley should be applauded for challenging widely established approaches to using power as a variable and suggesting ways to improve those approaches that are both promising and parsimonious.

Given the centrality of the concept of power to the study of international relations, the degree to which scholars have failed to arrive a consensus on how should be defined and measured is remarkable. Power, as Beckley notes, “is typically defined as the ability of a country to shape world politics in line with its interests” (8). However, because this understanding of power depends on a state’s conception of its interests, it cannot be measured directly. If we seek to determine a state’s power based on the outcomes it achieves in a given dispute, we must first know the relative preferences of the parties involved, and even then our measure of power would be restricted to that one, narrow-issue area.

¹ See, for example, Joshua R. Itzkowitz Shiffrin and Michael Beckley, “Debating China’s Rise and U.S. Decline,” *International Security* 37:3 (Winter 2012/13): 172-181, DOI: https://doi.org/10.1162/ISEC_c_00111.

Instead of looking at outcomes, most scholars measure power by assessing the aggregate resources a state possesses—particularly economic and military resources—based on the assumption that “countries with more wealth and more military assets at their disposal tend to get their way more often than countries with fewer of these resources” (8). Beckley supports such a resource-based approach for the compelling reasons that it allows for a single aggregate indicator to represent state power across issue areas and provides a way to measure a state’s relative power *ex ante*, thus enabling scholars to study the effect of power on dispute outcomes.

Beckley parts company with other scholars, however, by arguing—with considerable justification—that the dominant use of gross indicators such as GDP as proxies for military power is fundamentally flawed. This argument builds on the work of Klaus Knorr and others, and rests primarily on two important considerations that feature prominently in Beckley’s earlier work on the relative rise of China.² First, gross measures, such as GDP or military expenditures, can mislead by failing to consider variation in efficiency across states. Differences in worker productivity, in the existing availability of certain technologies and research and development (R&D) efforts, in established research institutions, and in related factors can produce considerable variation in the efficiency with which any amount of spending on a state’s military can be translated into actual capabilities that confer some advantage on the battlefield.

Second, gross measures of power put disproportionate weight on a state’s size. Gross measures of income correlate with a state’s population, making many poor but large states appear more powerful than smaller yet wealthier ones. As Beckley concedes, there is a logic to this, since larger countries can aggregate greater resources than smaller ones. Large countries like China and India can assemble far greater military capabilities than small states like Luxembourg, no matter how much greater Luxembourg’s per capita wealth might be. Scale matters. However, attention only or disproportionately to a country’s size obscures how differences in per capita resources influence state capabilities. The larger the state, the more it must spend to maintain a basic level of welfare and security for its citizens. Attention only to size distorts the capabilities of very large but less economically and technologically developed states, which are typically forced to devote the lion’s share of their resources to maintaining even low levels of welfare, leaving relatively few resources to devote to building up military capabilities. Some states like North Korea may be more willing to squeeze the population in order to devote a greater share of the nation’s resources to their military, but there are limits to how much this can be done before domestic stability is put at an unacceptable level of risk.

Beckley offers two suggestions for overcoming these problems with using gross indicators of power. The first is to devise a careful ledger of state resources and costs to produce a more precise net indicator of capabilities. As Beckley acknowledges, however, such an endeavor would require voluminous data, and there are no existing data sets available that would provide the necessary data for such an accounting across a sufficient set of states and time periods. Scholars can and should work toward producing such data sets. Beckley’s second suggestion, that scholars rely on the relatively simple indicator of Gross Domestic Product (GDP) x GDP per capita as a proxy for net capabilities, offers a way to roughly adjust for welfare and security costs using data that is easily and readily available. Most IR scholars have simply used rough proxies such as GDP to represent state power, or have relied on data sets like Composite Index of National Capability (CINC) that combine a number of variables—including GDP, population, and state military budget—to take a rough measure of a

² For example, Beckley cites Klaus Knorr, *The War Potential of Nations* (Princeton: Princeton University Press, 1956); see Michael Beckley, “China’s Century? Why America’s Edge Will Endure,” *International Security* 36:3 (Winter 2011/12): 41-78, DOI: https://doi.org/10.1162/ISEC_a_00066.

state's aggregate capabilities. The logic to using these measures is that wealth is fungible, and states can convert economic and other resources into military capabilities, making proxies that measure overall gross resources an acceptable way to compare relative power across states. GDP x GDP per capita adjusts the gross indicator GDP according to per capita income, thus reducing the impact of a country's overall size without discounting it entirely.

As a rough test of whether gross and net indicators better measure state power, Beckley compares GDP x GDP per capita to GDP and CINC data in three ways. First, he examines several historical cases of great-power rivalry in which there was considerable disparity between gross and net measures of resources. Second, he conducts a large-*n* analysis of wars and disputes to see which indicators most often predict outcomes. Third, he examines past studies that used GDP and CINC data to see whether replacing those measures with GDP x GDP per capita improves goodness-of-fit.

The case studies are by far the most revealing. As Beckley points out, GDP and the CINC data set lead to nonsensical conclusions about the relative power in several critical cases, especially those involving China and Russia. According to those measures, "China looked like a superpower in the nineteenth and early twentieth centuries," and surpassed Britain, which, in reality, had humiliated China in the First Opium War (1839-1842) and Second Opium War (1856-1860). GDP x GDP per capita is a clearly superior measure, giving Britain three times China's power in 1870—a finding which is far more consistent with outcomes. The net indicator similarly outperforms GDP and CINC when used to examine pre-First World War Russia and Germany, late-nineteenth and early-to-mid twentieth century Japan and China, and the United States and the USSR during the Cold War.

The other tests, however—while supporting Beckley's overall claims—are less impressive in terms of the magnitude of difference between using the different indicators. Here Beckley may be overstating his case when he criticizes gross indicators as not simply inferior to net ones, but as "logically unsound," "empirically unreliable," and "severely mischaracterizing the balance of power in numerous cases, including in some of the most consequential geopolitical events in modern history" (9). The paper's large-*n* analysis based on Militarized Interstate Disputes (MIDs) data, for example, shows an important yet modest improvement in predictive power by using net indicators in place of gross ones. GDP x GDP per capita predicted 70% of the dispute outcomes from the MIDs dataset, which is certainly impressive. However, GDP and the CINC dataset were both able to predict 64% of these dispute outcomes (38). That represents an important gain from using net indicators, but not the sort of drastic improvement one might expect given Beckley's logical critique of gross indicators. Indeed, it raises the question of why GDP and CINC perform so well given the flaws in using them that Beckley points out.

A likely reason is that the difference between net and gross indicators becomes especially salient when dealing with large but relatively poor states such as China and Russia (and, as Beckley notes, disputes involving those states represent a large share of the overall cases where net indicators outperform gross ones). This would still be a critical finding—Russia and China are important states, and any improvement in our ability to measure the power of those two states would be an important achievement. However, it is important to note whether gross indicators perform adequately outside of cases involving states that are particularly populous, or where there is a large difference between where those states rank globally in terms of aggregate GDP and GDP per capita. If net indicators only outperform gross ones in a subset of cases, this would require a far less drastic reexamination of past studies that relied on gross indicators of power.

Finally, Beckley's analysis could have gone further by considering how the use of net indicators might shed light on power transitions. Even though, in the paper's case studies, Beckley examines long-term rivalries between great powers, he focuses mostly on how relative differences in net aggregate resources influence outcomes in any given dispute, rather than considering what might lead these relative differences in power to shift over time. The U.S.-USSR case, where Beckley notes how Soviet welfare demands and security costs such as aid to Eastern bloc allies outpaced production and gains in technology and productivity, is an exception. What is notable here is that net indicators suggest a long-term decline in the Soviet Union's position even as gross indicators suggest that the USSR was reaching increasing heights of relative power. As Beckley points out, this is consistent with some U.S. estimates from the era that the Soviet Union had already eclipsed the United States in many key areas of military and industrial capacity. Importantly, the problem was not simply one of taking account of welfare and other costs. Efficiency in this case is of paramount significance. The USSR, for example, "spent 2 to 3 percent of its GDP on military R&D, but lagged a generation or more behind the United States in fifteen of the twenty most critical military technologies" (36). The question is not just one of having sufficient "disposable" resources to pour into military capabilities, but also one of having the wherewithal to efficiently translate those resources into improved capabilities relative to an adversary. The United States wasn't just more powerful in terms of a snapshot of capabilities at a given moment. Many of the same factors Beckley points to in the article also helped the United States outpace the Soviets in terms of growth in power over time. A better understanding of this dynamic would be greatly useful, and worthy of further research.

Robert J. Reardon is an assistant professor of political science in the School of Public and International Affairs at North Carolina State University. He is co-principal investigator with the Consortium for Nonproliferation Enabling Capabilities (CNEC), a NNSA-sponsored research and education consortium on nuclear nonproliferation. Reardon's teaching and research focuses on international security, nuclear weapons strategy and proliferation, and U.S. grand strategy. From 2012-2014, he was a fellow at Harvard's Belfer Center for Science and International Affairs with a joint appointment in the International Security Program and the Project on Managing the Atom, and from 2010-2011 was a Stanton Postdoctoral Fellow in nuclear security at the RAND Corporation. He received his Ph.D. in political science from MIT, and his BA from Columbia University.

©2019 The Authors | [Creative Commons Attribution-NonCommercial-NoDerivs 3.0 United States License](#)