International Studies Quarterly (2020) **0**, 1–14

Bread Before Guns or Butter: Introducing Surplus Domestic Product (SDP)

RESEARCH NOTE

Therese Anders
Hertie School, SCRIPTS Cluster of Excellence

Christopher J. Fariss
University of Michigan

and

Jonathan N. Markowitz
University of Southern California

Scholars systematically mismeasure power resources and military burdens by using gross domestic product (GDP) as a proxy for the income states can devote to arming. The core problem is that GDP confounds two conceptually distinct forms of income into one additive indicator. Subsistence income represents resources needed to provide the "bread" necessary to cover the basic subsistence needs of the population. Surplus income represents the remaining resources that could be allocated to "guns" or "butter." Our new measure of surplus domestic product (SDP) corrects for this measurement error by decomposing subsistence income and surplus income from total GDP. Validation exercises demonstrate that SDP outperforms GDP at measuring the distribution of power resources. Though theoretically we expect states' decisions to arm are influenced by the distribution of power; empirical models using GDP find mixed support for this expectation. Strikingly, using SDP reveals strong support for this proposition.

> The physical product of hundreds of millions of peasants may dwarf that of five million factory workers, but since most of it is immediately consumed, it is far less likely to lead to surplus wealth or decisive military striking power.
>
> — Paul Kennedy, *The Rise and Fall of Great Powers*

International relations scholarship systematically mismeasures both power resources and military burdens because the operationalization of these variables depends on gross domestic product (GDP) as an indicator of the income states can devote to arming and projecting power. The core problem is that GDP confounds two conceptually distinct forms of income into one aggregate indicator. Subsistence resources are the income necessary to provide the minimal amount of "bread" that the population needs to survive. Surplus resources are the remaining economic income that can be invested in "guns" or "butter" (Garfinkel and Syropoulos 2019). As a consequence of conflating these two types of income into a single indicator, existing inferences about the distribution of power resources, states' capacity to build power projection capabilities, and the costs of arming are biased. In particular, the misuse of GDP as a proxy measure of power resources systematically overestimates power resources of low-income states with large populations and underestimates the rate at which these resources increase when low-income countries begin to experience economic growth. We address this bias by developing a new measure called surplus domestic product (SDP). SDP is created by decomposing GDP into surplus income and subsistence income.

This article yields three contributions for scholars of international relations and political science more broadly. First, we introduce surplus income (SDP) and subsistence income, which are both part of total income (measured in GDP). We demonstrate that, once we account for both types of income, SDP is a more conceptually appropriate measure of the power resources available for states to arm and project military force abroad. For illustration, scholars mismeasure military burdens by using military expenditures as share of GDP. When SDP is used in the denominator instead, we demonstrate that, for most of history, nearly all countries endured much higher military burdens than previously realized. Failing to take into account that only surplus resources

Therese Anders is a postdoctoral research fellow at the Hertie School Data Science Lab and data scientist for the SCRIPTS Cluster of Excellence "Contestations of the liberal script."

Christopher J. Fariss is an assistant professor in the Department of Political Science and a faculty associate in Center for Political Studies, both at the University of Michigan. He is a faculty associate at the Security and Political Economy Lab at the University of Southern California. He received his PhD in 2013 for the University of California San Diego.

Jonathan N. Markowitz is an assistant professor in the International Relations and Political Science Department at the University of Southern California, where he is also a cofounder and co-PI of both the Security and Political Economy Lab and East Grand Strategy Program.

Author's note: We would like to thank the audiences at the 2018 annual meetings of the International Studies Association and the American Political Science Association, and in particular Alex Braithwaite and David Sacko, for their helpful comments and suggestions. Special thanks go to Benjamin Graham and Andrew Coe for extremely helpful comments, as well as to Miriam Barnum and the University of Southern California Security and Political Economy Lab for excellent research assistance. Authors are listed in alphabetical order, and all authors made an equal intellectual contribution. Jonathan Markowitz secured funding for the project.

Anders, Therese et al. (2020) Bread Before Guns or Butter: Introducing Surplus Domestic Product (SDP). *International Studies Quarterly*, doi: 10.1093/isq/sqaa013

can sustainably be invested in the military concealed just how rapid and steep the fall in military burdens has been after the Cold War. This drop is especially pronounced for developing states around the world, particularly in Asia. Our results suggest that scholars underestimated the size of the peace dividend associated with the end of the Cold War and the gains from entering into hierarchical security relationships (e.g., Lake 2009). Our recommendation of using SDP while also accounting for subsistence income as a separate variable—instead of total GDP—is potentially critical to scholars who currently use GDP as a proxy for the capacity of states to devote resources to nonmilitary purposes, for instance, in work evaluating states' decisions to invest in education, healthcare, or other quasi-public goods (e.g., Tanzi and Schuknecht 2000; Lake and Baum 2001).

Second, in the process of building comprehensive data on surplus and subsistence income we provide new data for some of the most widely used variables in political science and economics, with coverage from 1800 to 2018.[1] In particular, we improve existing cross-sectional and temporal data coverage for both GDP and population (the components of our new SDP indicator). Currently, cross-national data on military expenditure as a percentage of GDP go back to 1950 (Nordhaus, Oneal, and Russett 2012). Our new data allow us to extend the data coverage for military expenditure as a percentage of economic resources (SDP or GDP) back to 1816 for most of the world's states.

Third, we use SDP to improve the measurement of relative power resources between states, allowing us to provide more accurate identification of the most powerful and potentially threatening states in the world each year. Based on existing scholarship, which relies on GDP as a proxy for power resources, China is ranked as the world's most powerful country in the early and mid-nineteenth century. However, until the 1990s, nearly all of China's income was used to sustain its large, impoverished population and little surplus remained to invest in guns or butter. Using SDP to measure relative power resources leads to a markedly different set of countries topping the rankings of powerful states. In the early nineteenth century, our estimates place the United Kingdom rather than China in the top spot, which comports much better with arming and power-projection behaviors of countries during this time period. Decomposing GDP into surplus and subsistence income thus provides new insights for scholars working on a broad range of topics related to the distribution of power such as arming, alliances, power transition, peace, and great-power politics.

While financial resources are not the only dimension of a state's power resources, they represent a particularly important, if not the most critical, dimension (Cappella-Zielinski 2016; Beckley 2018; Norloff and Wohlforth 2019). The importance of financial resources led prior scholarship to use GDP as the primary measure of states' power resources. Our claim is not that SDP measures every dimension of states' power resources, or that other dimensions of power should not be considered, but only that SDP outperforms GDP as an approximation for the income a state can devote to pursuing various objectives, which may include not only arming, but also investment in industrialization. As we demonstrate in this article, SDP is more closely related than GDP to alternative measures of power resources, such as industrial capacity-related components of the Composite Index of National Capabilities (CINC). We also develop and examine solely population-based measures of power resources. We show that GDP measures are more closely related to population measures than SDP. In fact, in the early to mid-nineteenth century, variation in GDP is almost entirely explained by variation in population—a reflection of the Malthusian constraint faced by most countries at the time. Thus, SDP captures a concept of power resources distinct from GDP or population.

Our new measure of SDP outperforms GDP in three validation exercises. First, SDP fares better than GDP when compared to alternative indicators that approximate states' relative power resources.[2] Second, SDP more accurately ranks states with the greatest power resources. We modify a recently developed operational strategy to incorporate SDP into measuring the level of potential threat in states' geopolitical environments (Markowitz and Fariss 2018). Specifically, we use SDP to measure relative power resources between pairs of states and weight these ratios by distance and preference compatibility. Third, we show that weighted power-resource ratios between states based on SDP produce more valid rankings of countries each state finds most threatening, as compared to the same approach using GDP.

For our primary empirical application, we generate an aggregate country-year measure of the potential threat each state faces in its strategic environment. We illuminate how the potential threat states face influences the degree to which they invest scarce resources into arming and power-projection capabilities. We demonstrate that when SDP is incorporated into this model, it outperforms the same model using GDP in predicting military investments. Strikingly, our model reveals that key variables suggested by existing theories of arming, such as relative power, geographic proximity, and preference compatibility cannot explain the degree to which states arm when power resources are measured using GDP. However, these same variables can explain the degree to which states arm, when power resources are measured using SDP and military burdens are measured using military expenditures as share of SDP. It is only the misuse of GDP as a measure of available power resources that makes existing theories of arming appear empirically unsupported. Our results are robust to using naval tonnage relative to SDP as an alternative measure of arms levels.

Decomposing GDP into surplus income and subsistence income allows us to examine patterns of SDP historically. We show that in the nineteenth and early twentieth centuries, military burdens—the percentage of income devoted to arming—were, on average, higher than suggested by existing research. This difference exists because, until recently, most states were able to generate only small amounts of economic surplus. States that generated surplus spent a large proportion of it on arming. As a result, many states in this earlier period had military burdens as high as 25 to 50 percent of SDP. To put this in perspective, even during the Cold War, the United States spent only about 10 percent of its SDP on the military.

Finally, SDP reveals that military burdens have fallen faster than previously realized. Newly industrialized states such as China are seeing large gains in economic surplus, which they could invest in arming. Yet, when measured as a percentage of SDP, the military burdens of these states came tumbling down over the last several decades. Today, most governments are prioritizing butter over guns, and, as a result, military burdens as a percentage of SDP are far lower than in the past. However, for the least developed states, military burdens are still much higher than previously realized.

[1] See Coyle (2014) for a discussion of the history of the statistic known as GDP.

[2] Following Beckley (2018), we define a state's power resources as the pool of resources a state can potentially invest in generating influence.

Subsistence and Surplus

Scholars long recognized the close relationship between states' income and the resources available for military use. For example, Sandler and Hartley argue that, "as GDP rises, a nation has both more resources to protect and greater means to provide protection" (Sandler and Hartley 1995, 60). However, scholars acknowledge that states' ability to generate military power depends not only on the size and sustainability of its resource base, but also on the degree to which the state is constrained from extracting and mobilizing those resources (Milward 1977; Lamborn 1983; Sandler and Hartley 1995). Our argument and measure of SDP build on these insights. While most previous research highlights how domestic political factors constrain the amount of resources a state can extract, we focus on how biological factors limit the amount of resources that are potentially available for extraction. Political constraints are important, but we cannot accurately estimate their impact without first creating a measure of the resources that could, given the political will, be extracted by the state.

We argue that, when estimating the amount of resources that are potentially available for extraction and arming, scholars should account for a state's surplus and subsistence income separately, rather than adding them together. Surplus income (SDP) is calculated by removing from GDP the resources the population must consume to survive. Biology determines the number of calories required for survival, and this lower caloric bound is largely stable across time and space. If citizens do not survive, they cannot use their labor to produce income.

It is possible for states to extract subsistence income from their population, and states sometimes do, particularly in times of crisis. However, this level of predation results in the population growing sicker and weaker—decreasing their ability to produce income and consequently reducing the resources available for the state to extract. Thus, the subsistence needs of the population place an upper bound to the amount of income states can sustainably extract. Resources remaining after subsistence needs of the population are met represent surplus income that can potentially be extracted by the government.

How do we estimate subsistence needs? The World Bank monitors health and wellbeing associated with extreme poverty at several thresholds. We argue that a $3 per-day threshold, which the World Bank calls "close to extreme poverty," is a conservative estimate of the subsistence resources required per capita.[3] It is conservative because even at $3 per day, people tend to be chronically malnourished. As a result, they are more likely to succumb to disease and generate less surplus income. The World Bank estimates that in low-income countries (defined as countries having a per-capita gross national income of $1,095 or just under $3 per day), 27 percent of the population were undernourished in 2015 (World Bank 2017). Moving from a $3 to a $1.90 per-day threshold is associated with chronic malnourishment—causing approximately 10 to 40 percent of children under the age of five to be underweight (Ezzati 2004, pp. 1949 and 1985).

We use the $3 per-day threshold to calculate the SDP available for the state to extract and spend on public or private goods—in particular, arming and power-projection capabilities—and the remaining subsistence income. While it is possible for states to extract subsistence income via taxation, we argue that the costs of doing so are very high and that it cannot be done sustainably. Taxing the population to below the subsistence threshold undermines economic productivity in even the very near future. Biological constraints are not the only limits states face in taxing their population (e.g., Milward 1977; Lamborn 1983). Historically, most states lacked the political will or capacity to extract the entire surplus. SDP estimates the upper bound on the resources a state *can* sustainably extract if it has the capacity and political will to do so.

To compute SDP for each state i in year t, we first calculate the *minimum subsistence value* v_{it}, which is the level of income necessary to sustain the state's population, and then use this value to determine how much surplus income remains. We let $v_{it} = [(365 \times \tau) \times population_{it}]$, where τ is the per-day, per-person subsistence threshold. Based on our discussion above, we use a subsistence threshold τ of $3 per day per person. We also assess the sensitivity of our results to per-day subsistence thresholds at $3, $2, $1, and $0 (standard GDP). If $GDP_{it} > v_{it}$, then $SDP_{it} = GDP_{it} - v_{it}$ and $subsistence_{it} = v_{it}$. If $GDP_{it} \leq v_{it}$, then $SDP_{it} = 0$ and $subsistence_{it} = GDP_{it}$.

For a state to have surplus income, it must generate enough subsistence income to meet the needs of the population. If the state's income does not exceed this *minimum surplus value*, it has a SDP of zero and only has subsistence income to work with. We thus decompose GDP into surplus income (SDP) and subsistence income (see Supplementary Appendix Section A for further details about this formalization). We account for both income values in the measurements and regression models we develop below.

Converging and Diverging Trends in the International System: GDP versus SDP

Having decomposed GDP into surplus income (SDP) and subsistence income, we ask the following: *Is SDP a more valid measure of power resources than GDP?* As a first concurrent validity test, we compare temporal trends in states' shares of global economic power for GDP versus SDP.[4] We demonstrate that SDP produces a better representation of historical trends and a more valid list of the global top ten powers than GDP for the past two hundred years. Second, we demonstrate the benefit of using surplus as a measure of power resources by showing that SDP correlates more strongly with alternative indicators of power resources than GDP.

Figure 1 displays the evolution of states' shares of global power resources for six countries, for SDP and GDP. Using SDP, wealthier states such as the United States and the United Kingdom are estimated to have a much higher share of global power resources than when using GDP.[5] The opposite is true for most poorer states. For countries such as China and South Korea, measuring power resources via GDP overestimates relative power resources prior to 1980 and underestimates the rate of their rise since. The effect of economic development is illustrated by the case of South

[3] The thresholds are calculated in constant 2011 purchasing power parity dollars. The World Bank's poverty threshold underwent adjustments over time—the threshold of $1.90 today is equivalent to $1.08 in 1993 and $1.00 in 1985 (Ezzati 2004). The statistic on malnourishment referenced here was originally calculated based on the $1.08 threshold and adjusted to reflect the World Bank's new poverty threshold of $1.90 (Ferreira 2015).

[4] Supplementary Appendices B and C present formal validity test descriptions.

[5] Figure 7 in the supplementary appendix provides scatterplots of SDP versus GDP for select years. Figure 8 in the supplementary appendix illustrates the relationship between SDP and per-capita GDP and demonstrates that, while economic development increases the correlation between SDP and per-capita GDP, individual countries vary considerably regarding the strength of this relationship.

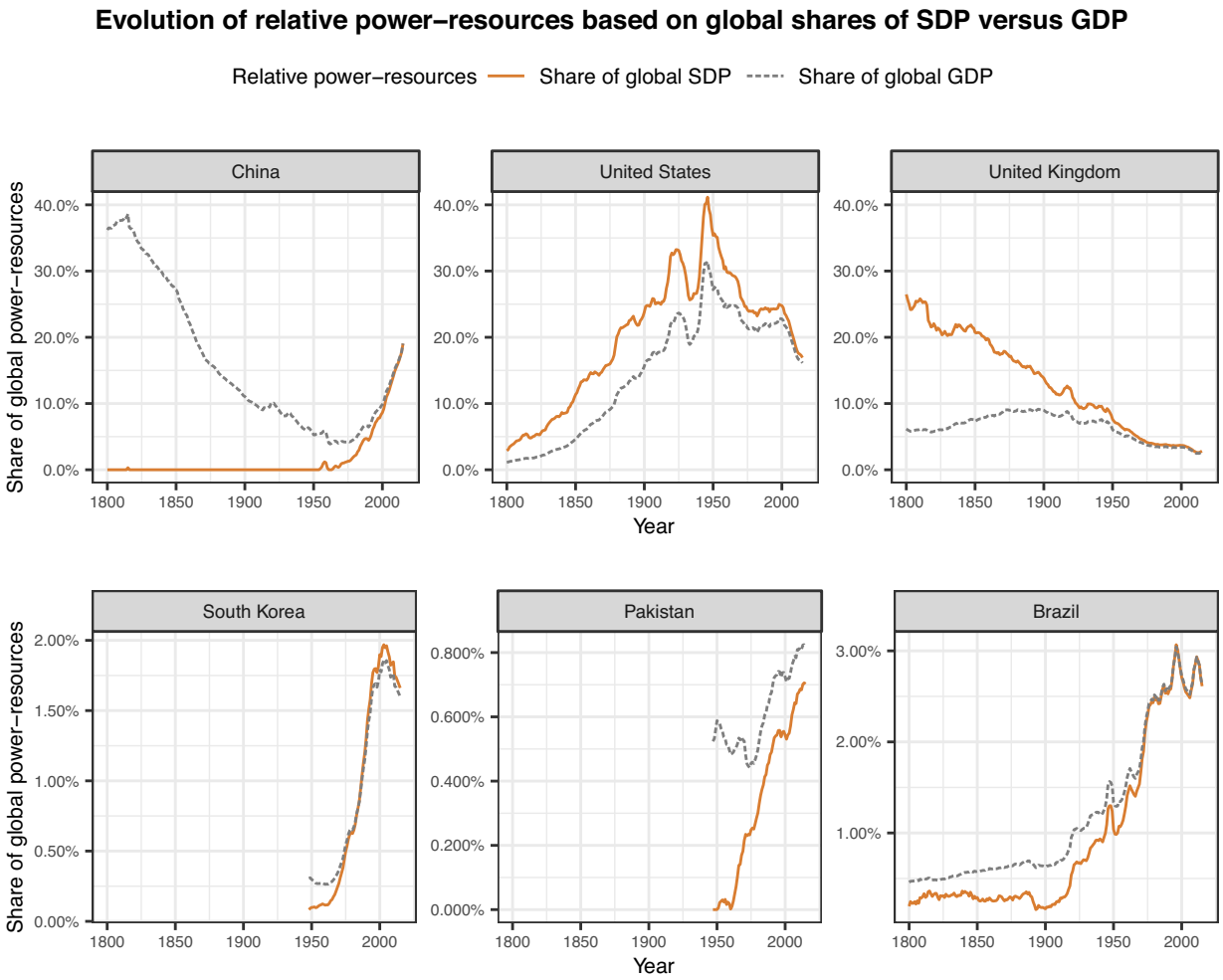

Figure 1. Evolution of economic power-resources based on a state's share of global SDP versus GDP
Note: SDP is based on a $3 per-diem subsistence level.

Korea. Prior to its burst of economic growth, the county's share of global GDP was substantially higher than its share of global SDP. Beginning in 1985, this trend reversed, and South Korea's share of global SDP started to become larger than its share of global GDP.

Figure 2 displays the top ten powers based on their average share of global SDP or GDP for four time periods. The difference is most striking for China. In the nineteenth century, the economic income produced by hundreds of millions of peasant farmers dwarfed the production of other states. Based on its share of global GDP, China would be considered the most powerful country in the period from 1816 to 1869. However, because most of this income was immediately consumed for subsistence needs, China had little surplus wealth to invest in arming or power projection. SDP takes this into account and does not place China in the top ten powers in the early nineteenth century. These cases provide evidence that SDP produces a more valid ranking of great powers than GDP over several historic periods.

For a second set of validations, we assess convergent validity by comparing a country's share of global SDP to several CINC component variables. While CINC has well-known drawbacks (Kadera and Sorokin 2004; Beckley 2018),[6] its components are useful for comparison as an established, widely used source of variables related to states' economic power. Because SDP and GDP are indicators of economic resources, we assess them in relation to four CINC variables that measure resources that could potentially be invested in military capabilities. The remaining CINC components, namely military expenditure and personnel, are related to actual military capabilities, which we use as a dependent variable to measure military investment below.

In the top panel of Figure 3, we display correlations between a country's share of global SDP or GDP and two components of CINC, which are yearly shares of global (1) iron and steel production and (2) primary energy consumption.[7] The year-by-year correlation coefficient with SDP ranges between 0.69 and 1 for iron and steel production and between 0.62 and 1 for primary energy consumption. Until WWII, a state's share of global SDP correlates more strongly with related measures of power resources than a state's share of global GDP. Between 1860 and 1960, the 95 percent confidence intervals do not overlap—suggesting a statistically significant difference between GDP and SDP.

In the bottom panel of Figure 3, we present analogous graphs for the correlation between the yearly share of SDP

[6] CINC's restrictive approach toward including countries as members of the international system leads to disagreements in the power estimates for some years (Supplementary Appendix C).

[7] The break in the correlation between industrial capacity and GDP in 1860 is caused by China entering the CINC dataset (with an approximate global population proportion of 47 percent according to CINC and 31 percent according to our computation).

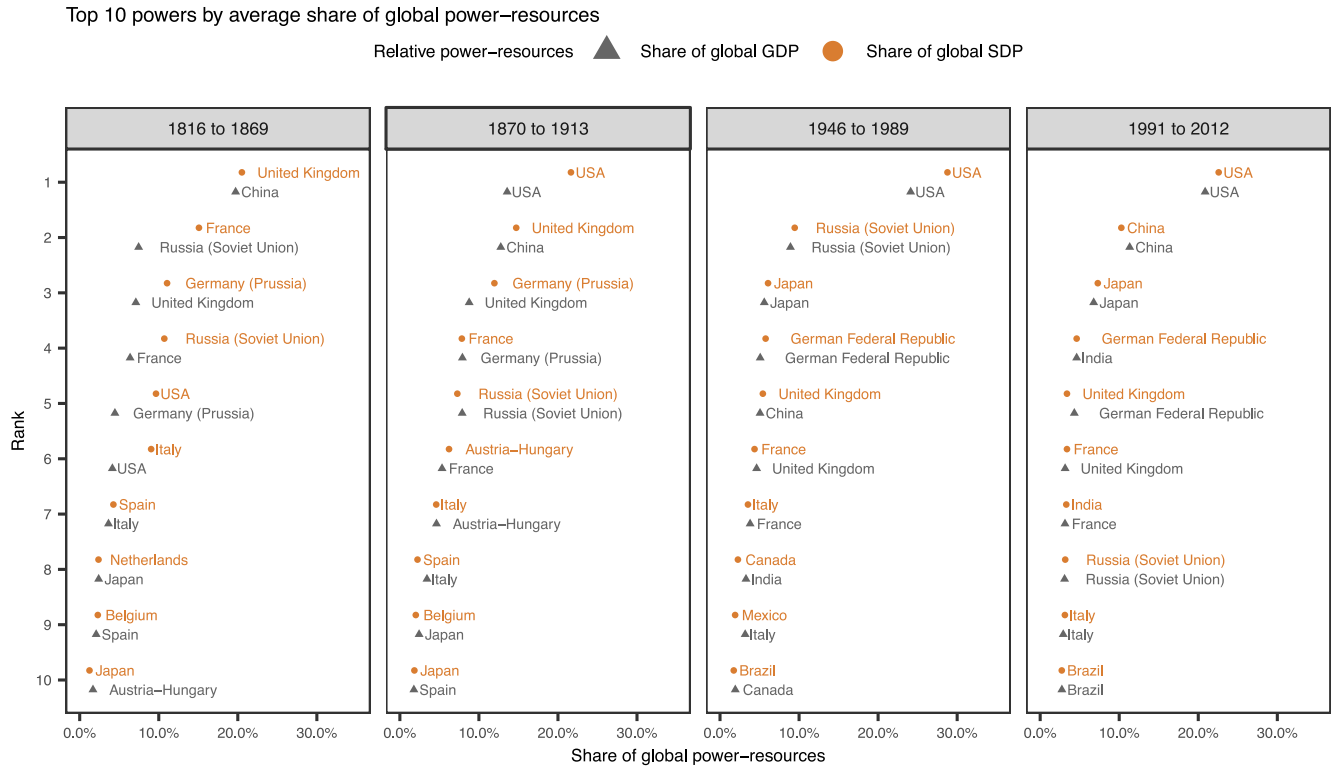

Figure 2. Top 10 powers ranked by their average share of global SDP or GDP
Note: SDP produces a more historically valid ranking of the great powers than GDP.

or GDP with a country's share of global (3) urban and (4) total population.[8] The contrast with the industrial capacity variables is striking: GDP is more closely related to a country's share of global population than SDP. The discrepancy between the two series changes over time. In the early nineteenth century, population is perfectly predictive of GDP, while SDP is not at all correlated with a country's population. In preindustrial years, and extending into the WWII period for many states, national wealth is primarily a function of how many citizens a government can exploit. The surplus resources most countries can extract from their population pre-1800 are nearly zero. Thus, the development of a force structure that goes beyond feeding soldiers and obtaining basic equipment is cost-prohibitive for most states. The Industrial Revolution changed this pattern because states began to produce economic wealth beyond the basic subsistence needs. These resources were, in turn, invested into the equipment and technology necessary to project force over distance.

The historical patterns depicted in Figure 3 demonstrate that accounting for subsistence income is crucial for the study of global power relationships, particularly in historical international relations analyses. By shifting the conceptual focus from gross to surplus domestic product, we can more accurately identify which states possess the greatest power resources and potential to generate the military capabilities to threaten other countries. States with low surplus economic resources are largely incapable of projecting power, even if their total GDP is large.

Next, we show how SDP, compared to GDP, represents a more valid measure of relative power resources between pairs of states. This dyadic-level measure, along with additional information about the state pair in question, allows us to assess the level of potential threat one state might expect from its interactions with other states.

Measuring Relative Power Resources and Potential Threat

To date, international relations scholarship had difficulty explaining patterns of arming. Even though theoretical expectations suggest that states arm in response to threats in the international system, empirical results are inconclusive (Sandler and Hartley 1995; Nordhaus et al. 2012; Cappella-Zielinski, Fordham, and Schilde 2017, 46). Our argument suggests that inconclusive results emerged because prior scholarship mismeasured power resources by relying on GDP. Our SDP-based approach addresses this systematic bias and reveals new patterns in line with theoretical expectations. States' efforts to arm (i.e., their military expenditures) should be scaled by the potential surplus resources they have available to arm. While previous scholars scaled military expenditures by GDP, we are the first to argue they should instead be scaled by SDP. Our findings demonstrate that a strong relationship between states' threat environment and efforts to arm exists, but only when military expenditures are scaled by SDP.

Relative power resources between two states is the key construct we utilize to understand patterns of arming and power projection. If military capabilities represent the latent power to hurt, then power resources represent the latent power to arm or states' military potential (Kennedy 1987). To account for the fact that not all dyadic economic power relationships are created equal, we additionally consider how other dyad-level features, specifically shared

[8] To alleviate concerns regarding the coverage of CINC's total population figures, we use our extended series of population estimates (Supplementary Appendix C).

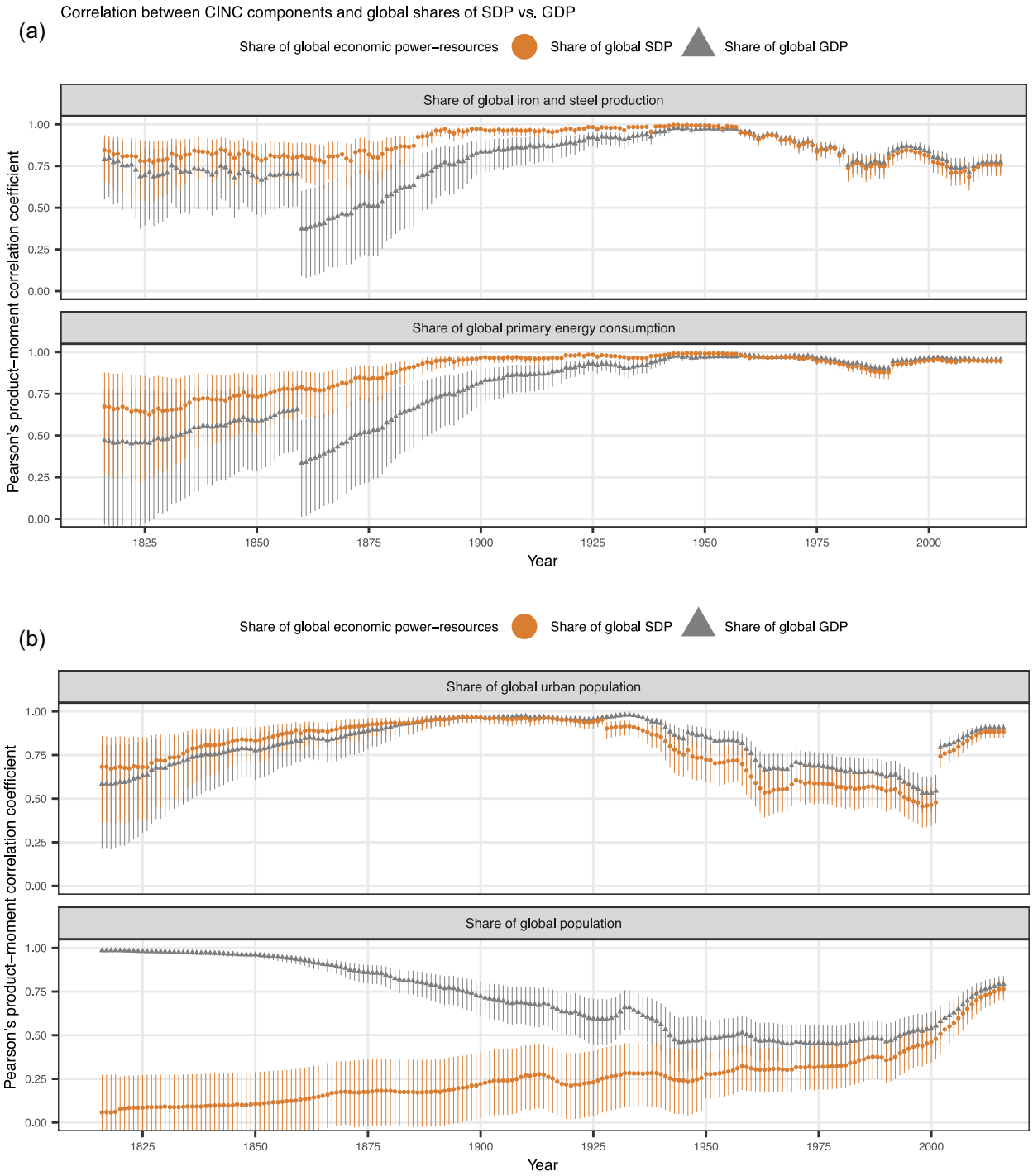

Figure 3. Yearly correlation coefficients with 95 percent confidence intervals
Notes: In each panel, we assess the degree to which a state's share of global SDP and GDP correlate with each of four component variables of CINC. Note that we utilize an alternative to CINC's population measure (Supplementary Appendix C).

preferences and the loss of strength gradient, mitigate the level of threat associated with differences in relative power resources between states.

Measuring the Difference in Relative Power Resources between States Using SDP

For each country-year unit, $i = 1, \ldots, N$ indexes states and $t = 1, \ldots, T$ indexes years. For every country-year unit, we assess information for each dyadic relationship between state i and all other states in the international system that year, indexed by $j = 1, \ldots, J$. For every state i, we measure i's annual surplus economic resources, SDP_{it}, as well as the surplus resources of each opponent, SDP_{jt}. For each ij pair, we compute power-resource ratios r_{ijt} as the proportion of the opponent state's SDP relative to the total SDP in each dyad, such that $r_{ijt} = \frac{SDP_{jt}}{SDP_{jt}+SDP_{it}}$.

The relative power-resource variable r_{ijt} is bounded between 0 and 1. When SDP_{jt} is large, the relative power resources of j compared to i will be greater than 0.5 and represent a state that is potentially threatening to i, because j has more resources to invest in arming and power projection. The least powerful state's power-resource ratio will be close to 1. The most powerful state's power-resource ratio will be close to 0. If two states have equal power resources, they find each other equally threatening and the power-resource ratio is 0.5. These ratios allow us to compare the relative power resources of one dyad to the relative power resources of another dyad, thus relativizing these comparisons and facilitating further measurement aggregation.

Relative Power-Resource Relationships between One State and All Other States

The relative distribution of power resources between two states is useful for dyadic-level analyses. However, we also want to understand how individual states respond to the total level of potential threat they face from all potential opponents in the international system. We define a country's level of potential threat as its expectations regarding other states' potential ability to harm it. All else equal, the more power resources a state has, the more potentially threatening it is. States shift from potentially threatening to actually threatening when they engage in behavior that is perceived as harmful, such as arming, coercing, or attacking. Our goal in this article is to measure the degree to which a given state is *potentially threatening* based on its relative power resources, rather than *actually threatening* based on its actions.

To measure the level of potential threat, we follow an existing operational procedure to create an aggregated country-year measure using each of the dyadic power-resource ratios.[9] This measure is simply the sum of all power-resources ratios, weighted by the loss of strength gradient and preference compatibility, between each state i and all opponent states j. This generates a country-year measure that summarizes all dyadic relationships of each country in each year and represents the unique level of potential threat each state faces.

We call this new variable $potential\ threat_{it}$, which is the sum of weighted relative power-resource ratios that relate state i to every opponent state j in year t.

$$potential\ threa\ t_{it} = \sum_{j=1}^{J} \left(r_{ijt} \times p_{ijt} \times w_{ijt} \right)$$

r_{ijt} denotes relative power resources between two states. Following Markowitz and Fariss (2018), we include weights for preference compatibility between two states, p_{ijt}, and the loss of strength gradient between two states, w_{ijt}.

The first weight, p_{ijt}, is based on the preference compatibility between states. $p_{ijt} = 0$ if state i and state j have compatible preferences with each other in year t; otherwise, $p_{ijt} = 1$ when states do *not* share compatible preferences. We operationalize joint democracy as our indicator of compatible preferences using the Polity IV data (Marshall, Gurr, and Jaggers 2016). If both states have a Polity2 value greater or equal to 6, they have compatible preferences and are not threatening to one another—and thus coded 0. If a state has incompatible preferences with other countries, they are potentially threatening, especially those with the greatest power resources. We consider alternative measures of preference compatibility in the supplementary appendix, which include other measures of joint democracy (Pemstein, Meserve, and Melton 2010; Boix, Miller, and Rosato 2013), defense pacts and alliances (Leeds et al. 2002; Gibler 2009), United Nations General Assembly voting similarity (Bailey, Strezhnev, and Voeten 2017), rivalry (Klein, Goertz, and Diehl 2006), energy consumption (Greig and Enterline 2017; Markowitz, Fariss, and McMahon 2019), bilateral trade (Barbieri, Keshk, and Pollins 2009; Barbieri and Keshk 2012), diplomatic exchange (Bayer 2006), and shared intergovernmental organization membership (Pevehouse, Nordstrom, and Warnke 2004).

The inverse logged distance between state capitals creates a second weight, $w_{ijt} = \frac{1}{\ln(d_{ijt})}$, which operationalizes the loss of strength of state j if it were to attempt to project power into state i. Short distances, d_{ijt}, between the capital cities of two states yield values closer to 1 and further distances yield values closer to 0.

Substantively, the operationalization captures the total power resources of state i relative to all of the opponent states j in the international system, weighed by distance and preference compatibility. In Supplementary Appendix E, we provide additional detail, a visual guide, and examples for each of the operational steps that generate the country-year values for the $potential\ threat_{it}$ variable.

Potential Threat Using SDP versus GDP

In this section, we compare the performance of SDP-based and GDP-based measures of potential threat. First, we compare the ability of distance- and preference-weighted power-resource ratios to correctly categorize country-year units as potentially threatening (concurrent validity). Next, we use SDP-based and GDP-based potential threat measures to explain variation in dependent variables that measure arming and power projection. We show that SDP-based models of states' military investments conform better with existing theoretical expectations than models based on GDP.

Assessing the Most Potentially Threatening States Using SDP versus GDP

Recall that the potential threat measure is the summation the distance- and preference-weighted power-resource ratios for all pairwise relationships for each country-year unit. If SDP better measures relative power resources than GDP, then the potential threat variable using SDP should more accurately rank countries of greatest concern to other countries. As a control variable, we also generate a potential threat variable that uses *population* to measure relative power resources within each pair of states. We show that the results using the population-based measure are very similar to the GDP version, which both contrast with SDP. This provides further evidence that poor, populous countries appear potentially threatening only because their large population gives them a large GDP.

Figure 4 displays the top ten potentially threatening states within the strategic environment of the United States by decade.[10] The upper panel illustrates the ranking based on distance-weighted relative power resources using SDP; the middle panel plots the ranking for an analogous measure

[9] We modify the formalization by Markowitz and Fariss (2018) to better capture the intuition that the global environment might be transitioning into an increasingly competitive space (Supplementary Appendix E).

[10] For graphs for Japan and the United Kingdom, see Supplementary Appendix Figures 3 and 4.

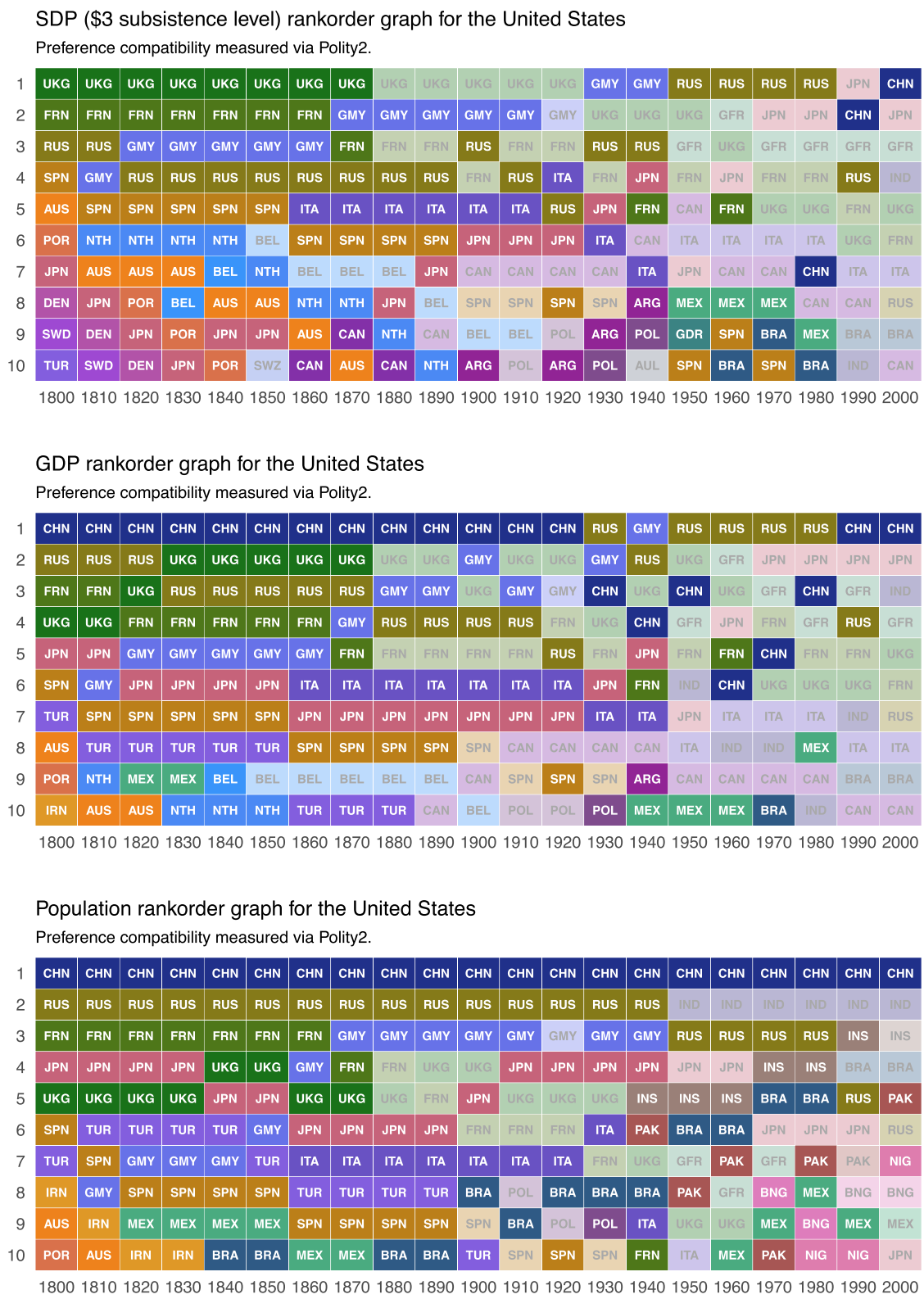

Figure 4. Top 10 potentially threatening states for the United States by decade
Notes: Lighter shading denotes non-threatening jointly democratic dyads.

using GDP; the lower panel shows this ranking for the same measure using population. States with the largest weighted power-resource ratios rank highest on the list as the adversaries that potentially threaten the United States. The upper panel's order of states differs substantially from the middle and lower panels. The distance-weighted relative power-resource measures incorporating GDP or population produce similar rankings and place countries that were unlikely to threaten the United States at the top. In particular, China is ranked as the most potentially threatening country

for the United States during the entire nineteenth and early twentieth centuries—a period China could not develop a military force structure capable of projecting force to the shores of the United States. The measure using SDP provides a more historically valid ranking of states posing a potential threat to the United States than the measures using GDP or population.

Using a lighter shading of tiles for states with compatible preferences, Figure 4 highlights the role of preference compatibility for the United States' assessment of potential threats. Recall that the p_{ijt} component of the potential threat variable down-weights power-resource ratios of jointly democratic dyads. The ranking corresponds to the top ten potentially threatening states for the United States based on economic might and geographic proximity, but this potential threat is mitigated if the country is democratic. Though they might have the economic capability to project power abroad, democratic states are not considered threatening by the United States because of the compatibility of their preferences. As former geopolitical rivals democratize, they stop contributing to the total potential threat faced by the United States. As a result, the strategic environment of the United States has become less threatening over time.

Figure 5 illustrates this downward trend. The height of each bar denotes the total level of potential threat faced by the United States each year. The potential threat faced by the United States fell sharply over the course of the nineteenth century and remains much lower today than in the past. Colored values indicate how much China and Russia contribute to the total level of potential threat for the United States. The left panel plots potential threat incorporating GDP as an indicator for economic resources; the right panel illustrates the same measure using SDP.[11]

The difference between measuring power resources via SDP versus GDP for the United States' threat assessment is striking. Historically, when the United States considered potential threats, it paid careful attention to states with the greatest power resources. For more than two hundred years, China was one of the largest *economies*, but only recently became one of the states with the greatest *power resources* in the world. SDP yields a more historically valid representation of countries' contributions to the total level of potential threat in the United States' geopolitical environment. Today, China is the largest contributor to the total potential threat faced by the United States. In fact, contemporary China makes up a larger proportion of the total potential threat faced by the United States than Russia did at the height of the Cold War. The rise of China as a major economic power in the late twentieth century dramatically increased the total potential threat that the United States experiences in its geopolitical environment.

Modeling Arming and Power Projection

As previously discussed, while international relations theory suggests that the level of potential threat states face should explain their efforts to arm, existing research finds only mixed empirical support for this proposition (Sandler and Hartley 1995; Nordhaus et al. 2012; Cappella-Zielinski et al. 2017, 46). We suggest a potential resolution for this puzzle by demonstrating that, once we scale military expenditures by SDP, which corrects for systematic measurement error inherent in using GDP, we find a strong relationship between potential threat and military burdens.

SDP not only does a better job of measuring the distribution of relative power resources—a core component in the level of potential threat states face—it also does a better job measuring the power resources that a given state could invest in arming. The example of China in 1990 illustrates this point. Military expenditures represented approximately 2.5 percent of China's GDP—a modest military burden. However, even if China's 1.135 billion citizens in 1990 could survive on just $2 per day and the state could seize the entire remainder of economic income, SDP would be half the value of its GDP. Hence, when military burden is measured using military expenditures as a percentage of SDP instead of GDP, China's military burden in 1990 was approximately twice as large as previously estimated.

Comparing SDP-based and GDP-based approaches, we apply our potential threat measure to investigate the relationship between threat and military burden. We estimate regression models in which we vary the measurement of SDP at $3, $2, and $1 per-day subsistence thresholds and compare those to GDP (equivalent to a $0 per-day threshold). We also control for states' subsistence income at these thresholds. We assess the relationship between the level of potential threat a state faces in its strategic environment (explanatory variable) and two measures of arming (dependent variables).

The first dependent variable is military burden, operationalized as military expenditure relative to income. Fearon (2018) argues that this is a reasonable proxy for states' resources that could be dedicated to arming and captures the magnitude of the social welfare costs of arming. In contrast to Fearon and others (Rasler and Thompson 1985; Khanna, Sandler, and Shimizu 1998), our preferred measure of military burden is a state's military expenditure as a proportion of the state's SDP, rather than GDP or gross national product (GNP). This operational choice better approximates surplus resources available for arming.

We employ a revised and extended series of military expenditure as a proportion of income. This new indicator is created by first converting military expenditure values from CINC into constant monetary units (Singer 1987). We then use new GDP and population estimates to measure the proportion of a state's income (SDP or GDP) devoted to the military.[12] This allows us to extend data coverage to cover most countries in the world from 1816 to 2012. As a robustness check, we assess the relationship between potential threat and a second dependent variable that captures states' investments in arming: power projection capabilities. We operationalize power projection capabilities via states' naval tonnage relative to income. Data on naval tonnage come from Crisher and Souva (2014).[13] States with higher military spending as a proportion of income have higher military burdens, as do states that have more naval tonnage relative to their income.

Figure 6 displays coefficients and 95 percent confidence intervals of standardized potential threat variables with and without the subsistence income control variable.[14] For each

[11] Preference compatibility is measured using Polity2 scores and supplemented with data from Boix et al. (2013) to reduce the number of missing values.

[12] Supplementary Appendix H describes in detail new estimates of GDP and population.

[13] Supplementary Appendix Figure 19 compares temporal trends of dependent variables.

[14] Table 1 contains regression results based on SDP for the $3 per-day subsistence threshold. Table 2 contains analogous results using GDP. For most models, we do not observe statistically significant or substantively meaningful interaction effects between potential threat based on SDP (or GDP) and population. We therefore limit the results presented in Figure 6 to additive model specifications. In the supplementary appendix, we demonstrate that the results are robust to excluding the control for population-based threat (Figure 23), omitting controls (Figure 24), and limiting observations to the post-WWII period (Figure 25).

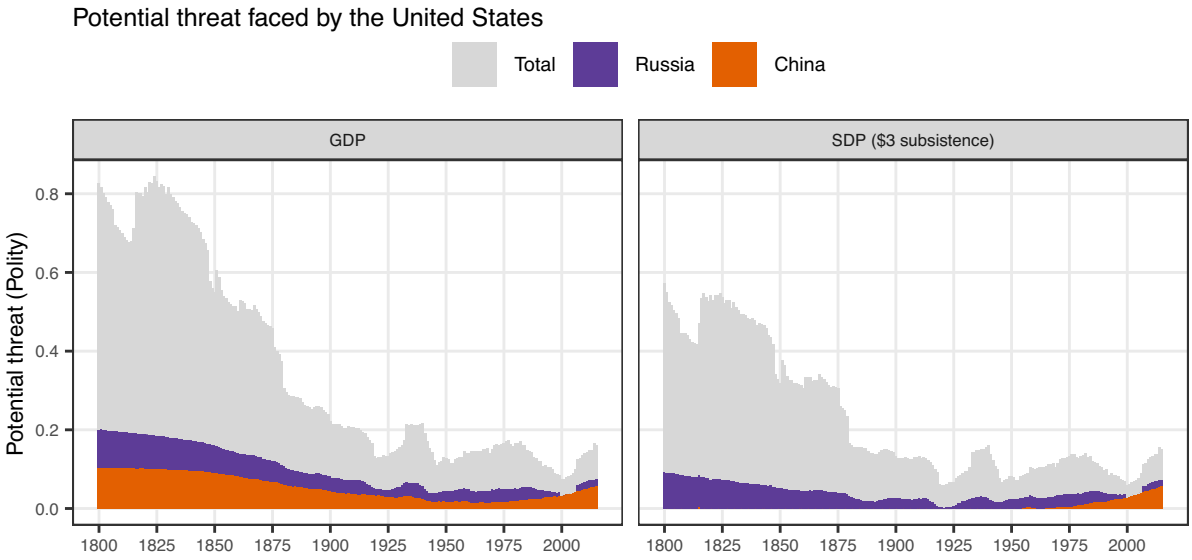

Figure 5. China's and Russia's contribution to the total potential threat faced by the United States for SDP versus GDP

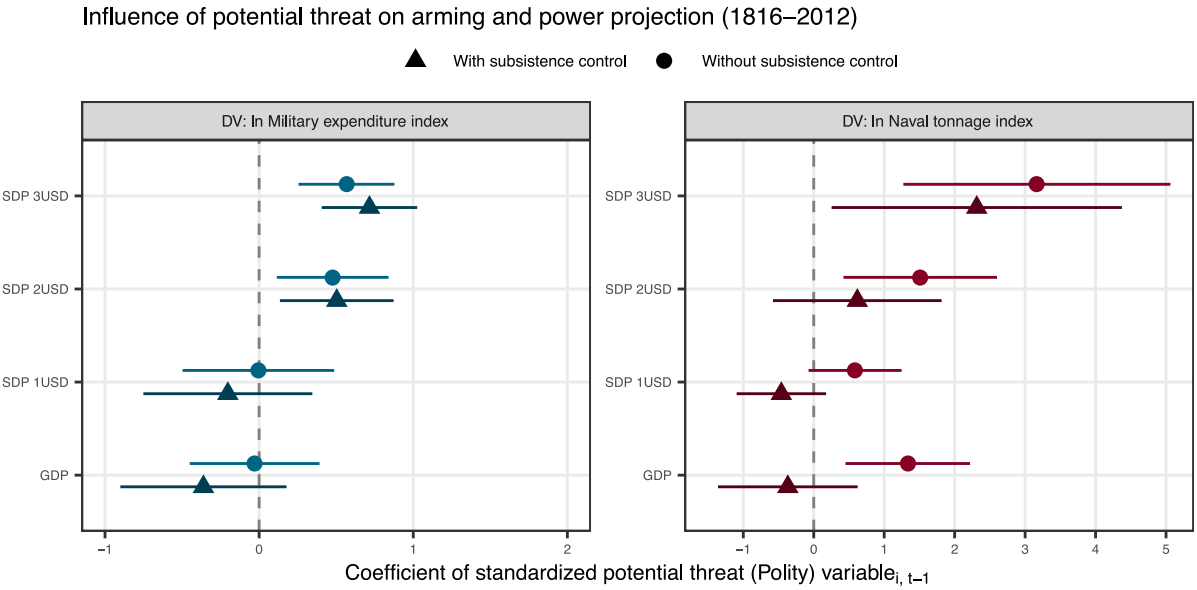

Figure 6. Coefficients and 95 percent confidence intervals of standardized potential threat variables for regression models of two dependent variables—the military expenditure index and naval tonnage index—on potential threat and control variables.
Note: See Tables 1 and 2 for model specification information.

dependent variable, we estimate a series of country-year fixed-effect regression models. Right-hand side variables are lagged by one year. All models include controls for the natural log of income (SDP or GDP), Polity2 score, and a measure of potential threat based on population. When applicable, we control for the natural log of subsistence (or population for GDP models).

Our preferred SDP-based approach to measuring power resources using a $3 per-day threshold produces statistically significant relationships between the level of potential threat states face and both measures of arming, while a GDP-based approach does not. As we adjust our measure of SDP to use lower subsistence levels, the coefficients become smaller and cease to be statistically significant at conventionally accepted levels. For many models, decreasing the subsistence level to $1, or switching to GDP, renders the effect of potential threat on military investments negative and insignificant for arming and power projection. Controlling for population-based potential threat increases the size of the estimated effects.

Overall, higher levels of potential threat are associated with larger investments in military and naval capabilities when measuring economic power resources using SDP. Crucially, all results depend on accurately measuring economic resources that states have at their disposal to invest in guns or butter. The conventionally used GDP measure does not yield a statistically significant association between the level of potential threat and arming—a

Table 1. Regression models relating different specifications of the potential threat variable to investments in arming and power projection

	Dependent variables									
	Military expenditure/SDP					*Naval tonnage/SDP*				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*
Potential threat (SDP)$_{i,t-1}$	0.63***	0.26**		0.72***	0.71***	1.82***	1.45*		2.31*	2.65*
	(0.09)	(0.10)		(0.16)	(0.16)	(0.51)	(0.58)		(1.05)	(1.20)
Potential threat (population)$_{i,t-1}$			−0.05	−0.73***	−0.75**			0.72	−1.53	0.29
			(0.13)	(0.22)	(0.26)			(0.53)	(1.10)	(1.14)
ln SDP$_{i,t-1}$		−0.08***	−0.09***	−0.06***	−0.06***		−0.05	−0.10	−0.02	−0.02
		(0.01)	(0.01)	(0.01)	(0.01)		(0.07)	(0.06)	(0.07)	(0.07)
ln subsistence$_{i,t-1}$		−0.21	−0.19	−0.37*	−0.38*		2.37*	2.82**	1.99	2.47*
		(0.19)	(0.17)	(0.18)	(0.18)		(1.06)	(0.97)	(1.07)	(1.05)
Polity2$_{i,t-1}$		−0.03**	−0.04***	−0.04***	−0.04**		0.01	−0.03	−0.01	0.03
		(0.01)	(0.01)	(0.01)	(0.01)		(0.05)	(0.05)	(0.05)	(0.05)
Interaction Potential threat$_{i,t-1}$					0.02					−1.28**
					(0.08)					(0.45)
Fixed effects	CY	CY	CY	CY	CY	CY	CY	CY	CY	CY
Observations	11,616	11,616	11,616	11,616	11,616	12,033	12,033	12,033	12,033	12,033
Adjusted R^2	0.02	0.08	0.08	0.10	0.10	0.0003	0.02	0.01	0.02	0.04

Notes: Power resources are measured using SDP at a per-diem subsistence level. The loss of strength gradient is conceptualized as curvilinear using the formula $1/\log(distance)$. Interest compatibility based on joint democracy using Polity scores.
Statistical significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Clustered standard errors by country (Satterthwaite correction) in parentheses. Potential threat variables are standardized. CY denotes two-way fixed effects. Period of observation: 1816–2012.

Table 2. Regression models relating different specifications of the potential threat variable to investments in arming and power projection

	Dependent variables									
	Military expenditure/GDP					*Naval tonnage/GDP*				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*
Potential threat (GDP)$_{i,t-1}$	0.23	−0.05		−0.36	−0.59	0.23	−0.04		−0.37	−0.23
	(0.18)	(0.15)		(0.28)	(0.39)	(0.14)	(0.14)		(0.51)	(0.49)
Potential threat (population)$_{i,t-1}$			0.01	0.36	0.18			0.03	0.39	0.59
			(0.19)	(0.39)	(0.32)			(0.16)	(0.56)	(0.61)
ln GDP$_{i,t-1}$		−0.05	−0.03	−0.17	−0.22		−0.01	0.01	−0.14	−0.13
		(0.25)	(0.25)	(0.31)	(0.32)		(0.23)	(0.21)	(0.33)	(0.32)
ln population$_{i,t-1}$		0.17	0.15	0.32	0.25		1.59***	1.58***	1.76***	1.83***
		(0.18)	(0.20)	(0.26)	(0.23)		(0.32)	(0.31)	(0.38)	(0.38)
Polity2$_{i,t-1}$		−0.05***	−0.04***	−0.05***	−0.05***		−0.02	−0.02	−0.02	−0.02
		(0.01)	(0.01)	(0.01)	(0.02)		(0.01)	(0.01)	(0.01)	(0.01)
Interaction potential threat$_{i,t-1}$					0.23					−0.18
					(0.19)					(0.10)
Fixed effects	CY	CY	CY	CY	CY	CY	CY	CY	CY	CY
Observations	11,616	11,616	11,616	11,616	11,616	12,033	12,033	12,033	12,033	12,033
Adjusted R^2	−0.03	−0.01	−0.01	−0.01	0.002	−0.02	0.08	0.08	0.08	0.08

Notes: Power resources are measured using GDP. The loss of strength gradient is conceptualized as curvilinear using the formula $1/\log(distance)$. Interest compatibility based on joint democracy using Polity scores.
Statistical significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Clustered standard errors by country (Satterthwaite correction) in parentheses. Potential threat variables are standardized. CY denotes two-way fixed effects. Period of observation: 1816–2012.

result that runs counter to existing theoretical expectations, but is consistent with the mixed empirical findings in the current literature. Only when measuring power resources via SDP can we explain arming decisions based on the level of potential threat in states' geopolitical environment.

Evaluating Military Burdens

Figure 7 illustrates the evolution of states' military burdens. Scaling military expenditures by SDP rather than GDP reveals that, historically, military burdens were higher than existing research suggests (Fearon 2018). In particular, many Asian states that spent a relatively small percentage of their GDP on the military were, in fact, laboring under extraordinarily high military burdens—spending 25 to 50 percent of surplus.

Figure 8 shows these trends for select countries. Scaling by SDP reveals that military burdens of poor states are much higher than the conventionally used measure of military expenditure as a percentage of GDP suggests. This divergence is particularly large for poor, populous countries such as China. Over time, military burdens fall for most states—especially for major powers—but these costs remain high for

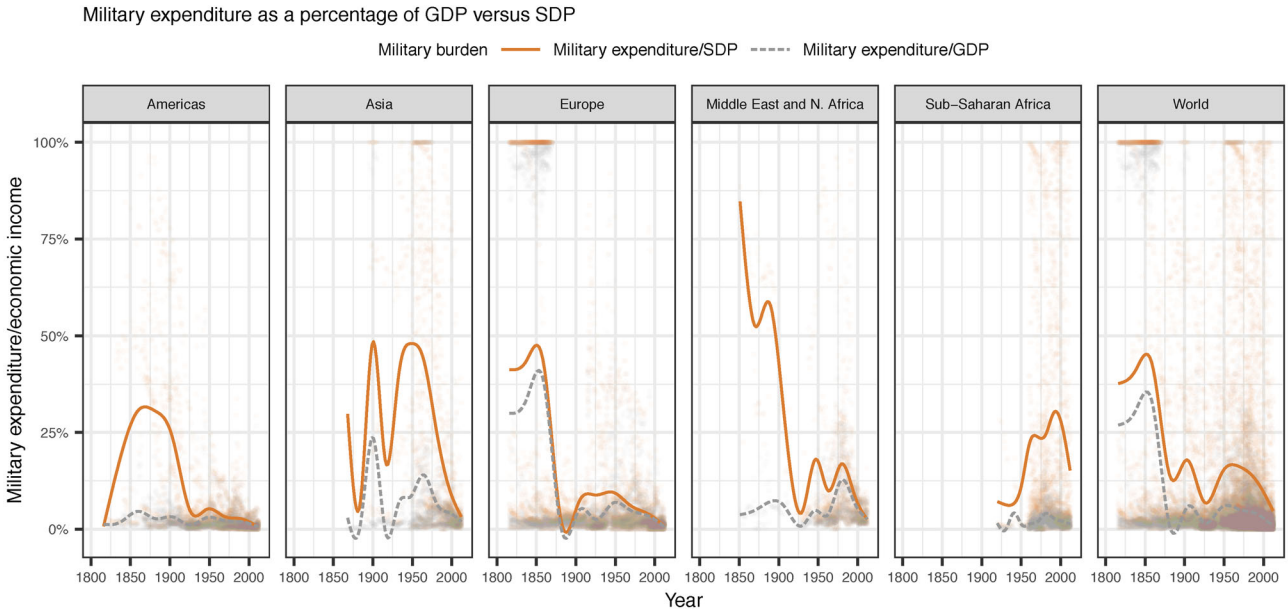

Figure 7. Change in military burden over time for regions: the Americas (including the US and Canada), Europe (including Russia), Asia, the Middle East and North Africa, and Sub-Saharan Africa
Notes: Lines represent the smoothed average over all countries in the region for two indicators of military burden: military expenditure as a proportion of SDP versus as a proportion of GDP.

poor states where most GDP is needed to cover basic subsistence and SDP is low.

The good news for states in the Western Hemisphere and Europe is that military burdens are much lower than in the past. Additionally, despite alarmist warnings of impending arms races and conflict, most states in Asia today face dramatically lower military burdens than they did in the past two hundred years. However, this decline becomes apparent only upon using SDP instead of GDP. For illustration, as a share of GDP, South Korea's military spending decreased only slightly from an average of 3.7 percent during the Cold War (1954–1991) to 3.0 percent after (1992–2012). However, as a share of SDP, military spending plunges from 9.9 percent during the Cold War to 3.2 percent after. Scaling by SDP reveals a sharp decline in South Korea's willingness to prioritize guns over butter and implies that leaders in Seoul believe that the level of threat they face fell enough to justify lower military burdens (Lind 2011). While South Korea may choose to increase its military burden in the future, they will do so from a historically low baseline. This point informs the debate over the degree to which states are balancing China's rise; in recent years, the willingness of states in the region to bear high military burdens is generally lower than commonly recognized.

Conclusion

GDP is a widely adopted measure of the financial resources that states can potentially invest in guns or butter (Coyle 2014). We introduce the concept of SDP, which separates the subsistence income, or "bread," needed for the population to survive from the surplus income, or "butter," that can potentially be extracted and invested.

Using GDP as a measure of power resources instead of SDP systematically overestimates the economic resources available to governments of poor, populous countries and underestimates the speed with which these resources increase during the early stages of industrialization (i.e., during the stage when countries first begin to produce significant surplus and governments could extract income without starving citizens). Similarly, using military expenditures as a share of GDP to measure military burdens leads scholars to underestimate the size of the military burdens born by poor states. These conceptual errors are particularly problematic for historical-comparative work because, for most of human history, virtually all states had incomes at or near subsistence levels—at least on average—creating a large divergence between estimates of SDP and GDP. This is a major issue. For illustration, as recently as 2007, the gap between SDP and GDP was still large for more than half the world's states, which were classified as either low income (gross national income [GNI] per capita below $995) or lower-middle income (GNI per capita between $996 and $3895, World Bank 2018). At the time, more than 70 percent of the world's population lived in such states.

Thus, using new data on SDP, we reveal that poor, populous states are far less powerful than generally assumed and that low-income countries historically and today face more severe guns-butter tradeoffs and higher military burdens than GDP-based measures suggest. Previous scholarship dramatically underestimated the benefits of factors that allow poor states to lower defense burdens, such as hierarchy and the liberal peace. Additionally, our results offer a potential solution to the puzzle of why previous scholarship found only mixed support for one of the core propositions of international relations theory: *states arm against potential threats*. Once we apply SDP to correct for the systematic measurement error associated with GDP, we find strong empirical support for this proposition.

In addition to our theoretical contributions, we provide new data that extend cross-national coverage of GDP, SDP, and population back to 1816 for nearly every country. These new data allow scholars to apply our measure of SDP to reexamine a broad range of research questions in which

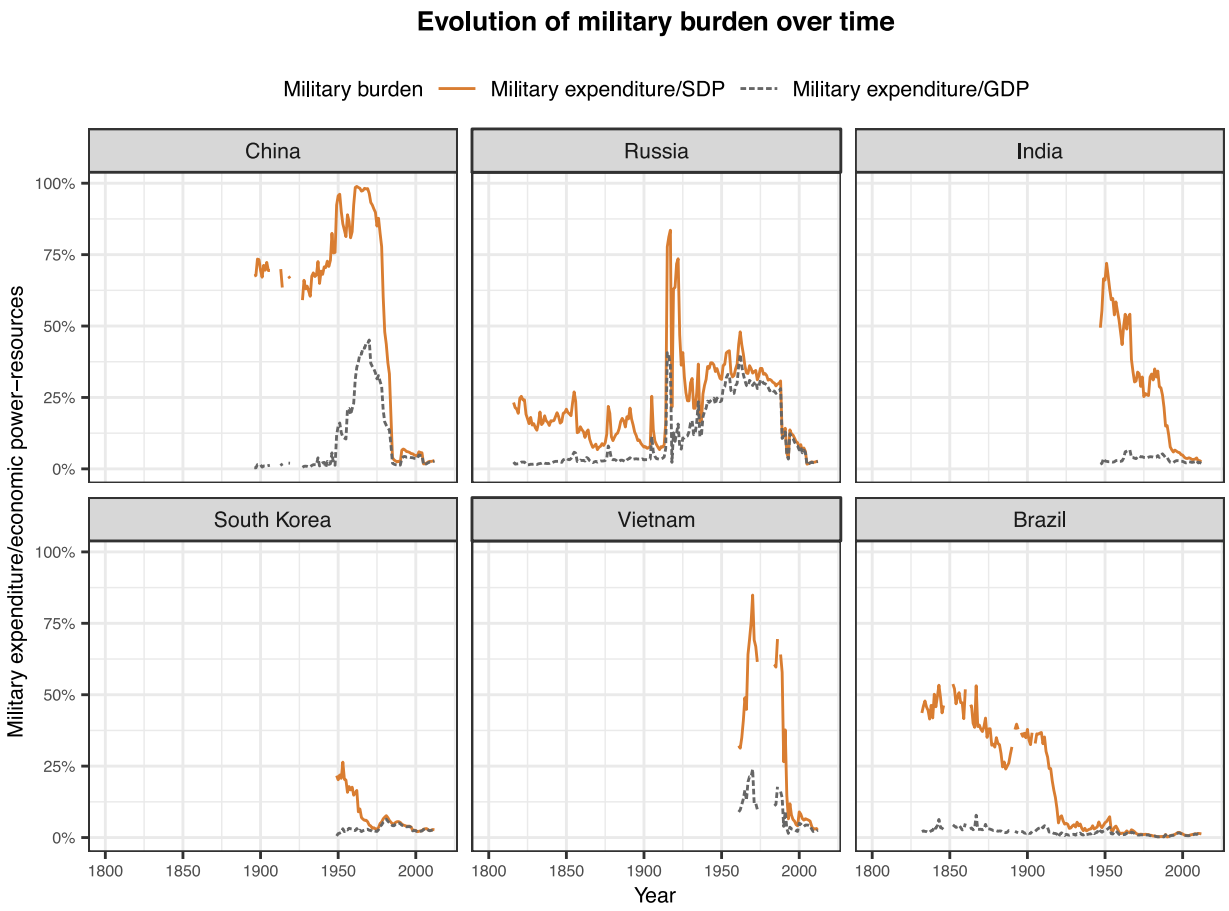

Figure 8. Change in military burden over time for select countries

GDP is frequently used as a proxy for potential or actual state capacity.

An area directly parallel to states' capacity to arm is states' capacity to repay debts. For illustration, Malawi in 2015 had a plausibly manageable debt burden of 39.5 percent of GDP (Reinhart and Rogoff 2010; World Bank 2017). However, Malawi's debt burden amounts to a crushing 226.5 percent of SDP. The annual payments on a debt of that magnitude consume a significant proportion of the country's surplus income, even if they constitute a seemingly manageable share of total economic income. This matters for understanding the ability of states to manage debt and engage with international institutions such as the World Bank and the International Monetary Fund.

More broadly, SDP represents the resources a government may potentially draw upon to build physical infrastructure, establish the rule of law, or provide public services such as education and health care. SDP is not a direct measure of government capacity, but it measures the upper bound of the income states may sustainably extract from citizens to develop that capacity.

When measuring the upper bound of states' extractable income, SDP, compared to GDP, showcases just how constrained state capacity is in low-income countries, compared to middle-income and higher-income countries. However, SDP also reveals that, in the early stages of economic development, government capacity expands more rapidly than currently realized—increasing the compound returns to growth for countries near subsistence levels. SDP-based assessments are unlikely to lead to exclusively pessimistic or exclusively optimistic new conclusions regarding the prospects for peace, prosperity, and democracy in the developing world. Thus, the analysis of SDP has the potential to radically reshape our understanding of the extent to which different fiscal strategies are plausible or desirable in lower-income countries.

Supplementary Information

Supplementary information is available at the *International Studies Quarterly* data archive.

References

Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61 (2): 430–56.

Barbieri, Katherine, and Omar Keshk. 2012. "Correlates of War Project Trade Data Set Codebook." Last Accessed January 6, 2019. http://correlatesofwar.org.

Barbieri, Katherine, Omar M.G. Keshk, and Brian M. Pollins. 2009. "Trading Data: Evaluating Our Assumptions and Coding Rules." Conflict Management and Peace Science. Last Accessed January 6, 2019. http://cmp.sagepub.com/content/26/5/471.abstract.

Bayer, Resat. 2006. "Diplomatic Exchange Data Set V2006.1." Last Accessed January 6, 2019. http://correlatesofwar.org.

Beckley, Michael. 2018. "The Power of Nations: Measuring What Matters." *International Security* 43 (2): 7–44.

Boix, Carles, Michael Miller, and Sebastian Rosato. 2013. "A Complete Data Set of Political Regimes, 1800–2007." *Comparative Political Studies* 46 (12): 1523–54.

Cappella-Zielinski, Rosella. 2016. *How States Pay for Wars*. Ithaca, NY: Cornell University Press.

Cappella-Zielinski, Rosella, Benjamin O. Fordham, and Kaija E. Schilde. 2017. "What Goes Up, Must Come Down? The Asymmetric Effects of Economic Growth and International Threat on Military Spending." *Journal of Peace Research* 54 (6): 791–805.

Coyle, Diane. 2014. *GDP: A Brief But Affectionate History*. Princeton: Princeton University Press.

Crisher, Brian, and Mark Souva. 2014. "Power At Sea: A Naval Power Dataset, 1865–2011." *International Interactions* 40 (4): 602–29.

Ezzati, Majid. 2004. "Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors." World Health Organization. Last Accessed January 6, 2019. http://apps.who.int/iris/bitstream/10665/42792/1/9241580348_eng_Volume1.pdf.

Fearon, James D. 2018. "Cooperation, Conflict, and the Costs of Anarchy." *International Organization* 72 (3): 523–59.

Ferreira, Francisco. 2015. "The International Poverty Line Has Just Been Raised to $1.90 a Day, But Global Poverty Is Basically Unchanged. How Is That Even Possible?" *Let's Talk Development* (blog), *World Bank*, October 4. http://blogs.worldbank.org/developmenttalk/international-poverty-line-has-just-been-raised-190-day-global-poverty-basically-unchanged-how-even.

Garfinkel, Michelle R., and Constantinos Syropoulos. 2019. "Problems of Commitment in Arming and War: How Insecurity and Destruction Matter." *Public Choice* 178 (3–4): 349–69.

Gibler, Douglas M. 2009. *International Military Alliances, 1648–2008*. Washington, DC: CQ Press.

Greig, J. Michael, and Andrew J. Enterline. 2017. "National Material Capabilities (NMC) Data Documentation Version 5.0." Correlates of War Project. Last Accessed January 6, 2019. http://www.correlatesofwar.org/data-sets/national-material-capabilities.

Kadera, Kelly M., and Gerald L. Sorokin. 2004. "Measuring National Power." *International Interactions* 30 (3): 211–30.

Kennedy, Paul. 1987. *The Rise and Fall of the Great Powers: Economic Change and Military Power from 1500 to 2000*. New York: Random House.

Khanna, Jyoti, Todd Sandler, and Hirofumi Shimizu. 1998. "Sharing the Financial Burden for U.N. and NATO Peacekeeping, 1976–1996." *Journal of Conflict Resolution* 42 (2): 176–95.

Klein, James P., Gary Goertz, and Paul F. Diehl. 2006. "The New Rivalry Dataset: Procedures and Patterns." *Journal of Peace Research* 43 (3): 331–48.

Lake, David A. 2009. *Hierarchy in International Relations*. Ithaca, NY: Cornell University Press.

Lake, David A., and Matthew A. Baum. 2001. "The Invisible Hand of Democracy: Political Control and the Provision of Public Services." *Comparative Political Studies* 34 (6): 587–621.

Lamborn, Alan C. 1983. "Power and the Politics of Extraction." *International Studies Quarterly* 27 (2): 125–46.

Leeds, Brett, Jeffrey Ritter, Sara Mitchell, and Andrew Long. 2002. "Alliance Treaty Obligations and Provisions, 1815–1944." *International Interactions* 28 (3): 237–60.

Lind, Jennifer. 2011. "Democratization and Stability in East Asia." *International Studies Quarterly* 55 (2): 409–36.

Markowitz, Jonathan N, and Christopher J Fariss. 2018. "Power, Proximity, and Democracy: Geopolitical Competition in the International System." *Journal of Peace Research* 55 (1): 78–93.

Markowitz, Jonathan N., Christopher J. Fariss, and R. Blake McMahon. 2019. "Producing Goods and Projecting Power: How What You Make Influences What You Take." *Journal of Conflict Resolution* 63 (6): 1368–402.

Marshall, Monty G., Ted R. Gurr, and Keith Jaggers. 2016. "Polity IV Project. Political Regime Characteristics and Transitions, 1800–2015, Dataset Users' Manual." Vienna, VA: Center for Systemic Peace.

Milward, Alan S. 1977. *War, Economy, and Society, 1939–1945*. History of the World Economy in the Twentieth Century. Vol. 5. Berkeley: University of California Press.

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 (3): 491–513.

Norloff, Carla, and William C Wohlforth. 2019. "Raison De L'hégémonie (The Hegemon's Interest): Theory of the Costs and Benefits of Hegemony." *Security Studies* 28 (3): 422–50.

Pemstein, Daniel, Stephen A. Meserve, and James Melton. 2010. "Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type." *Political Analysis* 18 (4): 426–49.

Pevehouse, Jon, Timothy W. Nordstrom, and Kevin Warnke. 2004. "The Correlates of War 2 International Governmental Organizations Data Version 2.0." *Conflict Management and Peace Science* 21 (2): 101–19.

Rasler, Karen A., and William R. Thompson. 1985. "War Making and State Making: Governmental Expenditures, Tax Revenues, and Global Wars." *American Political Science Review* 79 (2): 491–507.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2010. "Growth in a Time of Debt." *American Economic Review* 100 (2): 573–8.

Sandler, Todd, and Keith Hartley. 1995. *The Economics of Defense*. New York: Cambridge University Press.

Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985." *International Interactions* 14 (2): 115–32.

Tanzi, Vito, and Ludger Schuknecht. 2000. *Public Spending in the 20th Century. A Global Perspective*. Cambridge: Cambridge University Press.

World Bank. 2017. "World Development Indicators." Last Accessed January 6, 2019. http://data.worldbank.org/data-catalog/world-development-indicators.

———. 2018. "Classifying Countries By Income." Last modified October 4, 2018. http://datatopics.worldbank.org/world-development-indicators/stories/the-classification-of-countries-by-income.html.