Original Manuscript

Journal of Conflict Resolution
2024, Vol. 0(0) 1–28

DOI: 10.1177/00220027241232964
journals.sagepub.com/home/jcr
Sage

Measuring Arms: Introducing the Global Military Spending Dataset

Miriam Barnum[1], Christopher J. Fariss[2], Jonathan N. Markowitz[3], and Gaea Morales[3]

Abstract

Military spending data measure key international relations concepts such as balancing, arms races, the distribution of power, and the severity of military burdens. Unfortunately, missing values and measurement error threaten the validity of existing findings. Addressing this challenge, we introduce the *Global Military Spending Dataset* (GMSD). GMSD collates new and existing expenditure variables from a comprehensive collection of sources, expands data coverage, and employs a latent variable model to estimate missing values and quantify measurement error. We validate the data and present new findings. First, correlations between economic surplus and military spending are currently higher than at any point in the last two-hundred years. Second, updating DiGiuseppe and Poast's (2018) analysis, we find larger substantive effects. Specifically, we find that the (negative) effect of a democratic ally on military spending is three times larger, and the (positive) effect of an increase in GDP is five times larger than previously estimated.

Keywords

power, capabilities, military power, measurement, latent variable modeling

[1]Department of Political Science, Purdue University, West Lafayette, IN, USA
[2]Department of Political Science, University of Michigan, Ann Arbor, MI, USA
[3]School of International Relations, University of Southern California, Los Angeles, CA, USA

Corresponding Author:
Christopher J. Fariss, Department of Political Science, University of Michigan, 426 Thompson St, Ann Arbor, MI 48109-1382, USA.
Email: cjf0006@gmail.com

Data Availability Statement included at the end of the article

Introduction

While the world has become more peaceful, global military spending is now three times greater than at the height of the Cold War. Recently, in response to Russia's invasion of Ukraine, the German government approved an unprecedented increase in military spending of 100 billion Euros. These developments motivated renewed interest in the drivers and implications of military spending by policymakers and scholars. Findings on the relationship between threat and arming, arms races and war, and military expenditures and alliance obligations hinge on the completeness and accuracy of existing military spending data. Moreover, scholars use data on military spending to measure important concepts such as balancing, arms races, the distribution of power, and the severity of military burdens.

Existing arming datasets suffer from two key problems that undermine scholars' ability to draw valid inferences. First, there are at least 76 available variables from 9 dataset collection projects (Correlates of War (COW), Stockholm International Peace Research Institute (SIPRI), International Institute for Strategic Studies (IISS), and several others) that publish estimates of state military spending levels. Disagreement on the actual expenditure value for a given country-year is common, even between datasets produced by the same project. Not only is there no central repository containing all of these datasets, there is also no easy way of comparing across datasets given that some country-year estimates are reported using constant USD while others use current USD, GBP, or local currencies. Second, the extent of measurement error in the existing data is unknown, and there is no way to assess which existing datasets are most reliable. Thus, while scholars acknowledge potential measurement error, there is no easy way to incorporate this uncertainty into analyses, leaving open the possibility that existing findings are driven by measurement error.

We offer a systematic solution to these problems by introducing the *Global Military Spending Dataset* (GMSD). First, by gathering a central repository of 76 variables from 9 dataset collection projects, we provide the most comprehensive and complete set of published datasets on military spending ever assembled. Second, our latent variable model provides researchers with a principled means of systematically comparing agreement/disagreement across these data sources, which use a variety of different estimation approaches (from original sources) and currency units. This alone represents a major contribution that will save scholars many hours of effort on data collection and management. Third, by generating distributions of predicted expenditure values for each country-year, rather than just point estimates, our measurement model provides a means of quantifying uncertainty. These estimated distributions account for disagreement between data from different sources and variation in the level of coverage within each source. By focusing on uncertainty estimates, scholars can begin to systematically investigate whether previous findings are robust or potentially driven by measurement error.

Incorporating measurement uncertainty into observational models is an important validation step. As Fariss et al. (2022) discuss, models that do not incorporate uncertainty cannot rule out the possibility that statistical associations (or lack thereof) between military spending and other variables are either false-negative results (Type 2 error or attenuation bias) or false-positive results (Type 1 error). In observational models with multiple indicators, bias due to measurement error is not always attenuating if un-modeled, higher-order interactions exist.

Fourth, we fill in missing values by using a dynamic latent variable model and a more complete set of sources. Nearly one in five observations in the COW National Military Capabilities dataset has a missing value for military spending. Even supplementing with information from the 8 other data collection projects, there are still no observed values for 2579 out of 17,855 country-year-units (14.4 percent).[1]

To demonstrate the features and validity of our newly estimated arming data, we complete several empirical evaluations. First, we present a comparison of the distribution of observed arming values and the distribution of predicted intervals from our estimated arming variables. Second, we present the distribution of uncertainty estimates for several countries over time. Third, we present the range of uncertainty estimates for each of the 76 original arming variables across the nine dataset projects in addition to spatial-temporal coverage of each dataset.

Finally, we employ these data to generate two key insights that enhance our understanding of relationship economic and military power. First, correlations between economic surplus and military spending are currently higher than at any point in the last two-hundred years. Second, updating DiGiuseppe and Poast's (2018) analysis, we find much larger substantive effects. Specifically, we find that the (negative) effect of a democratic ally on military spending is three times larger, and the (positive) effect of an increase in GDP is five times larger than previously estimated.

Overall, these empirical results showcase the strengths of our new expenditure estimates. We hope that scholars find these estimates useful to both incorporate uncertainty into existing findings and expand the coverage for new analyses. Having established that addressing the problems associated with existing arming data has important implications for all areas of social science research, we now turn to demonstrating that we have used rigorous and appropriate methods to construct a new dataset measuring military spending. We close with several demonstrations of how our new estimates can be applied by scholars to improve research and discuss several productive pathways for future research.

Why Measures of Arming are Important

Military spending is a process of conceptual importance to the study of conflict and many other areas of inquiry in international politics. But how we study military spending is affected by how we measure it. Moreover, measuring military spending is of critical importance because international relations scholars use these measures to operationalize a broad range of concepts that they then use as independent or dependent variables.

First, military spending is directly employed as a measure of states' arming and arms racing. Some scholars have evaluated the impact of structural factors such as economic growth (Cappella Zielinski, Fordham, and Schilde 2017; Goldsmith 2003), the nature of military technology (Coe 2018; Coe and Vaynman 2020), and the threat environment faced by states (Anders et al., 2020; Böhmelt and Bove, 2014; Fearon, 2018; Lebovic, 1995; MacDonald and Parent, 2011; Nordhaus et al., 2012; Rider, 2013; Sample, 2018) on levels of arming. Others have focused on how domestic factors such as regime type (Bueno de Mesquita et al. 2004; Fordham and Walker 2005; Lebovic 2001), parochial economic interests (Fordham 2008 2019), internal threat or civil wars (Mohammed 1996), and civil military relations (Bove and Nisticò 2014; Flynn 2014; Ostrom 1978) influence states' absolute military spending and their military burdens (i.e., military spending scaled by GDP).

Second, scholars use military spending to operationalize various other concepts. For example, the ratio of a state's military spending to military personnel acts as a measure of a country's investment in capital intensive military technologies. This measure in turn has been used as a proxy for the quality of a state's military forces (Huth, Bennett, and Gelpi 1992), fighting technology (Gartzke and Rohner 2010), or the degree to which one chooses to invest in the capital intensive power projection capabilities necessary to be a great power (Fordham 2011, 592). Military spending has also been used to measure the benefits of alliances and hierarchy. Oneal and Whatley (1996) investigate whether alliances allow states to spend less on defense, while Lake (2007) shows that a key benefit of hierarchy is reduced defense spending by subordinate states.

Military spending is frequently used as a component in the operationalization of relative military power.[2] There is also literature on whether concerns about potential rivals' military spending inhibits trade (Gowa and Mansfield 1993; Kleinberg and Fordham 2013; Morrow 1997; Powell 1993). Military spending also plays an important role in explaining domestic outcomes such as economic growth (Carter, Ondercin, and Palmer 2021), food security (Jenkins, Scanlan, and Peterson 2007), and civilian victimization in conflict (Downes 2006). For a summary of these variables and how they are used, see Table 1.

Military Expenditure Data

We are grateful to build on existing arming datasets which represent a massive amount of human coding and scholarly effort. However, some of the information generated by these projects in published articles is no longer publicly available or is not linked with the information contained in existing datasets. The NMC dataset (Greig and Enterline, 2017, 2021) currently provides documentation of the historical source material used to generate the military expenditure variable, but the documentation is incomplete. No citation information is provided for approximately half the country-year entries for military spending from 1816 to 2000 and, even when citation information is provided, the relationship between the documentary sources and the recorded expenditure value is not always clear. Early publications associated with the project included bibliographical details for the source material (Singer 1987; Singer, Bremer, and Stuckey 1972), but neither those publications nor the current version of the dataset link these originally

Table 1. Common Variables Dependent on Military Spending.

Variable/Construct that Military Spending Estimates are Used to Measure	What Does the Variable Measure?	How is Military Spending Used to Operationalize the Variable?
Arms Levels (Cappella Zielinski et al., 2017; Goldsmith, 2003; Huth et al., 1992; Reiter and Stam, 1998; Wohlforth, 1999, 2009)	• The distribution of military power • A state's military capabilities • A predictor of victory in war	Used as a direct measure of arms levels
Military Burdens (Anders, Fariss, and Markowitz 2020; Carter, Ondercin and Palmer 2021; Fearon 2018; Jenkins, Scanlan and Peterson 2007; Lake 2007; Mohammed 1996; Nordhaus, Oneal, and Russett 2012; Oneal and Whatley 1996; Powell 1993; Rider 2013)	The share of national income resources dedicated to arming: • Proxy measure of threat (perceptions) • Intensity of arms racing • Guns vs. butter tradeoff	Ratio of military spending to population, GDP, or SDP
Military Capital Intensity (Coe 2018; Coe and Vaynman 2020; Fordham 2011; Huth, Bennett and Gelpi 1992)	• Degree to which a military substitutes capital for labor • Proxy measure for a military's level of technological sophistication	Ratio of military spending to military personnel
Composite Indicator of National Capabilities (Huth et al., 1992; Lake, 2007; Oneal and Whatley, 1996; Reiter and Stam, 1998; Sullivan, 2007; Wohlforth, 1999, 2009)	Global Distribution of Power: • States' power resources • Predictor of victory in war	Share of global military spending is one of six indicators averaged to produce CINC scores

cited sources with the individual dataset values (country-year unit values). We address these issues, first, by linking observed values across datasets, and second, by collecting and incorporating existing and new historical source material (note that the collection of existing and new historical material is ongoing).

Fortunately, other dataset projects provide information on military spending that we can compare with NMC (see Table 2). However, these datasets tend to have more limited temporal coverage and nearly all begin only after World War II. NMC is the only dataset project that provides global estimates prior to 1948. We recognize the need for further data collection, particularly in the pre-WWII period, to better identify potential shifts in the global distribution of military spending over time. The estimation

Table 2. Existing Military Expenditures Datasets.

Source/Author	Name of Dataset (*abbreviation*)	Number of Countries	Number of Expenditure Variables	Years Covered
International Institute for Strategic Studies	The military balance (*IISS*)	177	46	1961–2021
Stockholm International Peace research Institute	SIPRI Database (*SIPRI*)	171	2	1949–2022
U.S. Arms control and disarmament Agency and state Department	World military expenditures and arms Transfers (*WMEAT*)	159	17	N/A
William Zimmerman	Words and deeds in Soviet military expenditures (*Zimmerman/USSR*)	1	4	1955–1983
B. Guy Peters	Political System's Performance data: Sweden (Peters/ Sweden)	1	1	1865–1967
Charles Lewis Taylor and Joachim Amm	National capability data: Annual series (*NCD*)	133	4	1950–1988
John Gillespie and Dina Zinnes	Military Defense expenditure data (*MDED*)	63	1	1948–1970
Jan Faber	Annual data on nine economic and military characteristics of 78 nations (*SIRE NATDAT*)	78	1	1948–1983
National military capabilities (correlates of war Project)	NMC v6 data (*NMC [COW]*)	217	1	1816–2012/ 1816-2016

procedure we establish here can be extended to accommodate new sources of data as they are coded from the historical record.

Though we continue collecting historical information, we have already collected all existing datasets that are available electronically, merged and harmonized these datasets, and used a dynamic latent variable modeling approach to estimate a unified measure of spending with uncertainty based on a similar model developed for measuring Gross Domestic Product (GDP), GDP per capita, and population (Fariss et al. 2022). The latent variable model and additional dataset projects post-1945 allow us to estimate country-year observations missing in the NMC data through the dynamic structure of the model in the same manner as Fariss et al. (2022). Estimates for the pre-WWII period tend to have wider predictive intervals (more uncertainty) because of the more limited data. The uncertainty intervals tell us both how many observed dataset variables cover a country-year unit and how much agreement or disagreement exists within the sources. We describe these features in detail after reviewing each of the component datasets.

The Military Balance

Annual editions produced by the International Institute for Strategic Studies (IISS) (1961-Present) provide comprehensive data on defense economics and military and security capabilities (IISS 2022). IISS follows the North Atlantic Treaty Organization (NATO) concept of military expenditure which includes cash outlays by central or federal governments to meet the costs of national armed forces, divided into four categories: Operating Costs, Procurement and Construction, Research and Development, and Other Expenditure. Operating Costs, the largest category, covers salaries and pensions for military and civilian personnel, cost of maintaining and training units, service headquarters and support elements, and servicing military equipment and infrastructure. IISS cites official defense budgets as reported by national governments and international organizations, including the United Nations (UN), the Organization for Security and Co-operation in Europe (OSCE), and the International Monetary Fund (IMF). Additional data from the Organization for Economic Cooperation and Development (OECD), the World Bank, and regional banks supplement incomplete measures of military spending. We use optical character recognition to read in tables from annual reports and extract expenditure data, reported in either constant or current millions USD (varies by issue).

Stockholm International Peace Research Institute

Hosts a military expenditure database (1949–2020) that is updated as new sources become available. Expenditure includes capital spending on armed and peacekeeping forces, defense ministries and government agencies in defense projects, paramilitary forces (to be trained, equipped, and available for military operations), and military space activities, as well as spending on personnel (military and civil), retirement pensions, social services (for military personnel and dependents), operations, maintenance and procurement, research and development, construction, and military aid (reflected in donor country). SIPRI excludes spending on civil defense and past military activities (e.g., veteran's benefits, demobilization, conversion of arms facilities, destruction of weapons). SIPRI collects data from primary sources including national government documents, publications, questionnaires, secondary sources citing primary sources, and other secondary sources.

World Military Expenditures and Arms Transfers

A series of reports (10-year periods) providing data on annual military expenditures, arms transfers, armed forces, and select economic data for 159 countries, produced by the US Arms Control and Disarmament Agency until 1999 and now published by the State Department. Expenditure includes spending by all national defense agencies (except civilian programs), including mixed military-civilian activities (e.g., atomic, energy, space, research and development, and paramilitary forces). Source data is from national government publications and websites, the publications and data resources of US government agencies, standardized annual "Vienna Document" reporting to the Forum for Security Cooperation of the OSCE, and other international organization reports. Sources used to evaluate military spending include the Government Finance Statistics (IMF), the SIPRI Military Expenditure Database, *The Military Balance* (IISS), Jane's Defence Budgets and the "Defence Budget" sections of Jane's Sentinel Security Assessments, the European Defence Agency's (EDA) Defence Data Portal for non-NATO EDA member states, country armed forces profiles on defenceWeb for sub-Saharan African countries, and various media reports.

Words and Deeds in Soviet Military Expenditures, 1955–1983

Includes content analysis of finance ministers' speeches and estimates of Soviet defense spending (Zimmerman 1985; Zimmerman and Palmer 1983). We extract three military expenditure variables: defense spending (estimates by Cusack and Ward 1981), and 'high' and 'low' defense spending estimates from the United States Congress Joint Economic Committee. All data are reported in rubles.

Political Systems Performance Data: Sweden

Provides historical annual performance data for Sweden from 1865 to 1967 (Peters 1992), including demographic data and various government expenditures. We extract the defense expenditure data, deflated by the recorded price index (mean price index for years 1897–1902 as base) and reported in 1000 kronas.

National Capability Data

Annual Series includes economic, military, and population indicators for 133 countries, from 1950 to 1988 (Taylor and Amm 2006). Expenditures include both current and capital expenditures on the armed forces and defense department services, the cost of paramilitary forces and police trained or equipped by military operations, research and development, and costs associated with pensions of service personnel, including civilians. Military aid is included in the donor country budget, while civil defense, war debts, and veterans' pensions are excluded. Data on military expenditure were sourced from issues of SIPRI yearbooks (1972, 1979, 1983, 1986, 1987, and 1989) and were separated into groups, provided in constant millions of USD with different base years. We extract all reported military expenditure variables.

Military Defense Expenditure Data, 1948–1970

Provides defense expenditure data for 63 countries in either millions USD or national currency (with exchange rates provided) (Gillespie and Zinnes 1992). MDED is sourced from the UN Statistical Yearbook, U.S. ACDA publications (WMEAT), and the UN Statistical Bulletin for Latin America.

Annual Data on Nine Economic and Military Characteristics of 78 Nations (SIRE NATDAT), 1948–1983

Reports annual macroeconomic and military indicators, including defense expenditure, for countries independent since 1948 as well as Indonesia, Japan, and South Korea, independent in 1949 (Faber 1992). Data on defense expenditures were taken from the SIPRI Yearbooks and the annual Military Balance (IISS), and missing values derived from Statesman's Yearbooks. We extract annual observations of defense expenditures, in current millions USD.

National Military Capabilities

We use the military expenditure variable from the COW's NMC v6.0 data, covering 1816–2016 (Greig and Enterline 2017, 2021). The originators of the dataset used a large number of historical sources (Singer 1972 1987; Singer, Bremer, and Stuckey 1972). However, as discussed above, the relationship between the historic source material and the recorded data values is not always clear based on the current documentation.

Latent Variable Model of Arming

We are interested in estimating military spending at the country-year level. Military expenditure data presents several challenges because the datasets are incomplete, cover short periods of time, and are presented in many different monetary units-of-measurement. To overcome these challenges, we specify a dynamic latent variable measurement model that links all of the available information across different contemporary and historical sources of arms spending data. We essentially want to estimate the country-year distribution or simply the average of military spending across all the available observed dataset values so that we generate the best estimate of military spending for each of the country-year-units. However, unlike most latent variable models, we are not substantively interested in the estimated latent variable parameters themselves.[3]

The latent variable estimates are important because they are the parameters that link all of the country-year data together. However, to evaluate the model, we need to transform the latent variable estimates into meaningful monetary units of measurement so that we can compare the model estimates to the observed dataset values. We do this by using the estimated latent variable model to generate posterior prediction intervals for each of the observed variables in their original unit-of-measurement (transforming the latent variable back into the original monetary unit-of-measure of the observed dataset values for comparison). Note that we generate these prediction intervals for each observed item even for country-years where the original dataset has a missing value for that item. These are the posterior prediction intervals referenced throughout the rest of the article. The latent variable itself is essentially an average for each country-year unit (after re-scaling the observed values so they are each on the same scale as the latent variable). Because the average by itself does not have a real or direct monetary unit, we do not present this parameter directly.

To reiterate, once we estimate the model, we focus on the posterior predictive intervals that are each linear transformation of the latent variable. Importantly, these estimated distributions account for disagreement between data from different observed dataset values, and variation in the level of coverage within each dataset. We continue with a description of the model before turning to results.

For the model, $i = 1, \ldots, N$, indexes cross-sectional units and $t = 1, \ldots, T$, indexes time periods. For each country-year unit, $j = 1, \ldots, J$ indexes the observed variable dataset. Thus, itj indexes the observed military spending variables described above for every country-year dataset.

To connect each of the observed dataset values y_{itj}, we use a standard dynamic latent variable model setup, which is becoming more common in international relations and comparative politics research[4], and estimate separate latent traits for current and constant currency series. The dynamic model relates the estimate of the latent trait in year t to the value in the prior year $t-1$. The current latent trait $\theta_{cur[it]} \sim N(0, 1)$ for all i when $t = 1$ (the first year a country enters the dataset). When $t > 1$, the standard normal prior is centered around the latent variable estimate from the previous year such that: $\theta_{cur[it]} \sim N(\theta_{cur[it-1]}, \sigma)$. The constant latent trait is then estimated using a conversion factor, β_t, each year such that $\theta_{con[it]} = \beta_t * \theta_{cur[it]}$, making the two series comparable.

The latent traits are estimated by linking the latent variable parameters to the sets of observed arming variables for each country-year unit. All of the observed arming variables are continuous, so we use a Gaussian link function. For identification of the model, we do not estimate a slope parameter for some of the items (Fariss et al. 2022). Instead, we assume a one-unit change in the latent trait is equivalent to a one-unit change in the original observed variable.

There are two types of observed arming variables: (1) arming in current or nominal USD (not adjusted for inflation) and (2) arming in constant or real USD (adjusted for inflation). For variables not in USD, we convert these other currencies into USD before entering them into the model. We account for inflation directly in the model by altering the function that relates the latent trait to the observed values that are not adjusted for inflation. In this way, the latent trait is always in current or nominal terms, and we convert (inflate) it into real dollars by introducing a yearly slope parameter β_t for these observed dataset values. We combine all the observed data and parameters described above using the following: likelihood function:

$$L\left(\alpha,\beta,\tau,\theta|y_{itj}\right)=\prod_{i=1}^{N}\prod_{t=1}^{T}\prod_{j=1}^{J}\underbrace{N\left(\alpha_j+\theta_{con[it]},\tau_{con}\right)^{\left(v_j\right)}}_{\text{constant USD}}*\underbrace{N\left(\alpha_j+\theta_{cur[it]},\tau_{cur}\right)^{\left(1-v_j\right)}}_{\text{current USD}}$$

Recall that $\theta_{con[it]}=\beta_t{*}\theta_{cur[it]}$. We specify v_j as an indicator variable that determines which portion of the likelihood function a particular item should be passed through. For $v_j = 1$, the observed value of the *jth* indicator is in constant or real USD and does not need to be adjusted. For $v_j = 0$, the observed value of the *jth* indicator is in current or nominal USD. Table 3 shows the parameter type and prior distributions for each of the parameters that enter the latent variable model.

Results and Analysis

The latent variable model generates a normally distributed prediction interval for each of the observed dataset values y_{itj} and, importantly, the unobserved values that are included in at least one other dataset. Notationally, we denote the estimated posterior prediction intervals generated from the model as $\tilde{y}_{itj}$. The posterior prediction intervals are each a function of the value of the latent trait and the scaling parameter α_j which transforms the latent trait into the original unit of measure for y_{itj} (e.g., constant 2017, constant 2021 US dollars, etc) for each of the j datasets. Regarding uncertainty, the distribution for $\tilde{y}_{itj}$ will be larger if there are either more disagreement between observed dataset values, if there are fewer available dataset values, or both. In the subsections below, we assess how well $\tilde{y}_{itj}$ approximates y_{itj}, the conditions under which the intervals are relatively large or small, and ultimately what this entails for analysis.

Table 3. Prior Distribution for Latent Variables and Model Level Parameters.

Parameter	Prior
Country *i* latent variable estimate in first year *t* (current dollars)	$\theta_{cur[it=1]} \sim N\ (0, 1)$
Country *i* latent variable estimate in all other years (current dollars)	$\theta_{cur[it]} \sim N\ (\theta_{cur[t-1]}, \sigma)$
Country *i* latent variable estimate in any year (constant dollars)	$\theta_{con[it]} = \beta_t * \theta_{cur[it]}$
Latent variable standard deviation	$\sigma \sim HN\ (0, 1)$
Item *j* intercept (centered on the empirical mean)	$\alpha_j \sim N\ (\bar{y}_j, 4)$
Yearly slope parameter for inflation	$\beta_t \sim HN\ (0, 1)$
Model level uncertainty for constant items	$\tau_{con} \sim HN\ (0, 1)$
Model level uncertainty for current items	$\tau_{cur} \sim HN\ (0, 1)$

Note: N represents the Normal distribution. HN represents the Half Normal distribution.

Comparison of the Distribution of Observed Dataset Values and the Distribution of Posterior Predicted Intervals

We begin with an overview of the measurement model estimates generated for each variable, which are posterior prediction intervals ($\tilde{y}_{itj}$), compared to the original dataset values (y_{itj}). Figure 1 displays box and whisker plots of the distributions of posterior point estimates $E(\tilde{y}_{itj})$ (the means of the posterior prediction intervals) and observed values for all variables y_{itj} (76 total variables) in each of nine dataset projects. The top row shows variables reported in constant USD, and the bottom shows those reported in current USD. Each grouping of three box plots shows (from left to right) the distribution of observed dataset values, posterior point estimates for only country-year-units that have a corresponding observed value, and posterior point estimates for all country-year-units. This overview suggests that our measurement model is doing an excellent job recovering observed dataset values across the nine projects. However, this presentation does not yet incorporate the uncertainty estimates that the measurement model produces, which we explore next.

Uncertainty of Estimates by Country Over Time

If our latent variable model is effectively estimating the observed dataset values (y_{itj}), then observed values should fall within the estimated intervals generated from our model for most country-years ($\tilde{y}_{itj}$). To demonstrate that our model provides useful estimates for military expenditures, we compare our estimates for country-year-units

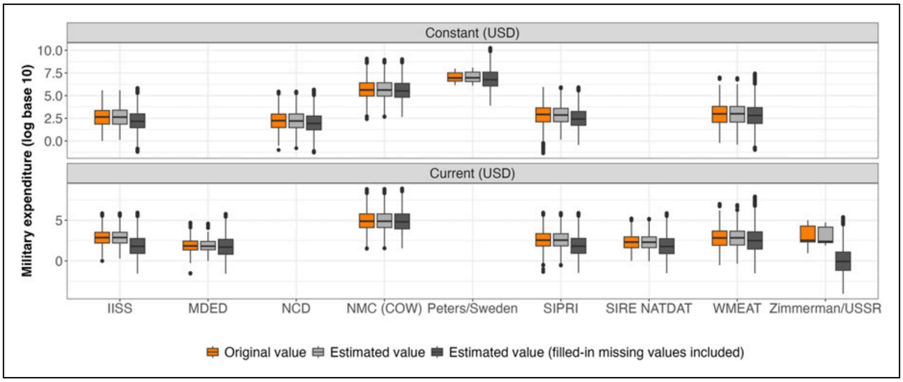

Figure 1. Agreement between observed country-year-variable-values (orange) and posterior predicted point estimates for which the observed value is observed (light grey). Dark grey boxes show the distribution of all posterior predicted point estimates (including estimates for which the original value is missing). Across all variables, these estimates have a lower median due to bias in missingness in the original data — military spending is less likely reported for country-years with very low expenditures.

with varying levels of dataset coverage. We begin with three prominent country examples that span the full time period covered by our model: the United States (Figure 2), China (Figure 3), and the United Kingdom (Figure 4).[5] These graphs display the estimated posterior prediction intervals $\tilde{y}_{itj}$ (gray lines) generated from our model for each country-year unit, and the location of the observed dataset values y_{itj} (orange points) within the estimated intervals.

In country-years where no orange points appear, there is no recorded value in the original dataset but there is information available for estimation because of the dynamic structure of the latent variable model. For country-year-units that have an observed value in at least one of the input datasets, the intervals are relatively small. The further away (in time) a unit is from an observed dataset value, the greater the uncertainty and the larger the posterior prediction interval of $\tilde{y}_{itj}$. We see this particularly in the early

Figure 2. Posterior prediction intervals (grey lines) with ±1 SD confidence bands and observed values (orange points) for China. Note substantial disagreement (large Z-scores) between posterior predictive intervals and observed values from both NMC (upper right) and WMEAT (lower right). There is much more agreement for SIPRI (lower left).

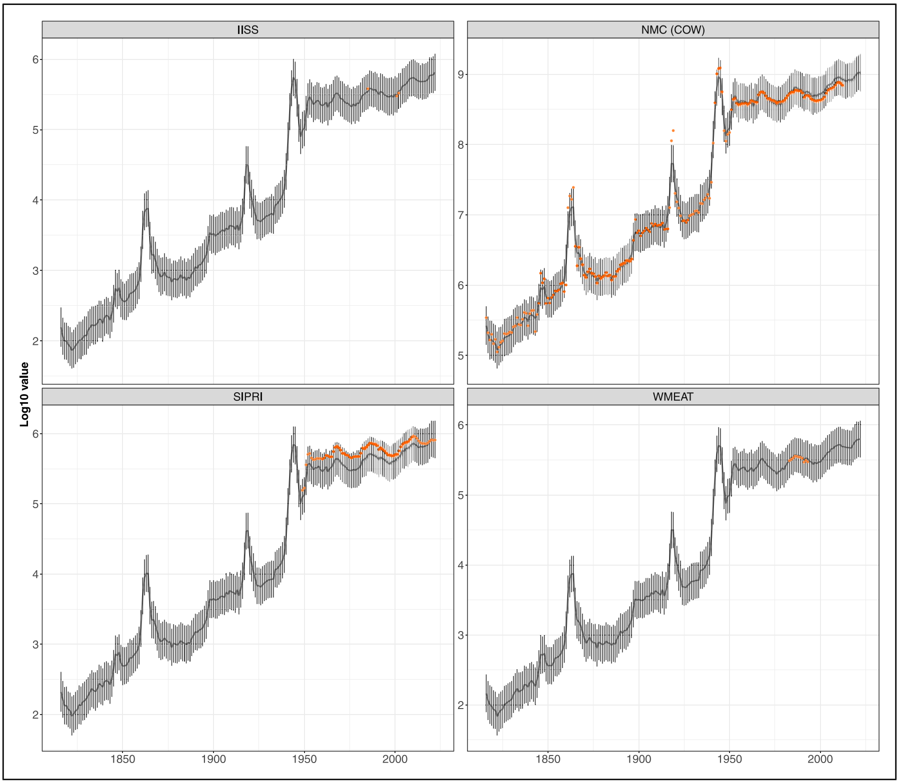

Figure 3. Posterior prediction intervals (grey) with ±1 SD confidence bands and observed values (orange) for the US. Note substantial agreement (small Z-scores) between observed dataset values and the posterior predictive intervals.

time periods when dataset coverage is sparser. In general, our estimated intervals capture the observed dataset values most of the time for these examples and additional ones in the appendix. Notice also that the trend of the posterior prediction intervals over time are identical across each panel, but that the scale of the y-axis in each panel is different. Again, this is because each of the observed dataset values uses a different unit of measurement (e.g., 2021 USD, thousands of 1990 USD, etc).

Uncertainty of Estimates by Dataset

We consider the relative level of uncertainty for units with no observed values relative to units with observed values in 1 or more datasets. In general, our estimated intervals are larger for country-year-units without any data coverage (the largest intervals are for country-year-units that are farthest away from a unit with an observed dataset value).

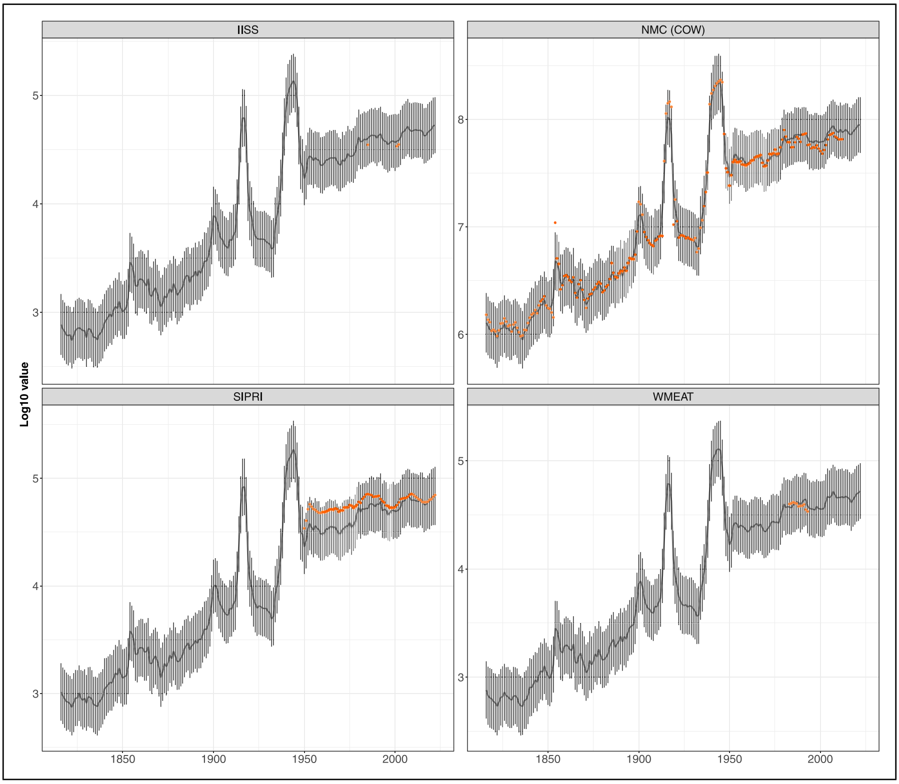

Figure 4. Posterior prediction intervals (grey) with ±1 SD confidence bands and observed values (orange) for the UK.

Conversely, these intervals decrease in size for country-year-units as additional observed dataset values become available with the caveat that the intervals will increase if the observed dataset values disagree.

If our model is doing a good job of capturing the observed dataset values, then our estimated intervals will contain the observed dataset value unless there is a high degree of disagreement. First, we consider the relative level of uncertainty for each of our country-year estimates (see Figures 2–4), then we use this information to understand how well our estimates capture the observed dataset values (see Figure 5). We do this with country-year standard deviations in the first assessment and country-year Z-scores in the second assessment (next subsection).

The left panel of Figure 5 shows that, as the amount of information (number of observed variables) for a country-year unit increases, the standard deviation of the predictive interval for that unit decreases. Because the model is dynamic, even units with zero observed dataset values have some information—the latent variable estimate

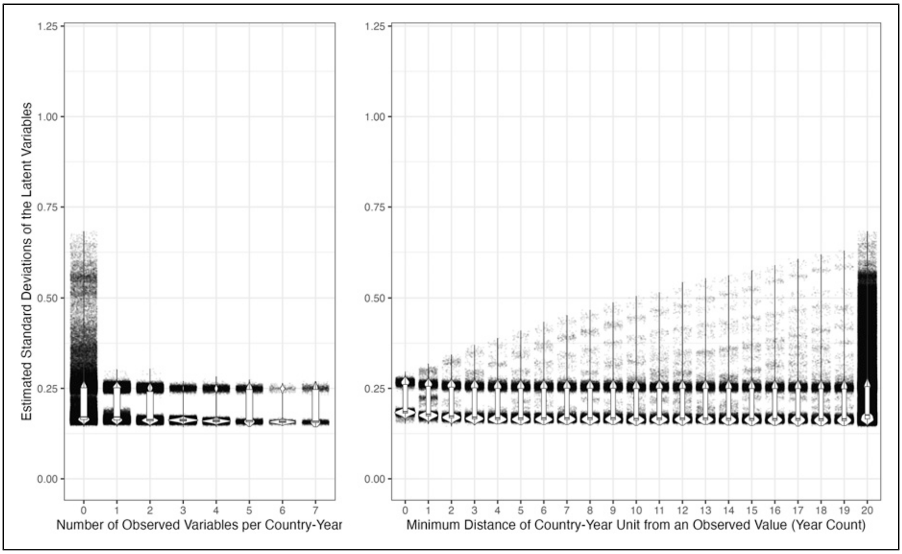

Figure 5. Relationship between the amount of information — the number of observed variables per country-year unit — and levels of uncertainty for each latent country-year estimate. As the amount of information increases, uncertainty decreases. Because the latent variable model is dynamic, even country-year-units with zero observed variables have some information — the prior year's latent variable estimate — with which to estimate a value of the latent variable. Such units have the largest standard deviations. Furthermore, the greater a country-year unit's distance from an observed variable, the larger the estimated standard deviation.

from the previous year—with which to estimate the latent variable. As seen in the left-most column, these cases have the largest standard deviations. The right panel explores additional variation in the level of uncertainty among these country-year cases with zero observed dataset values. As the distance from an observed dataset value increases, the estimated country-year unit standard deviation increases.

Validation of Predicted Intervals and Dataset Values

Next, we use the country-year standard deviations to estimate the precision of the new posterior predictions relative to the original observed variables (Gelman and Hill, 2007; Reuning et al., 2019). That is, how closely do our country-year intervals correspond to the observed dataset values and what does it mean when they do not?

First, we calculate country-year Z-scores for every observed dataset value: $z_{itj} = (y_{itj} - E(\tilde{y}_{itj}))/(\sigma_{\tilde{y}_{itj}})$. These Z-scores tell us how far away an observed dataset value y_{itj} is from the center $E(\tilde{y}_{itj})$ of the estimated country-year distribution, which is then standardized by the overall size of the country-year distribution $\sigma_{\tilde{y}_{itj}}$ for the posterior distribution of $\tilde{y}_{itj}$. We use these country-year Z-scores to tell us where the observed dataset values reside relative to the estimated country-year distributions from the latent variable model. See the appendix for tabular presentation of the results and additional discussion.

For all datasets, 88.5 percent of the observed dataset values fall within $\pm$ 1 standard deviation of the estimated country-year ranges (see Figure 5). 97.9 percent and 99.2 percent, respectively, of the observed dataset values fall within $\pm$ 2 or $\pm$ 3 standard deviations of the country-year ranges. Overall, the country-year intervals estimated from the measurement model closely cover or approximate the observed dataset values. See the appendix for more details.

There are some notable outliers however, and some datasets do better than others. Outliers are more common in the post-1948 period, where there are generally more observed dataset values per country-year unit. This is because of dataset disagreements. When observed dataset values disagree substantially, some of these observed values may fall outside the standardized ranges. For instance, compared to other datasets, there is much more disagreement for the SIRE NATDAT dataset relative to the model estimates. This is a feature of the unified measurement model, which combines information from many sources and accounts for disagreement through the quantification of measurement error.

Prior to 1948, only two datasets (three variables) provide coverage. One, from Peters (1992) only covers Sweden (see variable descriptions above). The other two are from the NMC dataset in constant and current USD. Unsurprisingly, the observed dataset values for these three variables tend to fall close to the center of the country-year distributions estimated from the latent variable model, as there are few variables and therefore few opportunities for disagreement between observed dataset values. But the posterior predictive distributions are much larger in this time period because there are so few available dataset values. As we continue to collect new historical sources of information, we expect to observe more disagreement between sources in this earlier time period, but this is a task left for future research.

Next, we demonstrate how our data can enhance understanding of the relationship between economic and military power. First, we show how the relationship between economic and military power has changed over time and discover new puzzles regarding this relationship. Second, we replicate and update Giuseppe and Poast (2018) to reveal that the effect of two key independent variables – democratic alliances and GDP – are much stronger than prior findings suggest.

First, existing theories suggest that the more severe states' conflicts of interest are, the greater the share of resources they will need to devote to arming and thus the higher their military burdens will be (Fearon 2018, 550–554). Conversely, the less severe states' conflicts of interest, the smaller the share of resources they need to devote to arming and thus the lower their military burdens will be. Scholarship suggests that countries' conflicts of interest have grown less severe over time as they become more democratic (Huth and Alee 2002; Markowitz and Fariss 2018; Nordhaus, Oneal, and Russett 2012), and are able to prosper via production and trade without needing to resort to conquest and colonialism (Brooks, 2005; Coe and Markowitz, 2021; Gartzke and Rohner, 2010; Lind, 2011; Markowitz et al., 2019, 2020). Along with the end of the Cold War and Pax Americana (Lake 2007; Wohlforth, 1999), these trends have been offered as explanations for why the world is safer and more peaceful. We know empirically that the world has grown dramatically wealthier as global surplus domestic product is roughly 3 times larger than at the end of the Cold War and that global military burdens have decreased dramatically during this same period (Anders, Fariss, and Markowitz 2020). While financial resources are not the only dimension of a state's power resources, they represent an important dimension (Beckley, 2018; Cappella Zielinski, 2016; Markowitz et al., 2024; Markowitz and Fariss, 2013; Norloff and Wohlforth, 2019) In the appendix, we show how measures of military burdens have evolved over time using the new estimates of military spending. Thus, given these trends, it would be reasonable to expect the correlation between economic and military power to be weaker in more recent years as states on average are spending a smaller share of their resources on arming.

However, as Figure 6 shows, contrary to this expectation, the correlation between economic and military power has actually become stronger over time. Figure 6 plots rank-order correlation coefficients for the relationship between military spending and Surplus Domestic Product (SDP) (Anders, Fariss, and Markowitz 2020), respectively, for 1-year, 10-year, and 50-year rolling temporal windows (similar figures showing relationships with other variables such as naval tonnage (Crisher and Souva, 2014) in the appendix). We also showcase how to incorporate uncertainty from the distribution of the arming estimates when estimating the correlation coefficients for each time-period. Specifically, we measure arming by taking $m = 100$ draws from the posterior predictive distributions for one of the arming variables, then correlating these draws with each of the other observed variables. The figures display 95 percent credible

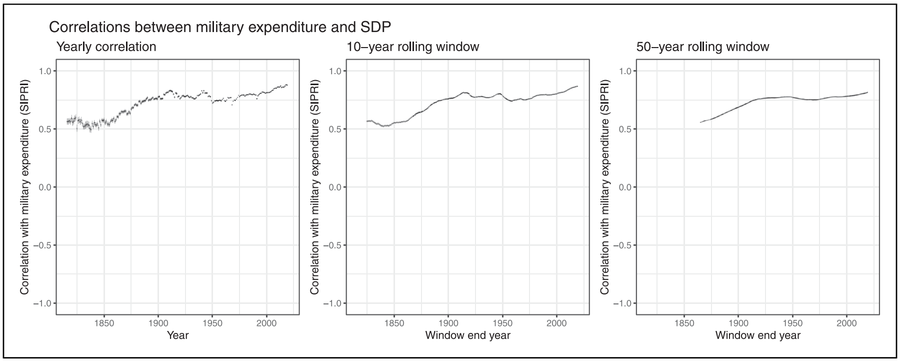

Figure 6. Distribution of Spearman rank-order correlation between the latent arming variable and SDP. Rolling correlations are calculated for 1-year, 10-year, and 50-year periods. No correlation coefficients are estimated for periods with fewer than 10 observations.

intervals generated from the distribution of 100 correlation coefficients for each rolling time period. Each of the plots illustrates considerable temporal variation in the magnitude of the correlation between these variables.

One might reasonably be concerned this relationship is driven by missing data or measurement error. Our new dataset fills in previously missing values (roughly 20 percent of all country-years since 1816) and provides credible estimates of the level of measurement error in the form of posterior prediction intervals. Thus, this evidence demonstrates the value of our data for (1) identifying new puzzles and (2) providing a means of assessing whether puzzling empirical patterns are simply the result of measurement error. Even using fairly wide uncertainty bounds, the correlation between economic and military power has generally strengthened rather than weakened over time.

Surprisingly, this correlation is stronger than at any time in the past two centuries. This was the case even before the recent increases in military spending driven by the war in Ukraine which are not yet included in our dataset because our data currently ends in 2019. So long as economic and military power remain strongly correlated, a more prosperous world will likely become a more heavily armed world. Researchers interested in arms control and global military burdens can investigate why this correlation has strengthened over time and what might cause it to weaken or strengthen in the future.

Extending the Analysis

We have demonstrated how the expanded coverage and incorporation of uncertainty influences understanding of broad international patterns. Here, we select one prominent study to showcase how to incorporate uncertainty estimates. We have replicated the main results (Models 1–4 from Table 2) from Digiuseppe and Poast (2018), using our new military expenditures data for the dependent variable, lagged DV, and to construct the measure of external threat used as a control by the authors (see Figure 7). We show that the substantive effects of democratic and nondemocratic defensive alliances are much larger (roughly 3 times as large) when using our military expenditure variable and incorporating measurement uncertainty into the model. This suggests the United States' alliance commitments may be playing an especially important role in restraining global military spending by reducing defense spending amongst its allies, who together with the U.S. represent approximately 75 percent of world GDP (Beckley 2015).

Additionally, incorporating uncertainty in measurement of the DV (as our new estimates allow) helps prevent the lagged DV from overwhelming the effects of other predictor variables, a common issue in panel data models. We see this in our replication's smaller estimated coefficients for the lag of military expenditure, and the larger estimated effect of alliances and GDP on military spending. We estimate an effect of GDP on military spending five times larger than DiGiuseppe and Poast (2018) found. This represents a major revision to our understanding of the effect of GDP on military spending, and is substantively important in light of our finding that the correlation

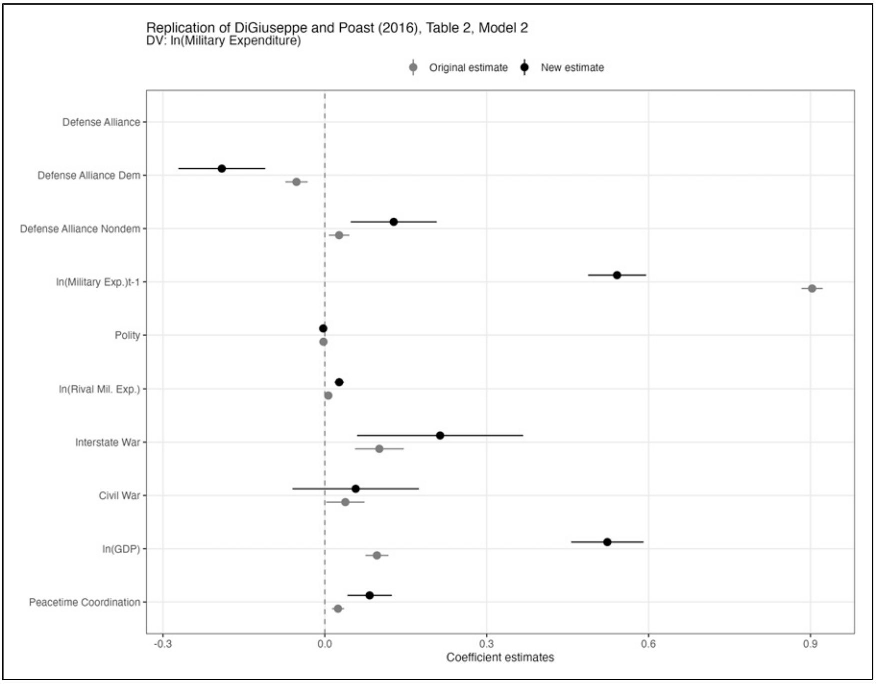

Figure 7. Coefficient estimates from replication of DiGiuseppe and Poast (2018), Table 2, Model 2. Estimate from the original paper shown in gray. Estimates using new military expenditure data from the GMSD shown in black. The GMSD's latent variable approach allows us to incorporate measurement uncertainty into the estimation process. Full regression table and coefficient plots for models 1–4 provided in the appendix.

between economic surplus and military spending is now stronger than at any time in the past two hundred years. If this trend holds, and major economies in East and South Asia continue to grow faster than European economies, then we would also expect an eastward shift in global military power. Moreover, while most European states have a powerful democratic ally, allowing them to reduce their own military spending given the credibility of U.S. security commitments, many of the largest and fastest growing economies in Asia (e.g., India, China and Indonesia) have no such ally. The same is true of smaller but rapidly growing regional economies (e.g., Vietnam and Bangladesh).

In combination with other efforts to produce new historical data, our military expenditure estimates allow scholars to extend analysis to additional time periods. While we restrict replication of the regression analysis to the time period covered by DiGiuseppe and Poast's study in order to directly compare results, we also illustrate the benefits of our new data by reproducing DiGiuseppe and Poast's Table 1 (descriptive statistics of military burdens for states with different types of allies) for additional time periods, using our expenditure estimates in combination with new historical GDP data to calculate military burdens (Fariss et al. 2022). Versions of this table for 1816–1899, 1900–1949, 1950–1999, and 2000–2019 are provided in the appendix. This analysis suggests that the core of DiGiuseppe and Poast's argument holds over the entire span of the data, but the aggregate effects differ across time periods as more countries democratize, increasing the availability of democratic allies. For example, prior to 1950, there are many fewer states with only democratic defensive allies, so the average effect of having any defensive alliance is to *increase* military burdens ($p < 0.01$ for both 1816–1899 and 1900–1949).

Conclusion

In this article, we identified and discussed the limitations of existing estimates of military spending. We then introduced a new latent variable model that estimates posterior predictions for this variable. These new estimates provide several advantages over existing variables: (1) they include multiple manifest indicators of expenditure and quantify the uncertainty for each of the resulting country-year estimates, (2) the modeling framework can be expanded to account for new data or knowledge about systematic errors in existing data, and (3) they extend the temporal and spatial coverage of existing data by estimating values that were previously missing for nearly 15 percent of all country-year observations. We close by discussing how our new data could be used in future research.

There is ongoing debate regarding whether countries with greater military capabilities are more likely to use military force. Scholars used military spending as a key measure of states' military capabilities (Fordham 2004), finding evidence that the effect of military spending is partly mitigated by force size. Spending that increases the size of the force is associated with increased use of force. However, as the U.S. military has become increasingly capital intensive, the number of military assets (i.e., troops and weapons systems) has declined, constraining the U.S. ability to use force. Our data can be used to investigate whether this relationship holds for other states.

This question is particularly relevant to policymakers, as Russian and Chinese military spending has skyrocketed over the past two decades. Meanwhile, Russia's military personnel numbers have slightly increased and China's have declined substantially, while the number and quality of both countries' power projection assets (e.g., ships and aircraft) has dramatically increased. Our data allow researchers to investigate not only the effects of increased military spending on states' propensity to use force, but also the degree to which force size, capital intensity, and power projection capabilities mitigate this relationship.

Our data allows for evaluating the relationship between economic shocks and military spending. Cappella Zielinski, Fordham, and Schilde (2017) find that "economic decline has a much larger impact on the military budget than economic growth." As the authors point out, these shocks have the potential to dramatically shift the distribution of power if some major powers experience a much deeper recession, or slower recovery, than others. This question is particularly relevant for policymakers as

COVID-19 demonstrated that major economic shocks can hit with little warning and have long lasting effects. Furthermore, trends suggest a critical and underappreciated pathway through which climate change might influence international security. Climate change generates natural disasters such as mass floods or droughts, resulting in economic shocks that disproportionately burden some countries, with the potential to produce a rapid and durable shift in the distribution of global military power.

We began this article with the observation that even though the world has become more prosperous and peaceful over the past 70 years, global military spending is now more than 3 times greater than at the height of the Cold War. This suggests a puzzle that researchers should address using our data: *If the world is more prosperous and peaceful than ever before, why are nations spending so much on arms?* More generally: *What is the relationship between wealth, war, and states' incentives to invest in weapons?* It is possible that as states grow wealthier they spend more on arming to defend those assets—Fearon (2018, 550) references Theodore Roosevelt's conjecture that "'undefended wealth invites aggression,' suggesting that rich countries need to spend more to deter invasion.

In contrast, rising global prosperity may lead to greater expectations of peace, which should lower incentives to arm. A critical first step in addressing this puzzle is to estimate what a state's baseline level of military spending will be given the size of its economy and the level of threat it faces. Surprisingly, no existing scholarship has empirically established this baseline (Fearon 2018, 550). Our data can be employed to address this gap by generating the most systematic and empirically informed predictions regarding how much a state or many/all states will spend on arming. This critical research would provide an empirical foundation to forecast future levels of military spending for rising powers such as China. It will also allow researchers to investigate whether a more prosperous world is likely to be more heavily armed or conflict prone. The new estimates we have presented have the potential to contribute to the renewed and continued interest in the political economy of security (e.g, Cappella Zielinski et al., 2021; Markowitz et al., 2019, 2020; Markowitz and Fariss, 2013; Strange, 1970).

Finally, we discussed at length how to estimate uncertainty and the importance of incorporating it into models that include military expenditure estimates. But what is associated with the variation in the uncertainty we have now estimated? This is an important substantive question. We have shown how the number of observed dataset values and the level of agreement between the dataset values determine the level of uncertainty. But there are likely substantive reasons for missingness and disagreement (e.g., Colgan 2011; Fariss, Reuning, and Kenwick 2020). These are questions we have not yet had the chance to explore but our approach and estimates are useful for such explorations. For example, if we fail to observe changes in spending during armed conflict it could be because conflict makes the act of observation more difficult. Another possibility is that governments might strategically exaggerate (either inflating or deflating) their budgets because of an incentive to misrepresent. Complicating matters, much of the information used to generate estimates of military spending is based on observations of the military hardware and personnel that a country maintains. This is especially true in earlier time periods. For example, estimates of Chinese military expenditures during the nineteenth century come from diplomatic reports based on observations made by French and English diplomats. But we know that a massive amount of spending was taking place in China during the Taiping rebellion, the largest conflict in the nineteenth century. This historical context further justifies the need to estimate uncertainty intervals for all these important cases. There are also time periods when unobserved military spending might deviate substantially from spending in observed periods of time prior to or later than these periods. Our model cannot currently estimate such swings without additional information. But, if there is additional data to suggest such deviations, it can be incorporated into the measurement model. We hope that our measurement model of military expenditure data will spur both new analyses and new measurement research.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Miriam Barnum https://orcid.org/0000-0002-2988-6648
Christopher J. Fariss https://orcid.org/0000-0001-9837-186X
Jonathan N. Markowitz https://orcid.org/0000-0001-7429-8397

Data Availability Statement

Miriam Barnum; Christopher J. Fariss; Jonathan N. Markowitz; Gaea Morales, 2024, "Replication Data for: Measuring Arms: Introducing the Global Military Spending Dataset", doi:10.7910/DVN/RKJAKJ, Harvard Dataverse.

Supplemental Material

Supplemental material for this article is available online.

Notes

1. We do not consider countries that have no observed dataset values for arming expenditures. See the appendix for a list of these countries and for the proportion of years covered for all other countries.
2. Often via COW's CINC index. For critiques of this approach, see Beckley (2018), Kadera and Sorokin (2004), and Merritt and Zinnes (1988).

3. For examples similar to the model presented here, Fariss et al. (2022) (latent GDP and population), Fariss, Kenwick, and Reuning (2020) (one-sided government killing counts).
4. See for example recent work measuring latent non-proliferation policy preferences (Barnum and Lo 2020), consolidation of power in non-democracies (Ghandi and Sumner 2020), human rights (Fariss 2019), GDP and population (Fariss, Anders, Markowitz, and Barnum 2022), civilian control of the military (Kenwick 2019), bilateral investment treaty preferences (Montal, Potz-Nielsen, and Sumner 2020), democracy (Reuning, Kenwick, and Fariss 2019), and interstate hostility (Terechshenko 2020).
5. Additional countries shown in the appendix.

References

Anders, Therese, Christopher J. Fariss, and Jonathan N. Markowitz. 2020. "Bread Before Guns or Butter: Introducing Surplus Domestic Product (SDP)." *International Studies Quarterly* 64 (2): 392-405. doi:10.1093/isq/sqaa013.

Barnum, Miriam, and James Lo. 2020. "Is the NPT Unraveling? Evidence from Text Analysis of Review Conference Statements." *Journal of Peace Research* 57 (6): 740-751. doi:10.1177/0022343320960523.

Beckley, Michael. 2015. "The Myth of Entangling Alliances: Reassessing the Security Risks of U.S. Defense Pacts." *International Security* 39 (4): 7-48. doi:10.1162/ISEC_a_00197.

Beckley, Michael. 2018. "The Power of Nations: Measuring what Matters." *International Security* 43 (2): 7-44. doi:10.1162/isec_a_00328.

Böhmelt, Tobias, and Vincenzo Bove. 2014. "Forecasting Military Expenditure." *Research & Politics* 1 (1): 2053168014535909. doi:10.1177/2053168014535909.

Bove, Vincenzo, and Roberto Nisticò. 2014. "Military in Politics and Budgetary Allocations." *Journal of Comparative Economics* 42 (4): 1065-1078. doi:10.1016/j.jce.2014.02.002.

Brooks, Steven. G. 2005. *Producing Security: Multinational Corporations, Globalization, and the Changing Calculus of Conflict*. Princeton: Princeton University Press.

Bueno de Mesquita, Bruce, Smith Alastair, Randolph M. Siverson, and James D. Morrow. 2004. "Testing Novel Implications from the Selectorate Theory of War." *World Politics* 56 (3): 363–388. doi:10.1353/wp.2004.0017.

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. doi:10.1177/0022343317715301.

Cappella Zielinski, Rosella, Kaija Schilde, and Norrin Ripsman. 2021. "A Political Economy of Global Security Approach." *Journal of Global Security Studies* 6 (1): ogaa005. doi:10.1093/jogss/ogaa005.

Carter, Jeff, Heather L. Ondercin, and Glenn Palmer. 2021. "Guns, Butter, and Growth: The Consequences of Military Spending Reconsidered." *Political Research Quarterly* 74 (1): 148-165. doi:10.1177/1065912919890417.

Crisher, Brian, and Mark Souva. 2014. "Power at Sea: A Naval Power Data Set, 1865-2011." *International Interactions* 40 (4): 602-629. doi:10.1080/03050629.2014.918039.

Coe, Andrew J. 2018. "Containing Rogues: A Theory of Asymmetric Arming." *The Journal of Politics* 80 (4): 1197-1210. doi:10.1086/698845.

Coe, Andrew J., and Jonathan N. Markowitz. 2021. "Crude Calculations: Productivity and the Profitability of Conquest." *International Organization* 75 (4): 1058-1086. doi:10.1017/S0020818321000291.

Coe, Andrew J., and Jane Vaynman. 2020. "Why Arms Control Is So Rare." *American Political Science Review* 114 (2): 342-355. doi:10.1017/S000305541900073X.

Colgan, Jeff. 2011. "Venezuela and Military Expenditure Data." *Journal of Peace Research* 48 (4): 547-556. doi:10.1177/0022343311406156.

Cusack, Thomas R., and Michael Don Ward. 1981. "Military Spending in the United States, Soviet Union, and the People's Republic of China." *Journal of Conflict Resolution* 25 (3): 429-469. doi:10.1177/002200278102500303.

Digiuseppe, Matthew, and Paul Poast. 2018. "Arms versus Democratic Allies." *British Journal of Political Science* 48 (4): 981-1003. doi:10.1017/S0007123416000247.

Downes, Alexander B. 2006. "Desperate Times, Desperate Measures: The Causes of Civilian Victimization in War." *International Security* 30 (4): 152-195.

Faber, Jan. 1992. "Annual Data on Nine Economic and Military Characteristics of 78 Nations (SIRE NATDAT), 1948-1983." ICPSR - Interuniversity Consortium for Political and Social Research 9273. doi:10.3886/ICPSR09273.v1.

Fariss, Christopher J. 2019. "Yes, Human Rights Practices Are Improving Over Time." *American Political Science Review* 113 (3): 868-881. 10.1017/S000305541900025X.

Fariss, Christopher J., Jonathan N Markowitz, Therese Anders, and Miriam Barnum. 2022. "New Estimates of Over 500 Years of Historic GDP and Population Data." *Journal of Conflict Resolution* 66 (3): 553-591. doi:10.1177/00220027211054432.

Fariss, Christopher J., Michael R. Kenwick, and Kevin Reuning. 2020. "Estimating one-Sided-Killings from a Robust Measurement Model of Human Rights." *Journal of Peace Research* 57 (6): 801–814. doi:10.1177/0022343320965670.

Fearon, James D. 2018. "Cooperation, Conflict, and the Costs of Anarchy." *International Organization* 72 (3): 523-559. doi:10.1017/S0020818318000115.

Flynn, Michael E. 2014. "Military Leadership, Institutional Change, and Priorities in Military Spending1." *Foreign Policy Analysis* 10 (2): 103-126. doi:10.1111/fpa.12005.

Fordham, Benjamin O. 2004. "A Very Sharp Sword: The Influence of Military Capabilities on American Decisions to Use Force." *Journal of Conflict Resolution* 48 (5): 632-656. doi:10.1177/0022002704267935.

Fordham, Benjamin O. 2008. "Economic Interests and Congressional Voting on American Foreign Policy." *Journal of Conflict Resolution* 52 (5): 623-640. doi:10.1017/S0898588X17000116.

Fordham, Benjamin O. 2011. "Who Wants to Be a Major Power? Explaining the Expansion of Foreign Policy Ambition." *Journal of Peace Research* 48 (5): 587-603. doi:10.1177/0022343311411959.

Fordham, Benjamin O. 2019. “The Domestic Politics of World Power: Explaining Debates over the United States Battleship Fleet, 1890-91.” *International Organization* 73 (2): 435-468. doi:10.1017/S0020818318000449.

Fordham, Benjamin O., and Thomas C. Walker. 2005. “Kantian Liberalism, Regime Type, and Military Resource Allocation: Do Democracies Spend Less?” *International Studies Quarterly* 49 (1): 141-157. doi:10.1111/j.0020-8833.2005.00338.x.

Gandhi, Jennifer, and Jane Lawrence Sumner. 2020. “Measuring the Consolidation of Power in Nondemocracies.” *The Journal of Politics* 82 (4): 1545–1558.

Gartzke, Erik, and Dominic Rohner. 2010. “To Conquer or Compel: War, Peace, and Economic Development.” Institute for Empirical Research in Economics University of Zurich Working Paper 511. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1684000.

Gelman, Andrew, and Jennifer Hill. 2007. “Data Analysis Using Regression and Multilevel/Hierarchical Models.” Cambridge MA: Cambridge University Press.

Gillespie, John, and Dina Zinnes. 1992. “Military Defense Expenditure Data, 1948-1970.” ICPSR - Inter-university Consortium for Political and Social Research 5007. doi:10.3886/ICPSR05007.v1.

Goldsmith, Benjamin E. 2003. “Bearing the Defense Burden, 1886-1989: Why Spend More?” *Journal of Conflict Resolution* 47 (5): 551-573. doi:10.1177/0022002703254297.

Gowa, Joanne, and Edward D. Mansfield. 1993. “Power Politics and International Trade.” *American Political Science Review* 87 (2): 408-420. doi:10.2307/2939050.

Greig, J. Michael, and Andrew Enterline J. 2017. “National Material Capabilities (NMC) Data Documentation Version 5.0.” *Correlates of War Project*.

Greig, J. Michael, and Andrew J. Enterline. 2021. “National Material Capabilities (NMC) Data Documentation Version 6.0.” Correlates of War Project.

Huth, Paul K., and Todd L. Allee. 2002. “Domestic Political Accountability and the Escalation and Settlement of International Disputes.” *Journal of Conflict Resolution* 46 (6): 754–790. doi:10.1177/002200202237928.

Huth, Paul D., Scott Bennett, and Christopher Gelpi. 1992. “System Uncertainty, Risk Propensity, and International Conflict Among the Great Powers.” *The Journal of Conflict Resolution* 36 (3): 478-517. doi:10.1177/0022002792036003004.

IISS. 2022. *The Military Balance 2022*. London: Routledge.

Jenkins, J. Craig, Stephen J. Scanlan, and Lindsey Peterson. 2007. “Military Famine, Human Rights, and Child Hunger.” *Journal of Conflict Resolution* 51 (6): 823-847. doi:10.1177/0022002707308215.

Kadera, M. Kelly, and Gerald L. Sorokin. 2004. “Measuring National Power.” *International Interactions* 30 (3): 211-230. doi:10.1080/03050620490492097.

Kenwick, Michael R. 2019. “Self-Reinforcing Civilian Control: A Measurement-Based Analysis of Civil-Military.” *International Studies Quarterly* 64 (1): 71-84. doi:10.1093/isq/sqz092.

Kleinberg, Katja B., and Benjamin O. Fordham. 2013. “The Domestic Politics of Trade and Conflict.” *International Studies Quarterly* 57 (3): 605-619. doi:10.1111/isqu.12016.

Lake, David A. 2007. “Escape from the State of Nature: Authority and Hierarchy in World Politics.” *International Security* 32 (1): 47-79. doi:10.1162/isec.2007.32.1.47.

Lebovic, James H. 1995. "How Organizations Learn: U.S. Government Estimates of Foreign Military Spending." *American Journal of Political Science* 39 (4): 835-863. doi:10.2307/2111659.

Lebovic, James H. 2001. "Spending Priorities and Democratic Rule in Latin America." *Journal of Conflict Resolution* 45 (4): 427-452. doi:10.1177/0022002701045004002.

Lind, Jennifer. 2011. "Democratization and Stability in East Asia." *International Studies Quarterly* 55 (2): 409-436. doi:10.1111/j.1468-2478.2011.00652.x.

MacDonald, Paul K., and Joseph M. Parent. 2011. "Graceful Decline? The Surprising Success of Great Power Retrenchment." *International Security* 35 (4): 7-44. doi:10.1162/ISEC_a_00034.

Markowitz, Jonathan N., Miriam Barnum, and Christopher Fariss J. 2024. "Measuring Threat in the International System." *OSF working paper*. https://osf.io/q6a2h/.

Markowitz, Jonathan N, and Christopher Fariss J. 2013. "Going the Distance: The Price of Projecting Power." *International Interactions* 39 (2): 119–143. doi:10.1080/03050629.2013.768458.

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. doi:10.1177/0022343317727328.

Markowitz, Jonathan N, Christopher Fariss J., and Blake McMahon. 2019. "Producing Goods and Projecting Power: How What You Make Influences What You Take." *Journal of Conflict Resoultion*. doi:10.1177/0022002718789735.

Markowitz, Jonathan, Suzie Mulesky, Benjamin A. T. Graham, and Christopher J. Fariss. 2020. "Productive Pacifists: The Rise of Production-Oriented States and Decline of Profit-Motivated Conquest." *International Studies Quarterly* 64 (3): 558-572. doi:10.1093/isq/sqaa045.

Merritt, Richard L., and Dina A. Zinnes. 1988 *Validity of Power Indices" International Interactions* 14 (2): 141-151. doi:10.1080/03050628808434697.

Mohammed, Nadir A.L. 1996. "What Determines Military Allocations in Africa: Theoretical and Empirical Investigations." *Defence and Peace Economics* 7 (3): 203-231. doi:10.1080/10430719608404853.

Montal, Florencia, Carly Potz-Nielsen, and Jane Lawrence Sumner. 2020. "What States Want: Estimating Ideal Points from International Investment Treaty Content." *Journal of Peace Research* 57 (6): 679-691.

Morrow, James D. 1997. "When Do 'Relative Gains' Impede Trade?" *The Journal of Conflict Resolution* 41 (1): 12-37. doi:10.1177/0022002797041001002.

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. doi:10.1017/S0020818312000173.

Norloff, Carla, and William Wohlforth. 2019. "Raison De L'hégémonie (The Hegemon's Interest): Theory of the Costs and Benefits of Hegemony." *Security Studies* 28 (3): 422-450. doi:10.1080/09636412.2019.1604982.

Oneal, John R., and Hugh Carter Whatley. 1996. "The Effect of Alliance Membership on National Defense Burdens, 1953–88: A Test of Mancur Olson's Theory of Collective Action." *International Interactions* 22 (2): 105-122. doi:10.1080/03050629608434884.

Ostrom, Charles. 1978. "A reactive linkage model of the US defense expenditure policymaking process." *American Political Science Review* 941–957.

Peters, B. Guy. 1992. "Political Systems Performance Data: Sweden, 1865-1967: Version 1." *ICPSR - Interuniversity Consortium for Political and Social Research* 36. doi:10.3886/ICPSR00036.V1.

Powell, Robert. 1993. "Guns, Butter, and Anarchy." *The American Political Science Review* 87 (1): 115-132. doi:10.2307/2938960.

Reiter, Dan, and Allan Stam. 1998. "Democracy, War Initiation, and Victory." *American Political Science Review* 92 (2): 377–389. doi:10.2307/2585670.

Reuning, Kevin, Michael Kenwick R., and Christopher Fariss J. 2019. "Exploring the Dynamics of Latent Variable Models." *Political Analysis* 27 (4): 503–517.

Rider, J Toby. 2013. "Uncertainty, Salient Stakes, and the Causes of Conventional Arms Races." *International Studies Quarterly* 57 (3): 580-591. doi:10.1111/isqu.12058.

Sample, Susan G. 2018. "Anticipating War? War Preparations and the Steps-To-War Thesis." *British Journal of Political Science* 48 (2): 489-511. doi:10.1017/S000712341500068X.

Singer, J. David. 1972. "The "Correlates of War" Project: Interim Report and Rationale." *World Politics* 24 (2): 243-270. doi:10.2307/2009738.

Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985." *International Interactions* 14 (2): 115-132. doi:10.1080/03050628808434695.

Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." In *Peace, War, and Numbers*, edited by Bruce Russett, 19-48. Beverly Hills: Sage.

Strange, Susan. 1970. "International Economics and International Relations: A Case of Mutual Neglect." *International Affairs* 46 (2): 304-315. doi:10.2307/2613829.

Sullivan, Patricia L. 2007. "War Aims and War Outcomes: Why Powerful States Lose Limited Wars." *Journal of Conflict Resolution* 51 (3): 496-524.

Taylor, Charles Lewis, and Joachim Amm. 2006. "National Capability Data: Annual Series, 1950-1988: Version 1." ICPSR - Interuniversity Consortium for Political and Social Research 9904. doi:10.3886/ICPSR09904.V1.

Terechshenko, Zhanna. 2020. "Hot under the Collar: A Latent Measure of Interstate Hostility." *Journal of Peace Research* 57 (6): 764-776. doi:10.1177/0022343320962546.

Wohlforth, William C. 1999. "The Stability of a Unipolar World." *International Security* 24 (1): 5–41.

Wohlforth, William C. 2009. "Unipolarity, Status Competition, and Great Power War." *World Politics* 61 (1): 28-57. doi:10.1017/S0043887109000021.

Zimmerman, William. 1985. *Words and Deeds in Soviet Military Expenditures, 1955-1983*. Ann Arbor, MI.: ICPSR Inter-university Consortium for Political and Social Research. 8188. doi: 10.3886/ICPSR08188.v1.

Zimmerman, William, and Glenn Palmer. 1983. "Words and Deeds in Soviet Foreign Policy: The Case of Soviet Military Expenditures." *The American Political Science Review* 77 (2): 358-367. doi:10.2307/1958921.