An Actuary Walks into a War Room: Using a Model from Actuarial Science to Predict War Size

Miles D. Williams[[1]](#footnote-20)

2025-03-27

Abstract

**Abstract**: This document was produced using a Quarto (.qmd) template. It’s goal is to provide a better looking and functioning Word template for researchers to use when weaving together reports with their text and embedded code in Quarto.

***Keywords***: Conflict, Battle Deaths, Inverse Burr

# Introduction

Risk is the business of actuaries, and when it comes to war business is good. Actuaries typically work for insurance companies or firms, and their job is to model financial risk and ruin. That means they care about more than a single point estimate of the expected value of, say, an insurance claim. They care about the probability distribution over the possible claims the insured might make. Often, these claims follow a thick tailed distribution where the likelihood of extreme values is higher than you would see with more conventional probability density functions. Because insurance companies and firms want to avoid financial ruin, they take the odds of extreme events seriously and want their actuaries to use the most rigorous methods possible for modeling this risk.

When Cunen, Hjort, and Nygård (2020) recently used the inverse Burr distribution to model the sizes of international wars, they drew from the deep well of actuarial science. They even used a statistical package created specifically for the purpose of doing actuarial science – an R package called {actuar} – to apply the inverse Burr to war fatality data. What does it mean to approach the study of war through the lens of an actuary? What theoretical framework and assumptions are implied? Cunen, Hjort, and Nygård (2020) say very little about the actuarial roots of their methodology. This is a missed opportunity to showcase the potential on offer with Cunen et al.’s approach – as well as the pitfalls.

This study is motivated by the goal of better contextualizing the use of the inverse Burr for studying the risk of deadly wars and providing guidance and resources for researchers interested in implementing this method. The study proceeds in three parts.

First, the problem Cunen, Hjort, and Nygård (2020) turned to actuarial science to solve – the heavy-tailed distribution of war fatalities – is discussed. This problem has posed a technical challenge and, thus, been a source of contention among scholars interested in testing macro trends in war’s deadliness. Cunen et al.’s work is a recent contribution to a small but growing literature centered on the decline of war theory. Much of this literature has, in part, developed in response to the best-selling book, *The Better Angels of Our Nautre,* by Steven Pinker (2011). The merits of assessing the risk of deadly wars are self-evident, and by extension, so are the merits of micro-founding and explaining methods used to make this assessment.

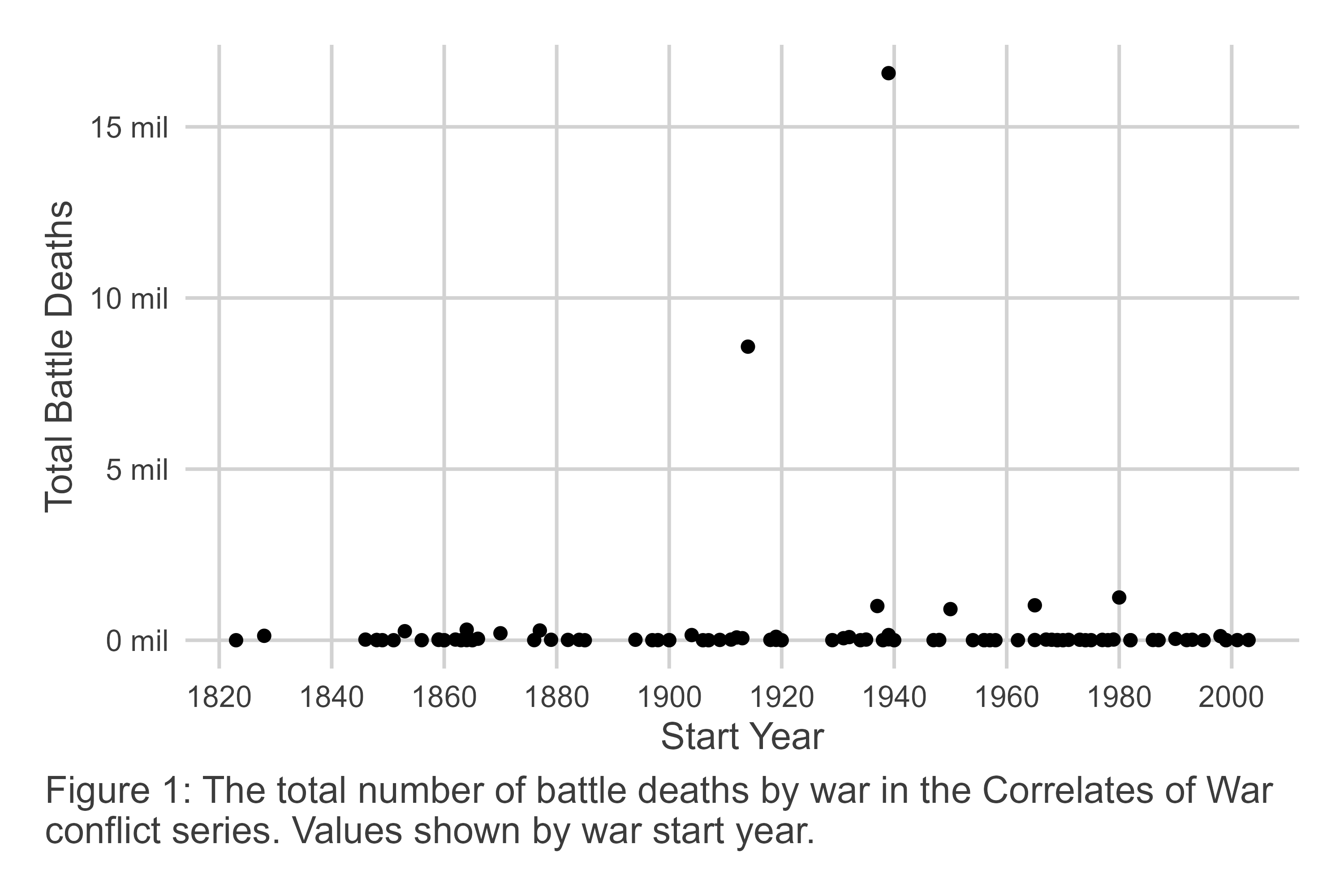
Second, readers are introduced to the inverse Burr model and how it can be used in conflict research. When Cunen, Hjort, and Nygård (2020) applied the model to the problem of modeling war size, they did not just borrow it from actuarial science, they also innovated by showing that the model could be parameterized with covariates. However, their novel approach lacks a consistent and easily accessible user interface for researchers, or a full explanation of what this actuarial tool implies for this methods appropriate use and evaluation. To facilitate better accessibility, a new statistical package developed in the R programming language is introduced that allows users to execute an inverse Burr regression model and summarize the results in ways consistent with standard modeling functions in the R language. However, because of the unique purpose for which the inverse Burr was intended, assessing its output requires a different approach than most political scientists are accustomed to. The model is ideal, not for generating a unique prediction for war size, but a conditional distribution of possible war sizes.

As the world is witnessing an upsurge in conflicts, the need for appropriate analytical tools and the right intuitions for how to use them is timely and necessary. Quantifying the risk of conflict escalation is important not just to academics, but to policy-makers as well. Cunen, Hjort, and Nygård (2020) introduced a helpful tool to conflict research for achieving this goal, but for it to be of greatest utility to the field, it is essential that researchers have guidance on how to use it and convenient tools for doing so.

# War, the Power-law, and the Alternative

One of the foundational questions in the quantitative study of war centers on variation in war deaths. Why are some wars short-lived and result in only a few thousand fatalities, while a others last years (even decades) and kill millions? (Richardson 1948, 1960) was among the first to approach this question using data. He was preceded in this effort by data collection done by Wright and Wright (1942), but Richardson was the first to bring “real statistical acumen,” as Braumoeller (2021, 274) put it, to the data Wright produced. One of the key observations Richardson made was the fact that many wars are quite small while only a few become especially deadly. In his 1948 work, Richardson noted that the spread of war deaths displays power-law like behavior, meaning that the wars that are the most deadly are exceptionally huge.

This observation is preserved in more recent and higher quality datasets such as the war series produced by the Correlates of War (CoW) project (Sarkees and Wayman 2010). Figure 1 shows the variation in total battle deaths recorded per war in the CoW data by the first year the war started. The data documents 95 international wars from 1816 to 2007, and, just like Richardson observed decades ago, this more recent dataset shows that most of the wars fought during the past two centuries were generally small. But a few were enormous. World War II, the deadliest war in the CoW dataset, had over 16 million recorded battle deaths. This makes up about 51.8% of all battle deaths in the data. World War I and World War II combined make up 78.6% of all battle deaths, and the top 10 deadliest wars account for nearly 95% of all battle deaths. This kind of distribution is far more extreme than a typical Pareto distribution where 20% of observations are responsible for 80% of outcomes.



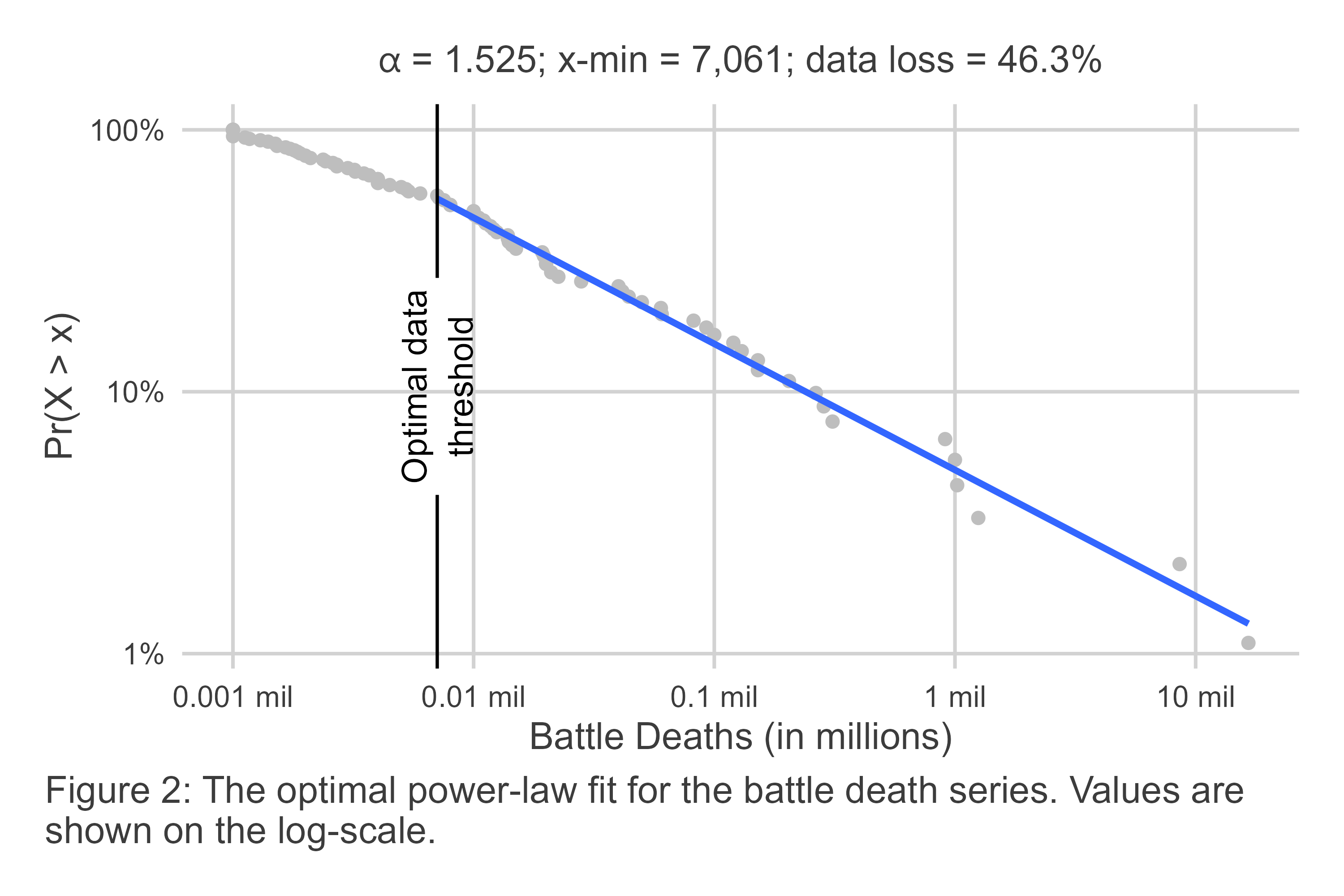
Data that is prone to the kind of massive up-swings in magnitude seen in war require specialized statistical tools to analyze. While methods differ a bit, most recent studies use the classic power-law model to study variation in war size (Braumoeller 2019; Cederman 2003; Cederman, Warren, and Sornette 2011; Cirillo and Taleb 2016; Clauset 2017, 2018; Spagat and Weezel 2020; Spagat, Johnson, and Weezel 2018). The model’s name takes inspiration from its functional form, which specifies that the likelihood of seeing an event greater than size is inversely proportional to the size of the event raised to the power :

The parameter , what can also be called the power-law slope, determines the likelihood of extreme events. Bigger values mean the likelihood of extreme events is lower, while smaller values mean the likelihood of extreme events is higher.

In practice, the power-law form has some unique behaviors and poses some challenges for fitting it to real world data. First, in terms of unique behaviors, one of the most consequential is that if the parameter is less than or equal to 3, finite variance for the variable cannot be guaranteed, and if is less than or equal 2, a finite mean cannot be identified. In short, depending on how extreme the distribution of the variable is, a power-law fit for the data may yield a result that says the expected magnitude of the variable is statistically indistinguishable from infinity. This issue is relevant for war. Many recent studies estimate s with values less than 2, and some scholars, like Braumoeller (2019), infer from this that extinction-level wars are not just possible but probable, with a higher likelihood than many assume.

Second, in terms of practical challenges, most real-world data are not uniformly explained by the power-law. This fact is noted by Clauset, Shalizi, and Newman (2009) who recommend identifying a minimum value of such that the data is power-law distributed for all . While this seems like a compromise, Clauset, Young, and Gleditsch (2007) argue that in many studies, power-law behavior in the extreme tail of a thick-tailed distribution is of greatest concern to researchers. Some data truncation is a necessary trade-off for ensuring an accurate fit for the power-law model for the largest events in the data.

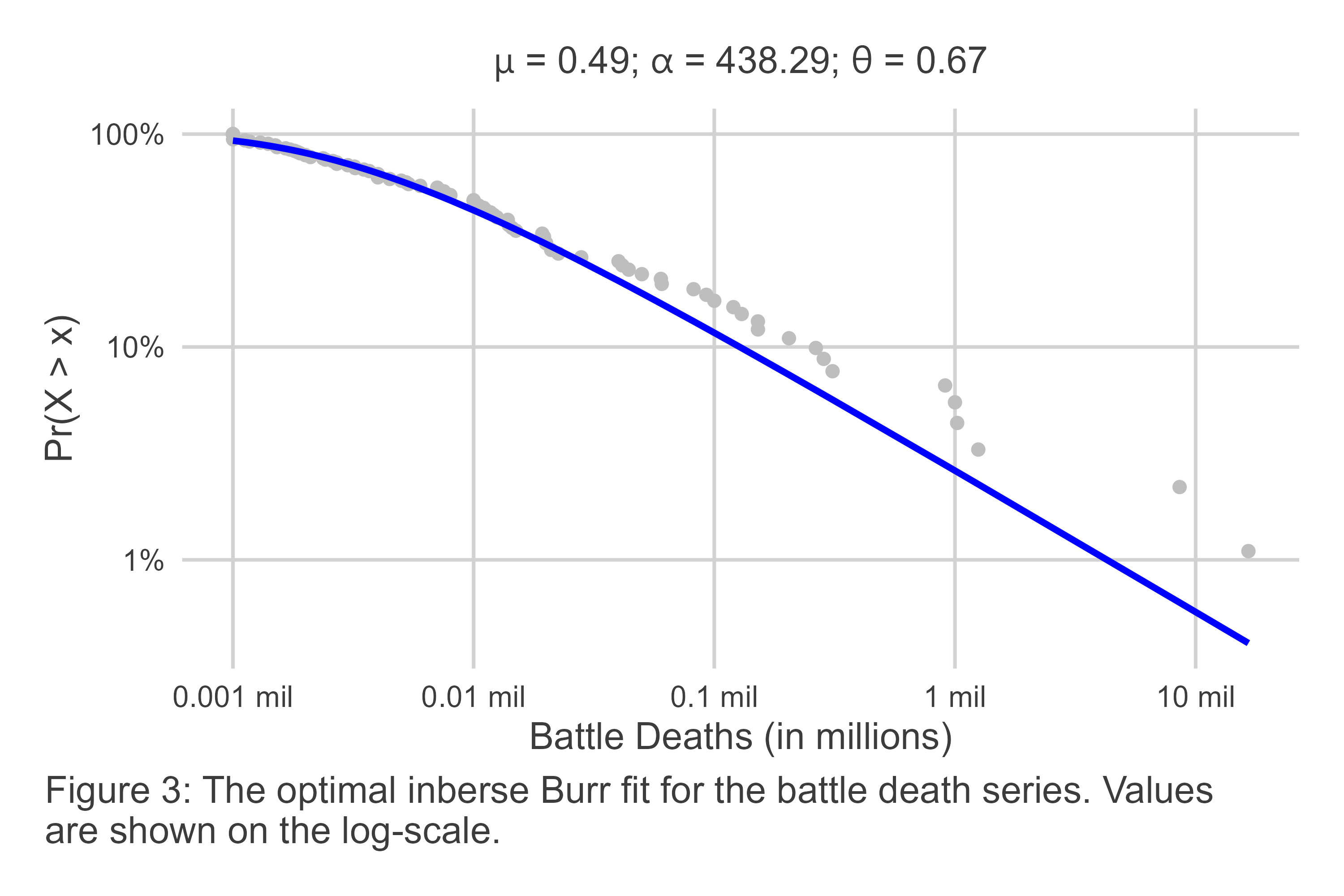
This practical concern is also relevant for studying war. Using the CoW dataset used to produce Figure 1, Figure 2 shows the optimal fit for a power-law model to the battle death series. Consistent with best-practice, the appropriateness of the power-law fit for data is shown using a log-log plot where the observed values of battle deaths are shown on the x-axis and the empirical CDF, , is shown on the y-axis. Based on the form of the power-law model, in a log-log plot you should see a negative linear correlation between and . Sure enough, this is what we observe in the data. This is with the caveat, however, that we observe power-law behavior for all large . Specifically, the parameters for the power-law model are and . While the power-law slope seems to make for a good linear fit for more extreme wars and the size of the minimum war size where the power-law applies is small, because of the extreme distribution in war deaths nearly a majority of the wars in the data *cannot be explained* by the model (46.3% fall below ).



Most scholars who have used the power-law to study variation in the size of war accept data loss as a reasonable trade-off for the ability to accurately fit the power-law for extreme conflicts. However, more recent work by Cunen, Hjort, and Nygård (2020) raises the concern that the conventional power-law model leaves substantial variation in war unexplained. Further, by removing observations, they worry that statistical precision is being needlessly sacrificed as well.

This loss in precision has special relevance for scholars studying war. Much of the recent up-surge in interest in examining macro trends in war size is driven in reaction to the best-selling book, *Better Angels of Our Nature*, by Pinker (2011). One of Pinker’s claims is that international wars have become less deadly over time. This argument, known as the decline of war thesis, is not original to Pinker (see Gaddis 1986), and a variety of studies have been published in the ensuing years that attempt to apply various statistical tests to assess its veracity. Most of these tests involve using the power-law. Some prominent work in this vein fails to find a statistically significant difference in power-law models fit before and after World War II (Braumoeller 2019) – the cut-point most scholars believe begets a decline in war’s likelihood of occurrence and deadliness due to a rise in international institutions, the liberal international order, and nuclear deterence. But Cunen, Hjort, and Nygård (2020), who use an alternative model known as the inverse Burr, do find evidence of a statistically detectable change in the chances of deadly wars in the second half of the 20th century. Unlike work that relies on the classic power-law, their approach allows them to use all of the data in their conflict series to draw this conclusion.

To demonstrate, Figure 3 shows the same data as presented in Figure 2. The difference is that the inverse Burr model suggested by Cunen, Hjort, and Nygård (2020) as an alternative for the power-law is fit to the data. The inverse Burr is more complex than the classic power-law, allowing it to accommodate power-law behavior in the extreme tail of the distribution while also explaining variation in smaller events that fall below the threshold required to fit the classic power-law to data.



The inverse Burr model has three parameters: a scale parameter , and two shape parameters, and . A good way to think about these parameters is that captures the central tendency of the data, captures the density of the data leading up to its central tendency, and captures the density of extreme values to the right of the data’s central tendency. Confusingly, the last parameter, , is approximately analogous to the power-law slope .

The as specified by the inverse Burr is given as:

where , , , and are all strictly positive. For very large , the fact that captures power-law behavior can be seen by the following approximate specification:

This functional form offers more flexibility, which is on display in Figure 3. Rather than being restricted to fitting a linear relationship between and in log-log space, the inverse Burr can fit a quasi-concave and monotone curve, allowing it to capture both small values of as well as very large values in a thick-tailed distribution.

While Cunen, Hjort, and Nygård (2020) tout the inverse Burr as an unproblematic alternative to the power-law, it under-predicts the likelihood of the most extreme wars in the CoW battle death series – a fact that is clearly visible in Figure 3. This is a point that others have raised (Braumoeller 2021), but it is premature to dismiss the inverse Burr simply for this reason. One additional promised benefit of the inverse Burr as opposed to the classic power-law is its ability to model the distribution of using covariates.

In their study, Cunen, Hjort, and Nygård (2020) showcase this strength by parameterizing their inverse Burr model using both a binary indicator for the time period (before or after a particular cut-point) for the and parameters, as well as continuous covariate (the average polity score among countries fighting a war) that is part of the parameter as well. In particular, they specify that:

along with , and a constant . The value represents average polity scores, and the parameters capture how democracy scores differently predict the central tendency before and after the specified cut-point.

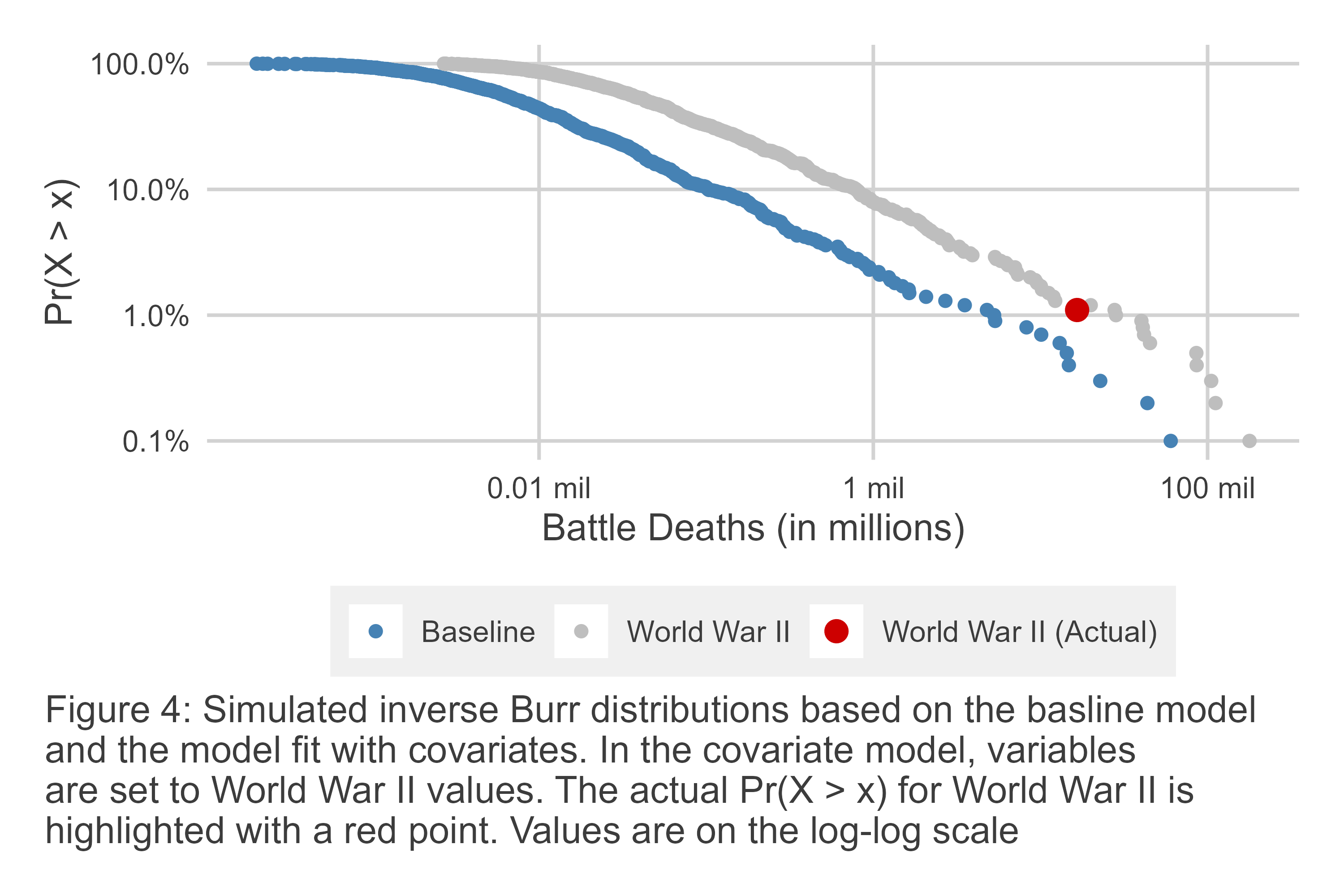
The ability to incorporate covariates into the estimation of an inverse Burr model potentially off-sets its inferior performance in the extreme tail of the CoW conflict series. As Cunen, Hjort, and Nygård (2020) note, their parameterized model offers a superior fit for the data relative to a basic inverse Burr model that assumes constant , , and .

The value of this approach goes beyond providing a better fit with greater statistical precision. Cunen, Hjort, and Nygård (2020) are able to offer affirmative evidence consistent with the decline of war theory, and provide substantive insight into the role of democracy in explaining variation in war size. First, using an innovative change-point algorithm, Cunen, Hjort, and Nygård (2020) identify 1950 as the best fitting cut-point in the CoW series. At that cut-point, the difference in inverse Burr fits before and after are statistically distinguishable, with the post-1950 data consistent with a decline in the chance of particularly large wars.

Second, Cunen, Hjort, and Nygård (2020) show that democracy has little correspondence with war size before 1950, but that democracy predicts smaller wars after 1950. This finding applies specifically to the central tendency , and means that when the average democracy score among countries fighting a war increases, the scale of wars declines.

Cunen et al.’s finding with respect to democracy is truly unique and novel. But the real gift their approach offers to conflict scholars is their methodology for identifying this relationship. The ability to model war size, not just as a stationary series, or by fitting separate models to discrete periods, offers researchers greater flexibility in identifying more nuanced relationships between factors of interest and the escalatory potential of war.

This, however, is where their study ends and the contribution of this one begins. If a tool like the inverse Burr model is to have more wide-spread acceptance among conflict scholars interested in testing theories about war size, more guidance is required than Cunen, Hjort, and Nygård (2020) offer in their paper. While they do have an online appendix and make their R code available for implementing their method, Cunen et al. offer little in the way of practical guidance or a convenient user interface for applied researchers. This study is motivated by the desire to fill this gap by offering these resources for conflict scholars. This is the subject of the next section.



Braumoeller, Bear F. 2019. *Only the Dead: The Persistence of War in the Modern Age*. New York: Oxford University Press.

———. 2021. “Trends in Interstate Conflict.” In *What Do We Know about War? Third Edition*, edited by Sara McLaughlin Mitchell and John A. Vasquez. Rowman; Littlefield.

Cederman, Lars-Erik. 2003. “Modeling the Size of Wars: From Billiard Balls to Sandpiles.” *American Political Science Review* 97 (1): 135–59.

Cederman, Lars-Erik, T. Camber Warren, and Didier Sornette. 2011. “Testing Clausewitz: Nationalism, Mass Mobilization, and the Severity of War.” *International Organization* 65 (4): 605–38.

Cirillo, Pasquale, and Nassim Nicholas Taleb. 2016. “On the Statistical Properties and Tail Risk of Violent Conflicts.” *Physica A: Statistical Mechanics and Its Applications* 452: 29–45.

Clauset, Aaron. 2017. “The Enduring Threat of a Large Interstate War.” Technical report. One Earth Foundation.

———. 2018. “Trends and Fluctuations in the Severity of Interstate Wars.” *Science Advances* 4 (2): eaao3580.

Clauset, Aaron, Cosma Rohilla Shalizi, and M.E.J. Newman. 2009. “Power-Law Distributions in Empirical Data.” *SIAM Review* 51 (4): 661–703.

Clauset, Aaron, Maxwell Young, and Kristian Skrede Gleditsch. 2007. “On the Frequency of Severe Terrorist Events.” *Journal of Conflict Resolution* 51 (1): 58–87.

Cunen, Céline, Nils Lid Hjort, and Håvard Mokleiv Nygård. 2020. “Statistical Sightings of Better Angels: Analysing the Distribution of Battle-Deaths in Interstate Conflict over Time.” *Journal of Peace Research* 57 (2): 221–34.

Gaddis, John Lewis. 1986. “The Long Peace: Elements of Stability in the Postwar International System.” *International Security* 10 (4): 99–142.

Pinker, Steven. 2011. *The Better Angels of Our Nature: Why Violence Has Declined*. New York: Viking Adult.

Richardson, Lewis F. 1948. “Variation of the Frequency of Fatal Quarels with Magnitude.” *American Statistical Association* 43: 523–46.

———. 1960. *Statistics of Deadly Quarels*. Chicago: Quadrangle Books.

Sarkees, Meredith Reid, and Frank Wayman. 2010. *Resort to War: 1816 - 2007*. Washington DC: CQ Press.

Spagat, Michael, Neil F Johnson, and Stijn van Weezel. 2018. “Fundamental Patterns and Predictions of Event Size Distributions in Modern Wars and Terrorist Campaigns.” *PloS One* 13 (10): e0204639.

Spagat, Michael, and Stijn van Weezel. 2020. “The Decline of War Since 1950: New Evidence.” In *Lewis Fry Richardson: His Intellectual Legacy and Influence in the Social Sciences*, edited by Nils Peter Gleditsch, 129–42. Springer.

Wright, Quincy, and Louise Leonard Wright. 1942. *A Study of War*. University of Chicago Press.

1. Denison University, Granville, OH USA. Contact at mdwilliams@denison.edu. [↑](#footnote-ref-20)