4

Lessons from Project Mars

QUANTITATIVE TESTS OF MILITARY INEQUALITY AND BATTLEFIELD PERFORMANCE SINCE 1800

Every military fact is also a social and political one.

ANTONIO GRAMSCI, *THE PRISON NOTEBOOKS*, 1929–35

THE MAHDIYA'S METEORIC rise and shocking collapse lend initial support to the notion that military inequality is bad for business. This chapter takes up the challenge of determining whether this stark pattern generalizes to the wider universe of belligerents and conventional wars fought since 1800. Drawing on Project Mars, this chapter subjects the proposed argument to an increasingly severe battery of statistical analyses to test its robustness across five different measures of battlefield performance. Inspired by the book's conceptualization of battlefield performance as combining elements of combat power and cohesion, these measures include a belligerent's loss-exchange ratio, the incidence of mass desertion and mass defection among its soldiers, and the fielding of blocking detachments designed to coerce reluctant soldiers to fight. A composite index of battlefield performance rounds out these measures. The intuition here is a simple one: our confidence that military inequality is a central driver of battlefield performance is earned by its ability to explain coherence across several different, and imperfectly correlated, behavioral indicators across time and diverse contexts.[1]

The chapter's empirical strategy unfolds over five stages. First, I detail how core variables, including prewar military inequality and battlefield

1. Seeking coherence across multiple outcomes and comparison groups can also reduce the sensitivity of findings arising from the non-random nature of military inequality. On this point, see Rosenbaum 2010, 118–20, 339.

performance, were constructed. Alternative explanations, including the relative balance of troops, regime type, and the military's own organizational design, are also detailed, alongside key control variables. Second, I provide simple descriptive statistics about the relationship between military inequality and the various measures of battlefield performance. Third, I subject two different measures of military inequality to a series of statistical tests to examine the strength of this association. I forego long, cumbersome, statistical tables and instead rely on simple figures to convey the substantive importance of military inequality. These tests also break with established practices in quantitative studies of military effectiveness by dividing the 1800–2011 time period into early modern (1800–1917) and modern (1918–) eras. I therefore test for military inequality's effects within these eras rather than pooling observations across them, a move that creates an additional hurdle for the proposed argument but one that better respects the contextual differences in warfare over time. Next, I use matching and a two-control group comparison to investigate how shifts from low to high inequality, and from medium to high inequality, affect battlefield performance within a reduced sample of most similar belligerents. This approach allows us to examine the effects of sudden increases in the "dosage" of inequality while screening out potential differences between belligerents, thereby isolating the effects of military inequality. A final section considers the fate of leading alternative explanations, which receive almost no empirical support within the expanded Project Mars universe.

To preview the chapter's findings, I find considerable support for the claim that prewar military inequality adversely affects battlefield performance across both eras. As inequality increases, the predicted likelihood that a belligerent will suffer lopsided casualties increases substantially. The same is true of both mass desertion and defection, where the effects of inequality are especially damaging in the modern era. Blocking detachments are also far more likely to make their appearance when inequality rises. These four wartime pathologies cluster in predictable fashion. Belligerents saddled with high or extreme levels of military inequality, for example, are far more likely to be wracked by multiple wartime deficiencies simultaneously than are similar belligerents with low levels of inequality. This is not simply a story of ethnic diversity, however. The number of ethnic groups in a given military has at best only a modest effect on battlefield performance. Instead, it is the state's prewar treatment of its constituent ethnic groups that conditions how its soldiers will fight. The roots of military power, these findings suggest, lie not in troop strength or technological advantages but in the political and social realities that mold and shape armies.

4.1. Building Project Mars

The chapter's empirical tests were made possible by the construction of Project Mars, a new dataset of 229 belligerents fighting 250 conventional wars between 1800 and 2011. I have relegated much of the discussion of coding rules, data sources, and auditing procedures to an extensive online appendix ("Sources, Data, Methods"). A few features are worth emphasizing here, however, as they bear on matters of data quality and reliability. A team of 134 coders, combing through materials in 21 languages, worked for nearly seven years to construct and then stress-test a battery of new measures for inequality, battlefield performance, and alternative explanations. Most of these research associates worked as part of Blue team, whose responsibility was to build-out Project Mars by tracking down initial data for its variables. A second, smaller, Red team then subjected Blue team codings to random audits to improve data quality, minimize errors, and raise intercoder agreement on difficult coding decisions. Measures for data quality and our confidence in our judgments were also constructed for core variables, including all measures of military inequality and battlefield performance. We also collected high, low, and mean estimates for core variables to capture uncertainty about these data and to reduce sensitivity to their often-contested, often-incomplete nature.[2] No cross-national effort of this ambition is entirely free of error, of course. These procedures do, however, enable the rigorous study of inequality over a large historical sweep, helping uncover patterns hidden for too long by the absence of relevant data.

4.1.1. Measuring Military Inequality

Coding teams collected two types of data to construct military inequality coefficients for all 825 belligerent observations. First, we took snapshots of each army's ethnic composition by calculating the proportion represented by each group among ground forces on the eve of war. We included all ethnic groups that represented greater than one percent of the army's personnel. For standing armies, we timed our snapshot to one year before the war to minimize the chance that leaders, anticipating conflict, rigged their armies to avoid future problems among restive ethnic groups. For non-standing armies, as well as expeditionary forces constructed to deal with emergent threats, we measured ethnic composition on the eve of the war's first battle, as the

2. This follows best practices established by the UCDP/PRIO Armed Conflict Dataset for post-1945 civil wars (UCDP/PRIO 2015).

opposing sides marshaled for armed combat.[3] In all cases, we include only ground forces in our calculations. We also incorporate colonial forces, auxiliaries, and volunteers from other states or groups if they served under the belligerent's command. We cast our net wide, drawing on a variety of sources to construct these demographic data. These include official histories,[4] regimental narratives,[5] formal orders of battle and tables of organization for units at the opening battle,[6] casualty lists,[7] and contemporary reports on the composition of enemy forces by military intelligence,[8] participating soldiers,[9] and enterprising journalists on the battlefield.[10]

Second, we coded the state's prewar treatment of each ethnic group according to the threefold typology outlined in the theory chapter. Inclusion, defined as the absence of state-orchestrated group-based discrimination or violence, was assigned a 0 value. All forms of state-directed discrimination, including political, economic, and within the military itself, were given a 0.5 value. The use of collective coercion and violence by the state against a specific ethnic group was assigned a value of 1. If an ethnic group was subjected to multiple forms of collective punishment, we assigned a 1 value. We generated our narratives of state treatment from political histories of these belligerents. For the post-1945 era, we cross-validated our accounts of ethnic treatment with data contained in the Ethnic Power Relations (EPR) and Minorities at Risk (MAR) datasets. We did not impose a minimum threshold for collective violence. In practice, however, we are capturing bouts of collective violence large enough to be visible in primary documents or secondary sources. These events tend to skew toward large-scale state-directed campaigns rather than one-time events. We anchored our measurement of state treatment in the five years preceding the war; that is, we scanned for evidence of state treatment prior to the war but not in the distant past. This

3. For example, we code the ethnic composition of General Gordon's Relief Expedition to Khartoum (1884–85) during the First Mahdi War, not the entire UK Army, using 1884 as a measurement benchmark. As a result, the same belligerent, usually a colonial power, can have sharply different military inequality coefficients depending on the location of hostilities. Staying with the UK example, its forces had radically different ethnic composition during the near simultaneous Boxer Rebellion (1900) in China and War of the Golden Stool (1900) against the Ashanti Empire.

4. See, for example, Airapetov 2014, 2015*a*, *b*.

5. See, for example, Bruce 1906.

6. See, for example, Haythornthwaite 2007.

7. See, for example, Lopukhovsky and Kavalerchik 2017.

8. See, for example, Burton 1908.

9. See, for example, McCormick 1859.

10. See, for example, Hardman 1860.

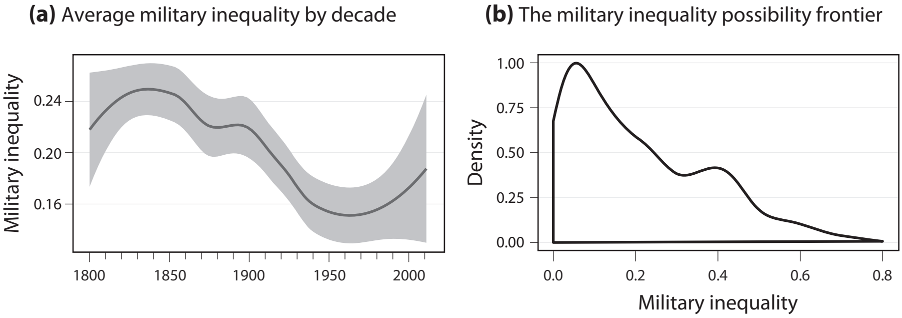

FIGURE 4.1. Summary Statistics: The Military Inequality Coefficient

coding rule ensures that exposure to state discrimination or repression was recent (or still ongoing) for soldiers as they marched into battle. Consistent with the book's argument, we do not assume that intergenerational trauma is at work. Instead, it is the recent lived experience of these soldiers that matters for how they regard the terms of their military service and their willingness to fight and die for the regime.

With these two types of data in hand, the construction of the military inequality coefficient (*Military inequality*) is straightforward. I simply interact each ethnic group's share of the army with its prewar treatment by the regime and sum the totals, creating a point estimate somewhere along a scale that ranges from 0 (perfect equality) to 1 (perfect inequality).[11] We do so for high and low estimates for each ethnic group, thereby constructing high and low estimates of military inequality for each belligerent. All analyses below use the mean values of military inequality.

To provide a sense of military inequality's distribution over time, I plot its mean value by decade for the 1800–2010 era in figure 4.1(a). A quick glance at the curve punctures the notion that military inequality has been crowded out by modernization and democratization globally. True, the mean level of a belligerent's military inequality coefficient has declined somewhat over time, falling from 0.23 in the early modern era to 0.17 in the modern one.[12] But military inequality has been, and remains, a durable feature of the international landscape, fluctuating within a narrow band between 0.24 and 0.16 for

11. The formula is $\sum_{i=1}^{n} pt_i$, where p is the ethnic group's share of an army's prewar strength; t represents the nature of the regime's prewar treatment of a specific group, where possible values are 0, 0.5, 1, representing inclusion, discrimination, and repression, respectively; and n is the number of ethnic groups in the army.

12. This difference is statistically significant at $t = 4.86$, 807.17df, $p = 0.000$.

over two centuries. It also appears to be enjoying something of a renaissance since 1945.

Figure 4.1(b) illustrates the military inequality possibility frontier by plotting the density of military inequality for all 825 belligerent observations. While the mean military inequality coefficient is 0.205, we can see how belligerents tend to cluster predominantly at lower levels of inequality. The number of observations becomes increasingly infrequent as military inequality advances, petering out in the 0.60–0.80 range. This is consistent with the earlier conjecture that armies cannot reach the theoretical maximum of military inequality due to command and control reasons. A basic minimum of soldiers must be drawn from core groups to maintain order, either through passive surveillance or active coercion, and to ensure that the army remains a cohesive fighting force. In practice, the absolute minimum appears to be twenty percent of a deployed force (or 0.80 on the military inequality continuum). Indeed, the highest military inequality coefficient recorded in Project Mars was 0.75, suggesting a natural upper bound for actual inequality.[13]

Empirical reality is messy, of course, and often works to confound our tidy coding protocols. While a fuller discussion of coding rules, along with the codebook itself, has been posted online, two issues deserve flagging here. First, colonial armies, that is, formations staffed principally by soldiers drawn from colonies, pose a particular conceptual challenge, standing at odds with standard renditions of armies as national in character. Put simply, what to do with soldiers of empire? Should Senegalese *tirailleurs* in French service, Eritrean *askari* in Italian employ, and the diverse soldiers of British West Indies and Bengal Army regiments be treated as members of the political community? On the one hand, these soldiers, as second-class colonial subjects, clearly faced ethnic discrimination (or worse), including the inability to command their own units and, of course, access the upper reaches of political power. On the other hand, these soldiers, often volunteers, were usually drawn from groups that enjoyed a greater share of prestige ("martial races") and political power within the colonial possession, suggesting a relatively higher degree of inclusion than other, less favored, groups. To square this circle, I calculated the military inequality scores for these armies twice: once with colonial groups treated as included and assigned a 0 for state treatment, and once as discriminated against, with a 0.5 value assigned.[14] This decision rule has the effect of adding a penalty for colonial status even if these soldiers were (relatively) privileged in their colonial political communities.

13. Individual units within the larger army may, however, exceed this threshold, as chapter 8 demonstrates.

14. If these soldiers were drawn from repressed populations, then a 1 value is assigned instead of the 0.5 value.

Second, some belligerents declared their independence on the eve of war—indeed, it may have been the precipitating event—or during its opening days. As a result, these independence-seeking belligerents lack the political history necessary for evaluating their treatment of ethnic groups within their boundaries. When confronted with a belligerent fighting in the first few years (or days) of its existence, we sought to code its initial treatment of ethnic groups from its first days regardless of whether it survived the war. In doing so, we sought as much temporal separation as possible between the state's treatment of its constituent ethnic groups and the war's opening to prevent inequality being driven by wartime dynamics. We also designated these belligerents with an indicator variable (*war birth*, see below) to demarcate them from the rest of the Project Mars universe for additional robustness checks.

To soften the issues arising from measurement difficulties, I constructed a second measure, *bands of inequality*, that assigns belligerents to one of four "bands" based on their military inequality coefficients. These bands are Low (0–0.20), Medium (0.21–0.40), High (0.41–0.60), and Extreme ($\geq$0.61). This simple classification scheme reduces bias from measurement difficulties while providing a natural grammar for speaking about the magnitude of inequality across belligerents. In the analyses below, I use bands as both a sensitivity check for military inequality and a simple way of interpreting how each increase in the "dosage" of inequality affects battlefield performance.

4.1.2. *Battlefield Performance*

Following the book's conceptual and theoretical discussions of battlefield performance, I draw on five specific measures to test the relationship between military inequality and wartime outcomes. These include: (1) a belligerent's loss-exchange ratio and, specifically, whether its LER drops below parity; (2) the incidence of mass desertion and (3) mass defection; (4) the use of blocking detachments; and (5) a composite battlefield performance index (BPI) that aggregates these four measures into a single family index. I detail each below.

Loss-exchange ratios are defined as the relative distribution of casualties inflicted versus suffered by a belligerent (or coalition) during a war. More specifically, loss-exchange ratios are calculated as the number of enemy soldiers killed by a belligerent divided by the number of soldiers lost by that belligerent to enemy fire.[15] A LER above one therefore indicates that a belligerent

15. For coalitional wars, each coalition's combined loss-exchange ratio is used instead of each belligerent's due to difficulties in assigning responsibility for inflicting casualties during a multiparty campaign or war.

is inflicting greater losses than it is suffering; a one designates parity; and values below one denote that the belligerent's forces are suffering greater casualties than they are inflicting. Care was taken when collecting data to concentrate only on soldier fatalities. Wounded soldiers, prisoners of war, missing individuals, and deaths from disease, a particular problem during the early modern era, were excluded from these counts to the best of our abilities. Civilian deaths, too, were excluded; these ratios are meant as a measure of force-on-force killing, not as an index of the lethality of the war itself. We generated high, low, and mean LER estimates for each belligerent or coalition to reduce sensitivity to competing fatality claims and reporting inaccuracies. To facilitate ease of interpretation, I constructed *LER below parity*, a dichotomous measure that records whether a belligerent's army lost more soldiers to enemy fire than it killed during the war.

Desertion is defined here as the unauthorized wartime withdrawal of soldiers, including entire units, from the battlefield or adjacent rear area with the intention of permanently abandoning the fight. Withdrawal can take two forms. Soldiers may attempt to return to their prewar life by hiding among the civilian population to escape state authorities. Renegade soldiers may also resort to brigandage in rear areas, or even at home, without coordinating with enemy forces. This definition excludes several types of behavior often conflated with desertion. Temporary absences, as when soldiers head home to plant or harvest crops, usually with tacit official approval, but then return when these duties are discharged, are excluded. Such practices were routine among Confederate soldiers during the American Civil War, for example.[16] Trench mutinies, including those that swept through nearly half the French Army after the disastrous Second Battle of Aisne (1917), are also excluded, as soldiers rebelled in place but did not abandon their posts.[17] Refusals to serve, as well as collective protests, are also excluded from this conceptualization. I also distinguish between mass desertion and simply chaotic retreats, where formations collapse under enemy pressure and their soldiers scatter temporarily before reconstituting their units to continue fighting. Libyan forces fighting in the 1978–79 Uganda-Tanzania War were known for their disorganized retreats, for example. Far from home, however, these forces eventually regrouped rather than risk mass desertion among hostile local populations.[18]

16. Enough soldiers deserted and remained absent, however, to exceed the 10 percent threshold for mass desertion. See Weitz 2005.

17. Pedroncini 1983; Doughty 2008, 510.

18. Pollack 2002, 373.

Mass desertion is therefore coded as occurring when ≥10 percent of an army's total deployed forces has decamped for home without authorization. This threshold is a pragmatic compromise designed to separate small-scale individual desertions that afflict nearly every army from large-scale desertion that can cripple war efforts.

In similar fashion, I measure *mass defection* as a dichotomous variable indicating whether ≥10 percent of a belligerent's fielded force switched sides during the war with the intention of taking up arms against their former comrades. I exclude side-switching by prominent military commanders if they acted alone. While these defections are important, commanders can shift allegiance for a variety of motives—personal enrichment and safety among them—that do not necessarily apply to the rank and file. In this view, mass defection is a particularly difficult act to complete successfully. Would-be defectors must evade their own fellow soldiers and then cross enemy lines, risking sanction by both sides. For some wars, mass defection may not be a realistic proposition. Suspicious commanders can station potential defectors well away from front lines while inflating the risks of defection, curbing enthusiasm. Enemy forces may not take prisoners or may treat them poorly, foreclosing this option. Desertion and defection are thus often substitutes, rather than complements, for at least some soldiers and wars. Indeed, consigning potential defectors to rear areas may reduce defection opportunities while increasing the odds of successful desertion.

I also constructed a dichotomous measure, *blocking detachments*, that records whether a belligerent deployed specialized armed formations to monitor and sanction its own officers and soldiers during wartime. These formations have five properties. First, these units are formally authorized by senior commanders; they represent official policy even if their origins can be traced to informal practices adopted haphazardly by frontline officers. Second, blocking detachments are typically stationed in the immediate rear of deployed forces to prevent unauthorized withdrawal and to prod reluctant soldiers into attacking. These units do not normally engage enemy forces, instead saving their fire for fellow soldiers. Third, these units have the formal authorization and military capacity to threaten and punish soldiers on their own remit. Potential sanctioning mechanisms include the forced return of soldiers to their units, dragooning into penal battalions, and execution, oftentimes in front of a soldier's comrades. In some cases, blocking detachments have the capacity to punish soldiers' families. Fourth, these units act as barriers between soldiers and rear areas, inhibiting the flow of information about battlefield progress and home-front conditions. Finally, while exceptions do exist, these units are usually staffed by personnel chosen for their perceived loyalty. The Zhili Clique fielded a special unit of exclusively child soldiers (the *Du Jun Dui*) to shoot deserters with cannon fire in its 1925 war against the Fengtian Clique in China, for example.[19]

I did not impose a minimal requirement for personnel size or fatalities inflicted by blocking units. Instead, three criteria were used to decide whether armies deployed blocking units. First, formal orders creating these units were issued by senior political or military officials. Second, these units were actually deployed on the battlefield in at least one major engagement. Third, these units executed soldiers, whether in a formal setting (i.e., a tribunal) or via shooting or bombardment during battle to stem desertion, block retreat, or drive soldiers forward.[20]

Finally, I constructed a summary measure, battlefield performance index (BPI), that pools these four measures into a single index. The BPI ranges from 0 to 1, where 1 denotes maximal battlefield performance, 0 indicates disastrous performance, and the presence of each of these four "pathologies" (LER below parity, mass desertion, mass defection, and blocking detachments) results in a 0.25 penalty subtracted from the belligerent's BPI score. Thus a 0.75 BPI value indicates the presence of one problem; a 0.50, two problems; a 0.25, three problems; and a 0, the presence of all four problems within the belligerent's army during the same war. More than simply a convenient summary index, BPI integrates elements of combat power and cohesion while capturing the intuition that these wartime behaviors may be correlated, at least partially, with one another.

4.1.3. Alternative Explanations

The mandate of Project Mars extended to the construction of new variables to test long-standing theories of military effectiveness. When possible, coding frameworks from existing datasets were applied to new belligerents and wars of Project Mars. In most cases, however, extensive data collection was required to build these new measures from scratch. Though the effort was considerable, so too was the payoff; we are now in a position to test a broad array of alternative explanations with a much deeper evidentiary base. I concentrate on three clusters of alternative explanations here.

As detailed in the introduction, perhaps the most oft-cited explanation for battlefield success is also the most intuitive: belligerents have greater combat power, and fewer cohesion problems, as their relative strength increases.

19. Guo and Qingchang 2003, 316.

20. Variation in the size, organization, and lethality of these units is an important area for future research.

To date, most empirical studies of war draw on the Correlates of War's Composite Index of National Capability (CINC) to measure relative distributions of military power. This index captures six measures of demographic, economic, and military strength, and provides an annual score that represents the belligerent's share of total global capability.[21] I break with tradition here, however, for two reasons. First, the addition of new belligerents to the Project Mars universe, some of them quite powerful, injects substantial measurement error into the calculation of both cumulative global capabilities and an individual belligerent's share. Second, and more fundamental, CINC scores do not provide any information about the military capabilities a belligerent actually deployed in a given war. Knowing a belligerent's relative standing in the global pecking order is unhelpful if they only deployed a fraction of the total strength, recruited armies "off the books," so to speak, by enlisting local or colonial populations, or surged their enlistment and industrial capacities beyond their CINC values during long attritional wars.

I therefore measure *relative forces* using a belligerent's (or coalition's) share of the total number of soldiers deployed during the war's first major ground battle. I focus on the first battle for several reasons. Military planners have traditionally, and often mistakenly, emphasized winning the first battle decisively as an important step to eventual victory.[22] Initially deployed forces are both more relevant than distant factors such as industrial production and have the advantage of not being confounded by endogenous wartime dynamics. That is, we avoid the danger of accidentally controlling for the effects of military inequality by using more aggregate measures, especially the total number of soldiers mobilized during wartime, that are themselves partially shaped by prior levels of military inequality.[23] I also include two additional measures of material power that avoid these inferential problems. *Great power* denotes whether a belligerent was a member of the exclusive club of leading major powers that dominated the international system due to their especially high levels of material capabilities, including militaries with global reach.[24] *Distance to battle* records the distance from the belligerent's capital to the

21. Sarkees and Wayman 2010, 26.

22. Nolan 2017.

23. On post-treatment bias, see Acharya, Blackwell and Sen 2016. In the supplemental analyses, I replace relative forces with the "bad control" of total number of soldiers deployed during the war as a robustness check.

24. United States (1899–); United Kingdom (1800–); France (1800–1940; 1945–); Germany/Prussia (1800–1918; 1925–45; 1990–); Austria-Hungary (1800–1918); Russia (1800–1917; 1922–); China (1950–); and Japan (1895–1945; 1990–) are considered to be Great Powers by the Correlates of War (Sarkees and Wayman 2010, 34–35). Great Powers represent 251 of 825 observations in Project Mars.

war's first engagement in kilometers (logged). This measure performs double duty. It captures a belligerent's ability to project power over distance, a hallmark of Great Powers.[25] Distance also accounts for the presumed decrease in opportunities for successful desertion as soldiers find themselves increasingly encamped among hostile populations and isolated from coethnic networks that can facilitate escape.[26] Alternatively, fighting near home might confer additional benefits, including heightened resolve, if not desperation, that pushes soldiers to fight harder to avoid a defeat that imperils the homeland directly.[27]

A second theoretical tradition centers around how political institutions shape battlefield performance. To measure *regime type*, I draw on the standard Polity2 indicator, where values range from −10 (the most autocratic) to +10 (the most democratic) and are taken in the year preceding the war.[28] Because these data are tied to the COW list of countries, I created new Polity2 scores for each new belligerent in Project Mars. For some, this process was admittedly clumsy. Polity2 scores were not designed with the wide range of political institutions, including tribal confederacies, khanates, and warlords, found among new Project Mars entrants. I also substitute regime type for a simple dichotomous measure of *democracy*, which denotes whether a belligerent's prewar Polity2 was ≥7, the standard threshold for an established democracy. The empirical expectation here is a simple one: the more democratic the belligerent, the better its loss-exchange ratio, the lower its incidence of soldier indiscipline, and the less likely it is to embrace fratricidal violence as official policy. *Democratic opponent* was also introduced to capture the perceived advantages of democracies by indicating whether the belligerent was facing an opponent with a Polity2 score of ≥7. Belligerents faced democratic opponents 121 times in these wars, about 15 percent of all war observations. Following existing theories, adversaries facing democratic opponents should exhibit battlefield problems with greater frequency, and possess lower BPI values, than states fighting non-democracies.

Battlefield performance may also be dictated by whether the belligerent initiated the war, perhaps through surprise attack, or was itself the victim of external aggression.[29] Initiators are likely to have superior loss-exchange ratios and lower incidence of mass indiscipline because they control the timing and pacing of opening offensives. Blocking detachments, too, should be less frequent among initiators than victims. *Initiator* therefore records whether

25. Boulding 1962.
26. McLauchlin 2014.
27. Castillo 2014.
28. Jaggers and Gurr 2004.
29. Slantchev 2004; Wang and Ray 1994.

the belligerent crossed a political boundary with the intent of seeking battle or was the first to openly attack the opposing side and inflict casualties. *Joiner* denotes whether the belligerent entered an on-going war as a third party. Joiners may possess especially high degrees of battlefield performance since they control the timing of their attack and are selectively choosing to engage an already-weakened foe.[30] From this standpoint, democratic initiators should be top performers, a claim I test using an interaction term between regime type and initiator (*regime type**initiator*).[31]

The institutional design of a belligerent's military may also shape battlefield fortunes. I therefore collected data on three dimensions of military organizations. First, *standing* denotes whether the belligerent had a permanent prewar army or whether it was levied for the purpose of attacking or defending. Standing armies are able to generate socialization pressures through persistent training and indoctrination that levied armies cannot match. As a result, standing armies should exhibit superior combat power and cohesion. Second, I built a sevenfold index that tracked how belligerents recruited their armies and specified their primary recruitment channel, from which $\geq$50 percent of its soldiers were recruited, and secondary channels, if present, through which the remaining 10–49 percent of its soldiers were drawn.[32] Two variables were built from this recruitment index. *Full volunteer* records whether the army was staffed solely by volunteers rather than conscripts, mercenaries, or slaves. These armies likely have higher motivation and better skills than their counterparts, resulting in superior combat power and cohesion. Existing theories suggest that conscripts are far more likely to desert and to suffer heavier casualties than professional volunteer armies, for example.[33] *Composite* is a dichotomous measure that records whether a belligerent's army was drawn from two or more different recruitment streams. Blending recruitment paths may render armies more vulnerable to indiscipline since their components may not mesh well together, creating vulnerabilities that opponents could exploit. Historically, colonial and other expeditionary armies were almost invariably composite in nature, drawing on multiple recruitment streams among different populations to recruit soldiers.[34]

30. Downes 2009.

31. Reiter and Stam 2002.

32. Belligerents could identify and mobilize soldiers from within their own societies via four channels: volunteers, conscripts, mercenaries, and slaves (or other coercion-based approaches such as abduction). Three additional channels were added to allow for recruitment outside of a belligerent's own borders or territorial possessions: volunteers, mercenaries, and slaves.

33. McLauchlin 2015; Horowitz, Simpson and Stam 2011.

34. Composite is only weakly correlated with *military inequality* and *bands* at 0.19 and 0.20, respectively.

4.1.4. *Controls*

Finally, I add several new control variables that might also influence battlefield performance. *Civil war* denotes whether the war was an armed conflict between two or more sides subject to the same prewar government that resulted in at least five hundred battle-related fatalities. This variable helps account for the possibility that mass desertion and side-switching might be especially prevalent in these kinds of wars given cross-cutting ethnic cleavages and the close proximity of soldiers to their families, homes, and support networks. Casualties might also follow a different logic in these wars. *Multiparty* captures whether the war was fought between more than two belligerents, which creates different implications for defection and desertion than bilateral conflicts.[35] I also created *war birth*, a binary variable that denotes whether a belligerent was fighting the war within the first two years of its existence. Faced with severe time and resource constraints, these new states might experience battlefield failures at a higher clip than their more established counterparts, especially if they are waging high-stakes wars of independence. These states possess slightly higher levels of military inequality than their peers, though the difference is only statistically significant in the early modern era. There are 79 war birth observations in the Project Mars dataset. A final variable, *non-COW belligerent*, identifies the 124 new belligerents added to Project Mars that are not included in COW's Inter-State War dataset.[36] These belligerents collectively account for 24 percent of all observations in Project Mars (193 of 825).

4.1.5. *Caveats*

All large-scale data collection efforts face limitations; Project Mars is no different. Two caveats deserve special mention here. First, these data provide only a snapshot of each belligerent's aggregate performance in a given war. While valuable for uncovering cross-national patterns, these data do not directly test the argument at the battle level and cannot identify intrawar dynamics.[37] Nor are these data fine-grained enough to test the proposed causal mechanisms, a task better suited for the qualitative paired comparisons detailed in subsequent chapters. A battle-level dataset remains something of a holy grail for conflict researchers, a truly Herculean endeavor best tackled collectively. As it stands, Project Mars represents the beginning of a conversation rather than the final word on these understudied battlefield outcomes.

35. Christia 2012.

36. The correlation between *war birth* and *non-COW belligerent* is modest (0.34).

37. For a similar cross-national approach to studying a sensitive battlefield topic, see Cohen 2016, 67–71.

Second, it is possible that political bias has crept into the data generation process, in several ways. Desertion, defection, and, above all, the use of violence against one's own soldiers, represent contested, often shameful, chapters of a nation's history that are usually sidelined in official narratives of the war. Few political regimes, especially those engaged in nation-building around wartime sacrifices, are likely to admit that "patriotic" soldiers switched allegiance or were coerced to fight. Powerful incentives therefore exist to minimize, or deny outright, the occurrence of soldier indiscipline. In extreme cases, reports of mass desertion and defection, and especially the use of blocking detachments, have been purged almost entirely from official narratives. To take one example, a growth industry now exists in Russia, fueled by the declassification of long-buried files, of books on so-called forgotten units that were overrun or destroyed during the Second World War and subsequently were stricken from official rosters to minimize losses.[38] At the same time, a victor's bias also exists, in which victorious powers rewrite the war's narrative to exaggerate the cohesion problems of their enemies while downplaying their own issues. This is especially likely in cases where the adversary is completely destroyed as an independent state, leaving no one to contest the skewed narrative.[39] In short, the combination of national sensitivities and political imperatives collude to sweep these behaviors under the rug. As such, Project Mars data on the incidence of mass desertion, defection, and blocking detachments likely exhibit a conservative bias that underestimates their actual frequency on the battlefield. They should, in other words, be viewed as the floor, not the ceiling, of possible estimates of their battlefield presence.

4.2. Descriptive Statistics

As the opening wedge in our empirical analysis, I summarize Project Mars data on military inequality and battlefield performance in table 4.1. In total, there are 825 belligerent observations, each of which summarizes an army's performance in a given war. Belligerents are sorted according to their respective band of inequality; the mean military inequality coefficient value is also provided for each band. We can clearly see the military inequality possibility frontier at work: there are 467 belligerents that entered their wars with low inequality (56 percent), 216 at medium levels (26 percent), 118 at high levels (14 percent), and only 24 (or 3 percent) at extreme levels of inequality. These data underscore the difficulty in fielding armies staffed solely or principally

38. In that vein, a heated, often partisan, debate has recently appeared on the role of blocking detachments and the extent of their violence against Red Army soldiers (Beshanov 2004; Glantz 2005*a*; Filippenkov 2016).

39. A state was coded as conquered or destroyed in 113 of 825 observations in Project Mars.

TABLE 4.1. Descriptive Statistics: Battlefield Performance by Bands of Military Inequality, 1800–2011

	Inequality		Battlefield Measures					Observations
Bands	Mean Military Inequality Coefficient	Number of Ethnic Groups in Army	LER Below Parity	Mass Desertion	Mass Defection	Blocking Units	BP Index	N
Low	0.079	4.22	24.6% (115)	21.1% (99)	8.3% (39)	9.4% (44)	0.841	467
Medium	0.286	5.57	45.4% (98)	43.5% (94)	24.5% (53)	24.1% (52)	0.656	216
High	0.461	5.67	66.1% (78)	68.6% (81)	40.7% (48)	33.1% (39)	0.479	118
Extreme	0.651	4.50	75.0% (18)	95.8% (23)	45.8% (11)	45.8% (11)	0.343	24
Mean	0.205	4.79	37.5% (309)	36.0% (297)	18.3% (151)	17.7% (146)	0.726	825

from repressed ethnic groups. The mean number of ethnic groups in each army also varies across different bands of inequality. Belligerents with low inequality and, somewhat curiously, extreme inequality both field armies that contain somewhat fewer ethnic groups than their medium and high inequality counterparts. This may partly reflect the success of inclusive states in fostering attractive national identities that subsume existing ethnic ones or, conversely, the outcome of repressive homogenization in extreme inequality belligerents. Conversely, it may simply be that these states had different preexisting ethnic makeups. I explore whether this difference in the number of ethnic groups, rather than their prewar treatment by the state, is driving variation in battlefield performance in the analyses below.[40]

Subsequent columns outline the percentage of belligerents within each band that suffered below-parity casualties, mass desertion and defection, and deployed blocking detachments during the war. The total number of occurrences for each battlefield pathology is provided in parentheses. Even a cursory glance at these data reveals how rising levels of inequality are associated with declining battlefield performance. Take, for example, the below-parity casualties. Only one-quarter of low inequality belligerents recorded loss-exchange ratios below parity. By contrast, 45 percent of medium inequality belligerents, nearly two-thirds of all high inequality belligerents, and a staggering three-quarters of all extreme inequality belligerents suffered greater casualties than they inflicted on enemy forces. In all, 140 of the 229 belligerents in the Project Mars dataset experienced wartime below-parity casualties at least once. Mass desertion follows an identical pattern. Low inequality belligerents experienced mass desertion in one-fifth of their total observations, compared with 44 percent of all medium inequality observations and two-thirds of high inequality ones. Mass desertion was a virtual certainty among extreme inequality belligerents, with only a single observation (the LTTE during its 1990–2002 war with Sri Lanka) failing to breach the 10 percent threshold denoting mass desertion. A full 127 of 229 belligerents experienced mass desertion at least once, emphasizing its widespread (and understudied) nature. Historically, both Chinese and Russian armies have been prone to bouts of mass desertion, with its occurrence noted in 46 percent (12/26) and 63 percent of their observations, respectively. Brazil, Kokand, and the Mahdiya also recorded mass desertion in nearly all of their wars, suggesting a chronic frailty within their armies.

40. The mean difference in the number of ethnic groups between low inequality and the remaining three bands is 1.31. Though substantively small, this difference is statistically significant at $t = 7.49$, $p \leq 0.000$.

Mass defection, too, becomes more common as we move up the rungs of inequality, though it remains less frequent than mass desertion. About 8 percent of low inequality belligerents' wartime observations were marred by mass defection, rising to one-quarter of medium inequality and 41 percent of high inequality belligerent observations. Mass defection occurred in just under half of all extreme inequality belligerents' wartime observations. A total of 89 different belligerents experienced mass defection during at least one of their wars. Some armies faced chronic side-switching: Uruguay, the Yemen Arab Republic, and Hungary all experienced mass defection in two-thirds of their respective wartime observations. China's soldiers defected during 12 of 26 wars, especially during the nineteenth century, while the Russian Army suffered mass defection in 7 of 49 wars. Similar distributions were recorded for the deployment of blocking detachments. Some 10 percent of low inequality belligerent observations saw the emergence of these formations, compared with one-quarter, one-third, and nearly half of all observations by medium, high, and extreme belligerents. Surprisingly, 70 different belligerents resorted to blocking detachments during their wars. The Asante Empire used these formations in three of their four wars; Russia and the Soviet Union did so for nearly 40 percent of their wars; and the Ottoman Empire deployed them in over one-third of its wars.

Finally, belligerents with low inequality score highest on the composite battlefield performance index (BPI). With a 1 representing the best possible performance, low inequality belligerents recorded a mean 0.841 score. Medium inequality belligerents trail far behind, at 0.656, while high inequality belligerents manage only a 0.479 average score, which indicates that two of these four battlefield deficiencies are present. Extreme inequality belligerents bring up the rear, recording a meager 0.343 score, indicating that their armies routinely experience three of these four problems in the same war.

These descriptive statistics provide initial evidence of the dangers of inequality across individual and composite measures of battlefield performance. Each step up the rung of inequality results in diminished performance across measures that are imperfectly correlated with each other, helping build our confidence that we have correctly identified a wide-ranging and robust relationship between inequality and battlefield outcomes.[41] Inclusion does not guarantee battlefield perfection, however. Low inequality belligerents still fall short of the ideal on the battlefield performance index even as they perform comparatively better than belligerents at higher bands of inequality. Intriguingly, these battlefield problems occur with varying probabilities.

41. The highest correlation is between desertion and defection (0.34), followed by desertion and blocking detachments (0.23) and defection and blocking detachments (0.17).

Casualties below parity and mass desertion are twice as common among wartime experiences than either mass defection or blocking detachments. Defection arises at only half the rate as desertion, confirming my earlier intuition that would-be defectors face greater logistical and other obstacles than deserters. Blocking detachments also appeared in 18 percent of all observations, a far higher rate of self-coercion than our current theories admit. Finally, these data suggest an upper bound on battlefield dysfunction. Only 14 different belligerents managed to score a 0 on the battlefield performance index; a 0 was recorded on only 20 occasions. Russia holds the unfortunate pride of place for this achievement, reaching this nadir four times in its history.

4.3. Statistical Analysis, Part I: The Military Inequality Coefficient

Descriptive data are useful for drawing initial inferences about military inequality and battlefield performance. We require more sophisticated approaches, however, if we are to test whether this relationship survives in the face of alternative explanations and control variables. I therefore draw on statistical regression (Ordinary Least Squares and logistic regression) to test whether increased military inequality is associated with diminished battlefield performance in the early modern (1800–1917) and modern (1918–2011) eras of conventional war. For each measure of battlefield performance, I first regress military inequality alone, and then introduce a more complicated model that incorporates the full range of alternative explanations and control variables.[42]

Table 4.2 reports the coefficients from these models for all four measures of individual battlefield outcomes as well as the combined battlefield performance index for the early modern era. The results are stark: Military inequality is positively associated with casualties below parity, the outbreak of mass desertion and defection, and the fielding of blocking detachments at the $p \leq 0.001$ level regardless of whether military inequality is alone or tested with the full model. The same holds true for the combined battlefield performance index; military inequality is associated with decreased performance in Models 9a and 10a at the $p \leq 0.001$ level. That these findings change only slightly when the full battery of measures for alternative explanations and controls are added to the mix highlights the robustness of this relationship.

But does it hold in the modern era? Indeed, the same pattern holds when we reestimate the same models for the post-1917 era (table 4.3). Military

42. These full models were used to generate figure 1.1 displayed in the introduction.

TABLE 4.2. Battlefield Performance in the Early Modern Era (1800–1917)

	LER Below Parity		Mass Desertion		Mass Defection		Blocking Units		BP Index	
	Alone Model 1a	Full Model 2a	Alone Model 3a	Full Model 4a	Alone Model 5a	Full Model 6a	Alone Model 7a	Full Model 8a	Alone Model 9a	Full Model 10a
Military inequality	4.528***	3.785***	5.145***	5.140***	5.411***	5.860***	4.141***	4.133***	−0.892***	−0.782***
	(0.757)	(0.725)	(0.629)	(0.686)	(0.746)	(0.796)	(0.689)	(0.893)	(0.064)	(0.065)
Regime type		−0.066†		−0.024		−0.017		0.039		0.003
		(0.036)		(0.029)		(0.034)		(0.030)		(0.003)
Initiator		0.180		−0.050		−0.380		−0.501		0.023
		(0.268)		(0.305)		(0.342)		(0.315)		(0.021)
Regime type*initiator		0.002		−0.003		−0.006		−0.042		0.002
		(0.041)		(0.035)		(0.049)		(0.046)		(0.002)
Democratic opponent		1.005**		0.144		−0.536		−0.097		−0.033
		(0.377)		(0.353)		(0.554)		(0.367)		(0.035)
Joiner		−1.083†		−0.042		0.262		0.005		0.015
		(0.623)		(0.466)		(0.591)		(0.711)		(0.043)
Relative forces		1.396**		−0.375		0.039		−0.098		−0.046
		(0.529)		(0.478)		(0.612)		(0.611)		(0.045)
Great power		−0.492†		−0.068		0.537		0.330		0.009
		(0.257)		(0.277)		(0.335)		(0.426)		(0.022)
Distance to battle		−0.183**		−0.041		−0.020		0.010		0.008
		(0.066)		(0.078)		(0.077)		(0.073)		(0.006)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Standing army		0.432		−0.492		−0.089		−1.174*		0.040
		(0.383)		(0.430)		(0.575)		(0.487)		(0.037)
Volunteer army		−0.030		0.554		0.904†		−1.416†		−0.020
		(0.426)		(0.440)		(0.532)		(0.762)		(0.035)
Composite army		−0.079		0.114		0.300		0.392		−0.023
		(0.239)		(0.329)		(0.344)		(0.456)		(0.025)
Civil war		−0.325		1.056***		1.248***		0.262		−0.102**
		(0.339)		(0.295)		(0.306)		(0.333)		(0.029)
Multiparty war		−0.713**		0.207		−0.242		−0.014		0.024
		(0.253)		(0.230)		(0.338)		(0.360)		(0.022)
War birth		−0.286		0.059		1.021		0.369		−0.044
		(0.572)		(0.609)		(0.719)		(0.607)		(0.055)
Non-COW belligerent		0.471		−0.629*		−0.470		0.215		0.012
		(0.318)		(0.316)		(0.403)		(0.462)		(0.027)
Constant	−1.519***	−1.404*	−1.648***	−1.092†	−2.971***	−3.501***	−2.664***	−1.812*	0.915***	0.856***
	(0.261)	(0.682)	(0.168)	(0.623)	(0.231)	(0.990)	(0.261)	(0.733)	(0.016)	(0.066)
Wald χ^2	35.76***	127.93***	66.81***	129.88***	52.56***	183.62***	36.17***	62.86***		
F Score									193.90***	36.66***
(Pseudo) r^2	0.107	0.234	0.132	0.176	0.147	0.229	0.091	0.140	0.363	0.423
N	482	482	482	482	482	482	482	482	482	482

Note: Standard errors clustered on 138 belligerents. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ † $p < 0.10$.

TABLE 4.3. Battlefield Performance in the Modern Era (1918–2011)

	LER Below Parity		Mass Desertion		Mass Defection		Blocking Units		BP Index	
	Alone Model 1b	Full Model 2b	Alone Model 3b	Full Model 4b	Alone Model 5b	Full Model 6b	Alone Model 7b	Full Model 8b	Alone Model 9b	Full Model 10b
Military inequality	4.469***	4.158***	6.616***	6.319***	5.312***	5.132***	5.646***	5.632***	−1.021***	−0.909***
	(1.010)	(1.1775)	(1.079)	(1.069)	(1.165)	(1.152)	(1.494)	(1.402)	(0.145)	(0.135)
Regime type		−0.029		0.008		−0.073		−0.070	0.003	
		(0.028)		(0.026)		(0.035)		(0.044)		(0.002)
Initiator		−0.450		−0.033		0.589†		0.435		−0.005
		(0.310)		(0.258)		(0.330)		(0.362)		(0.024)
Regime type*initiator		0.003		−0.000		0.0137		0.046		−0.001
		(0.041)		(0.038)		(0.0.39)		(0.045)		(0.003)
Democratic opponent		−1.236***		−0.320		−0.087		−0.469		−0.036
		(0.363)		(0.388)		(0.505)		(0.520)		(0.039)
Joiner		−2.350*		−0.662		−0.475		0.127		0.083*
		(1.110)		(0.700)		(0.833)		(0.719)		(0.034)
Relative forces		0.387		−0.241		−0.746		0.134		−0.005
		(0.661)		(0.611)		(0.810)		(0.856)		(0.071)
Great power		0.284		−0.330		0.173		0.344		−0.012
		(0.333)		(0.357)		(0.450)		(0.518)		(0.031)
Distance to battle		−0.132†		−0.036		0.145†		0.016		0.003
		(0.073)		(0.069)		(0.086)		(0.092)		(0.006)

Standing army		0.483		−0.341		$-0.901^{\dagger}$		0.559		0.008
		(0.514)		(0.484)		(0.491)		(0.667)		(0.052)
Volunteer army		−0.067		−0.480		0.151		−0.450		0.033
		(0.389)		(0.457)		(0.456)		(0.545)		(0.038)
Composite army		0.288		0.175		$0.843^{\star}$		0.358		−0.042
		(0.284)		(0.294)		(0.404)		(0.450)		(0.030)
Civil war		−0.245		0.297		$1.432^{\star\star\star}$		0.173		$-0.049^{\dagger}$
		(0.332)		(0.334)		(0.396)		(0.366)		(0.028)
Multiparty war		$-0.684^{\star}$		−0.090		0.343		−0.066		0.027
		(0.279)		(0.260)		(0.306)		(0.369)		(0.027)
War birth		0.022		0.672		0.220		$1.354^{\dagger}$		−0.094
		(0.459)		(0.421)		(0.472)		(0.683)		(0.061)
Non-COW belligerent		−0.050		−0.595		−0.143		−0.289		0.031
		(0.4521)		(0.480)		(0.406)		(0.610)		(0.046)
Constant	$-1.430^{\star\star\star}$	−1.089	$-2.058^{\star\star\star}$	$-1.223^{\dagger}$	$-2.651^{\star\star\star}$	$-4.083^{\star\star\star}$	$-2.707^{\star\star\star}$	$-4.046^{\star\star\star}$	$0.922^{\star\star\star}$	$0.917^{\star\star\star}$
	(0.256)	(0.855)	(0.270)	(0.709)	(0.284)	(0.979)	(0.321)	(1.097)	(0.023)	(0.084)
Wald χ^2	$19.57^{\star\star\star}$	$90.44^{\star\star\star}$	$37.60^{\star\star\star}$	$67.14^{\star\star\star}$	$20.80^{\star\star\star}$	$79.91^{\star\star\star}$	$14.285^{\star\star}$	$41.30^{\star\star\star}$		
F Score									$49.83^{\star\star\star}$	$12.93^{\star\star\star}$
(Pseudo) r^2	0.073	0.185	0.147	0.177	0.106	0.217	0.118	0.173	0.315	0.369
N	343	343	343	343	343	343	343	343	343	343

Note: Standard errors clustered on 124 belligerents. *** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ † $p < 0.10$.

inequality is once again positively associated with the occurrence of all four battlefield problems and remains statistically significant at the $p = 0.001$ level. This relationship persists even after a battery of measures for alternative explanations and relevant controls are included. Military inequality is also associated with lower scores on the composite battlefield performance index. In fact, the decrease in performance on the BPI is even steeper in the modern era, suggesting that the lethality of modern weapons may be increasing the penalties associated with military inequality. In brief, there is considerable evidence that rising levels of military inequality are associated with a range of negative battlefield outcomes across the early modern and modern eras of conventional war.

While military inequality surmounted the hurdle of statistical significance, it still remains to be seen whether it has substantively important effects on battlefield performance. I therefore turn to figures and percentage changes to illustrate military inequality's effects on each aspect of battlefield performance across each historical era.[43]

Figure 4.2 plots a belligerent's predicted probability of a loss-exchange ratio below parity, the incidence of mass desertion and defection, and the deployment of blocking detachments for armies during the early modern and modern eras. For each battlefield problem, and for each historical era, we can clearly see the positive relationship at work: an increase in military inequality is associated with an increased probability that a belligerent's army will suffer from reduced combat power or cohesion. Loss-exchange ratios, for example, plummet as military inequality increases. At extreme levels of military inequality, it becomes a near certainty that the belligerent's army will suffer lopsided casualties in the early modern era. During the modern era, extreme levels of military inequality are associated with a 75 percent predicted probability that loss-exchange ratios will be unfavorable. The same holds true for mass desertion. Once a belligerent breaches a military inequality coefficient of 0.60, mass desertion appears almost unavoidable in the early modern era and only slightly less likely in the modern one. The odds of mass defection, too, skyrocket in the modern era, with about a 35 percent likelihood at a 0.40 military inequality coefficient but a nearly 60 percent likelihood once the belligerent reaches 0.60. Consistent with the results above, the modern era especially punishes belligerents with high inequality; the likelihood of experiencing mass desertion or defection top out at higher probabilities than during

43. I report percentage point estimates in the text without 95 percent confidence intervals for readability; these are plotted on the relevant figures.

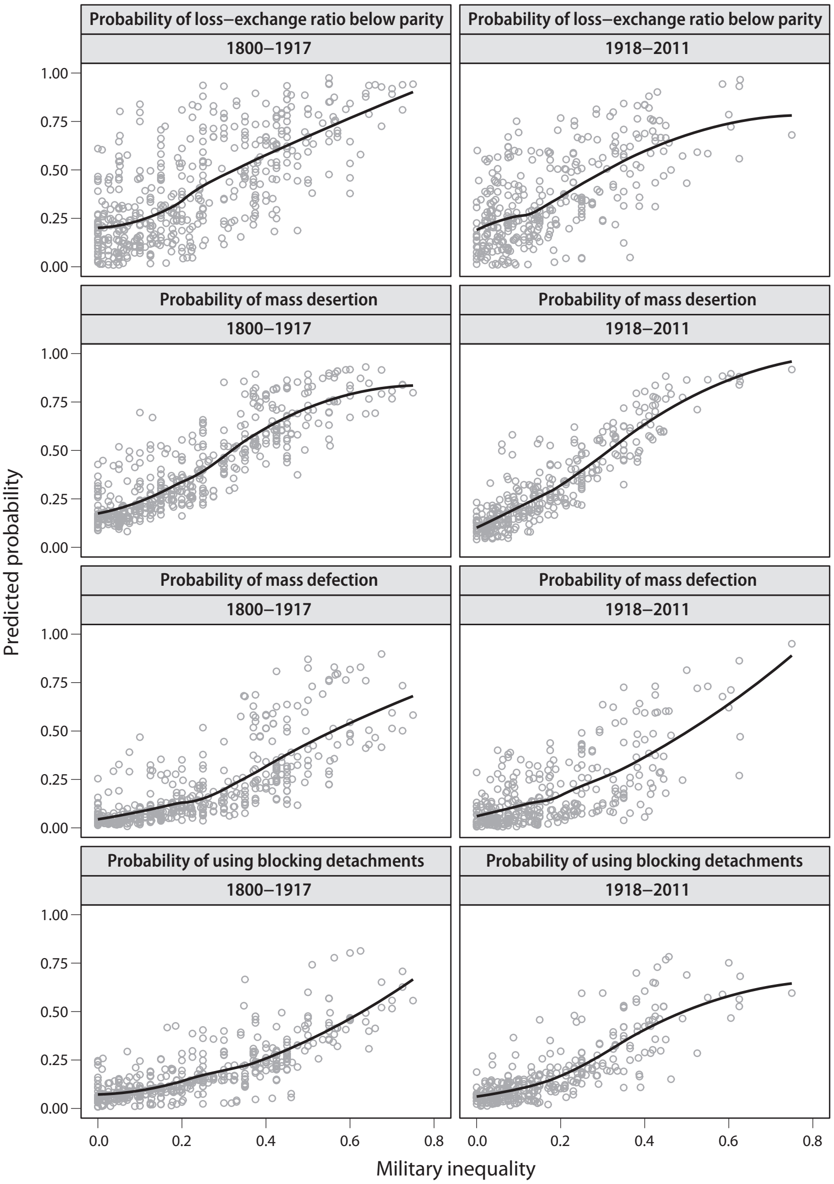

FIGURE 4.2. Military Inequality and Battlefield Performance by Historical Era

FIGURE 4.3. The Negative Relationship Between Military Inequality and the Battlefield Performance Index by Historical Era

the early modern era. Relatedly, belligerents turn to blocking detachments at lower levels of military inequality in the modern era, perhaps a tacit recognition of the greater probability of mass desertion and defection occurring within their ranks.[44]

Figure 4.3 extends these statistical tests to the composite battlefield performance index. Once again, higher values of military inequality are associated with diminished performance across both eras. Highly inclusive armies clearly attain the highest BPI values, though it bears emphasizing that these belligerents do not necessarily obtain a perfect index score. Once belligerents reach a threshold of 0.20 for military inequality, however, they have already lost a mean of 0.25 from their BPI score, indicating that their armies experienced at least one of the four battlefield problems tracked by Project Mars. By 0.40, the average BPI score has dropped to about 0.55, nearly halfway down the BPI scale. Belligerents with a military inequality score of 0.70 have witnessed their mean performance fall precipitously to about 0.30 on the BPI scale, a near-disastrous value. The penalties imposed by inequality are again somewhat higher in the modern era, particularly for those belligerents toward the more extreme end of the military inequality continuum. Compare, for example, the mean BPI score for belligerents at a military inequality value above 0.60 in the early modern and modern eras. These belligerents actually dip below a predicted mean 0.25 BPI value in the modern era, denoting

44. All predicted probabilities generated using the full regression models outlined in table 4.2 and table 4.3.

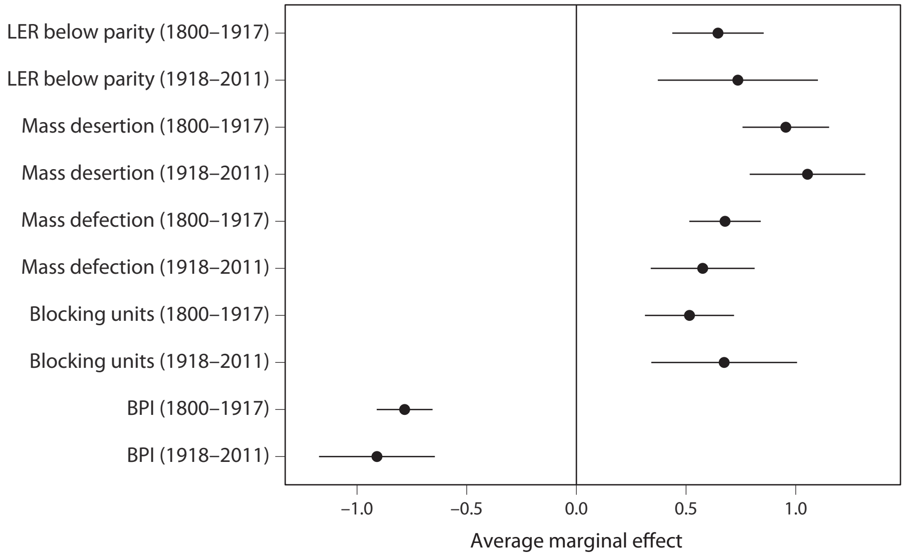

FIGURE 4.4. The Marginal Effects of Military Inequality on Battlefield Performance by Historical Era

a disastrous level of battlefield performance (narrowly) avoided by similar belligerents in the early modern era.

How substantively important is military inequality for explaining these outcomes? Figure 4.4 plots the average marginal effects and associated 95 percent confidence intervals of military inequality on each battlefield outcome for each historical era.[45] Average marginal effects can be interpreted as the predicted likelihood of an outcome given a one-unit change in the variable—here, a shift in military inequality from 0 to 1—while marginalizing over all other covariates. All reported estimates are substantively large, exceeding a 50 percentage point increase in every instance, and are statistically significant. In the early modern era, military inequality is associated with a 64.5 percentage point increase in the likelihood of suffering below-parity casualties; in the modern era, this estimate rises to a 73.6 percentage point increase. Military inequality's predicted effects on mass desertion are even larger, reaching staggering 95.4 and 105.4 percentage point increases in the early modern and modern eras, respectively. Mass defection also records large swings of 67.8 and 57.6 percentage point increases when shifting from perfect equality to perfect inequality in the early modern and modern periods. Blocking detachment deployment also proves sensitive to military inequality,

45. All average marginal effects are determined by a generalized linear model (GLM) using the full complement of measures for alternative explanations and controls.

registering 51.6 and 67.3 percentage point increases in the likelihood of their use during the early modern and modern eras, respectively. The BPI also reflects the negative effects of inequality. Average marginal effects of −0.783 and −0.909 are recorded for the early modern and modern eras, signifying that the shift from perfect equality to perfect inequality results in a near total collapse of battlefield performance for the average belligerent. A −0.783 decrease represents most of the BPI continuum, suggesting that belligerents with extreme inequality will have at least three of these four battlefield problems in the early modern era. The trend is even worse in the modern era; a −0.909 decrease translates into a performance near the bottom of the BPI scale in which almost all problems are present on average. In short, military inequality is not only statistically significant but also substantively important for explaining both individual and aggregate battlefield performance within each of these distinct historical eras.

4.4. Ethnic Diversity and Battlefield Performance

Can these findings be explained simply by an army's ethnic diversity? Perhaps armies face increased transaction costs and coordination problems as more ethnic groups are represented in the ranks. Language barriers, incompatible ethnic preferences, and disagreement over policies may all increase with the number of ethnic groups under arms, conspiring to drag down an army's battlefield performance.

To test this claim, I reestimated the models above with a new variable, *ethnic groups*, that records the (logged) number of ethnic groups present in the army at the war's outset. Figure 4.5 plots the average marginal effects associated with ethnic groups. For the most part, the substantive effects of the number of ethnic groups are modest and, in several instances, indistinguishable from zero. A one-unit change in ethnic groups—that is, a shift from one to nineteen ethnic groups in the army—results in an 8.8 percentage point increase in the predicted likelihood of suffering a below-parity loss-exchange ratio in the early modern era, for example. The predicted likelihood falls to 6.4 percentage points in the modern era and fails to reach conventional levels of statistical significance. Ethnic groups' largest effects are found on mass desertion, where a one-unit change is associated with an 11.3 and 16.1 percentage point increase in the predicted likelihood that soldiers will abandon the fight and return home. A similarly massive shift in the number of ethnic groups fails to budge the predicted likelihood of mass defection in the early modern era, with only a statistically insignificant 2.6 percentage point increase. Ethnic groups is associated with a 12.7 percentage point increase in the modern era, however. The relationship between the number of ethnic

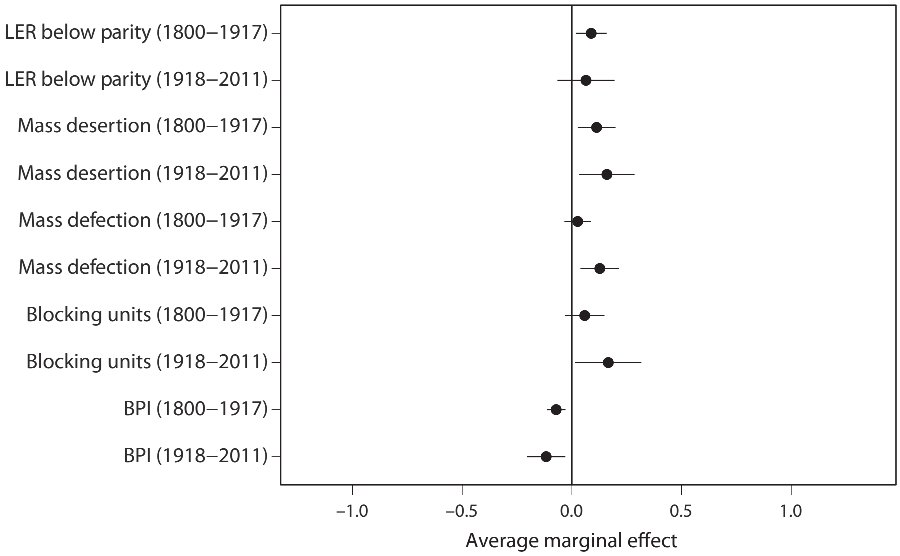

FIGURE 4.5. The Marginal Effects of Ethnic Diversity on Battlefield Performance by Historical Era.

groups and blocking detachments also fails to reach statistical significance in the early modern era; a 16.6 percentage point increase is noted in the modern era. Turning to a belligerent's BPI score, ethnic groups is associated with a slight −0.07 drop in the early modern era and a larger −0.12 decrease in the modern era. In short, the number of ethnic groups does appear to cause some friction, but the effect is dwarfed substantively by military inequality, which retains its expected direction and statistical significance at the $p \leq 0.001$ level or better for all regressions. The state's prewar treatment of ethnic groups, and their demographic weight within the army, are far more important for explaining battlefield outcomes than simply the number of ethnic groups within the ranks.[46]

4.5. Climbing the Ladder of Inequality: Statistical Analysis Using Bands of Inequality

Due to variation in the quality and quantity of historical evidence across belligerents, there is an unavoidable element of uncertainty around point estimates of military inequality. To mitigate these issues, I employ a second measure, *bands of inequality*, that collapses the military inequality continuum

46. The correlation between military inequality and ethnic groups is quite modest at 0.24.

into four levels ("bands") with clear cutpoints: low (0–0.20), medium (0.21–0.40), high (0.41–0.60), and extreme ($\geq$0.60).[47] This scaled variable helps soften measurement issues while permitting the identification of broad trends across different levels of inequality. It also has the advantage of capturing an intuitive understanding of military inequality as a series of gradations rather than a specific point estimate. In turn, bands sets up a threefold comparison of military inequality's effects when shifting from (1) low to medium bands; (2) medium to high bands; and (3) low to high bands. I use first differences to convey the substantive interpretation of the magnitude of the effects of these cross-band shifts, measured in percentage points, on the predicted likelihood that a given battlefield problem will manifest itself within a belligerent's army in each historical era.[48]

Replacing *military inequality* with *bands* and then reestimating the full models in table 4.2 and table 4.3 returns broader similar results as obtained above. All regressions yield a positive and statistically significant bands at $p \leq 0.001$ or better for each battlefield measure, including the composite BPI, for each era. What does this mean substantively for battlefield outcomes?

Beginning with loss-exchange ratios below parity, a shift from the low to medium band of inequality is associated with a 17.2 percentage point increase in the likelihood that a belligerent will suffer the brunt of casualties in the early modern era. A similar 17.6 percentage point increase is observed when shifting from the medium to high band of inequality. As a result, the shift from low to high levels of inequality is associated with a large 35 percentage point increase in the likelihood that a belligerent's loss-exchange ratio will be unfavorable (see figure 4.6(a)). For the modern era, a shift from the low to medium band of inequality is associated with a slightly smaller 14.1 percentage point increase that is surpassed by a 19.1 percentage point increase when moving from the medium to high band. Shifting from the low to high band is thus associated with a 33.2 percentage point increase in the predicted likelihood of lopsided casualties during the modern era.

47. In the following analysis, I combine extreme and high inequality belligerents into the same band due to the small number of belligerents with a military equality value of $\geq$0.60.

48. First differences were generated by *Clarify* using full models outlined in table 4.2 and table 4.3. All continuous variables except regime type and regime type*initiator were set at their mean to create meaningful values (Hanmer and Kalkan 2013). All dichotomous variables were set at median values; $K = 1000$ simulations were estimated (King, Tomz and Wittenberg 2000; Tomz, Wittenberg and King 2003). The baseline belligerent is a stable autocratic non-Great Power with half of the total fielded forces that possesses a permanent (i.e., standing) and non-volunteer army recruited through composite channels, and that is fighting about 400–450 kilometers from home in a non-civil war.

We find a similar pattern with mass desertion. In the early modern era, a shift from the low to medium band is associated with a 20.6 percentage point increase in the predicted likelihood of mass desertion. This climbs to a 23.2 percentage point increase when shifting from the medium to high band of inequality. Together, a move from the low to high band of inequality is associated with a 43.8 percentage point increase in the likelihood that mass desertion will plague a belligerent's army. The results are even larger for the modern era. Moving from the low to medium band of inequality is associated with a 26.7 percentage point increase; from the medium to high band, a whopping 30.6 percentage point increase. Holding all other variables at their mean, a shift from the low to high band of inequality is thus associated with a 57.3 percentage point increase in the predicted likelihood of mass desertion (see figure 4.6(b)).

The same trends hold true for mass defection, though the magnitude of the changes in predicted probability is smaller than observed for mass desertion. In the early modern era, a shift from the low to medium band of inequality is associated with a 10.8 percentage point increase in the predicted likelihood of mass defection, rising to 22.6 percentage points when moving from the medium to high band. All together, a shift from the low to high band of inequality is expected to increase the predicted likelihood of mass defection by 33.5 percentage points. This relationship attenuates sharply in the modern era, however. Moving from the low to medium band is associated with only a 4.3 percentage point increase, while the medium-to-high shift is predicted to increase the likelihood of mass defection by only 9.5 percentage points. Transitioning from the low to high band is therefore associated with a modest 13.8 percentage point increase (see figure 4.6(c)). These results confirm the earlier intuition that mass defection is harder to coordinate successfully than mass desertion, helping to explain its relatively infrequent nature in Project Mars data. Indeed, would-be defectors must not only escape from their own lines but successfully navigate surrender to enemy forces, a fraught endeavor.

Bands of inequality also shape the predicted likelihood of blocking detachments being fielded, albeit modestly. Beginning with the early modern era, a shift from the low to medium band of inequality is associated with an 8.1 percentage point increase, while a shift from the medium to high band of inequality is associated with another 13.3 percentage point increase. Moving from the low to high band of inequality is thus associated with a combined 21.4 percentage point increase in the predicted likelihood that blocking detachments will be ordered onto the battlefield. The magnitude of these expected probabilities jumps in the modern era. We observe a 10.2 percentage point increase when moving from the low band to medium; an 18.5 percentage point increase

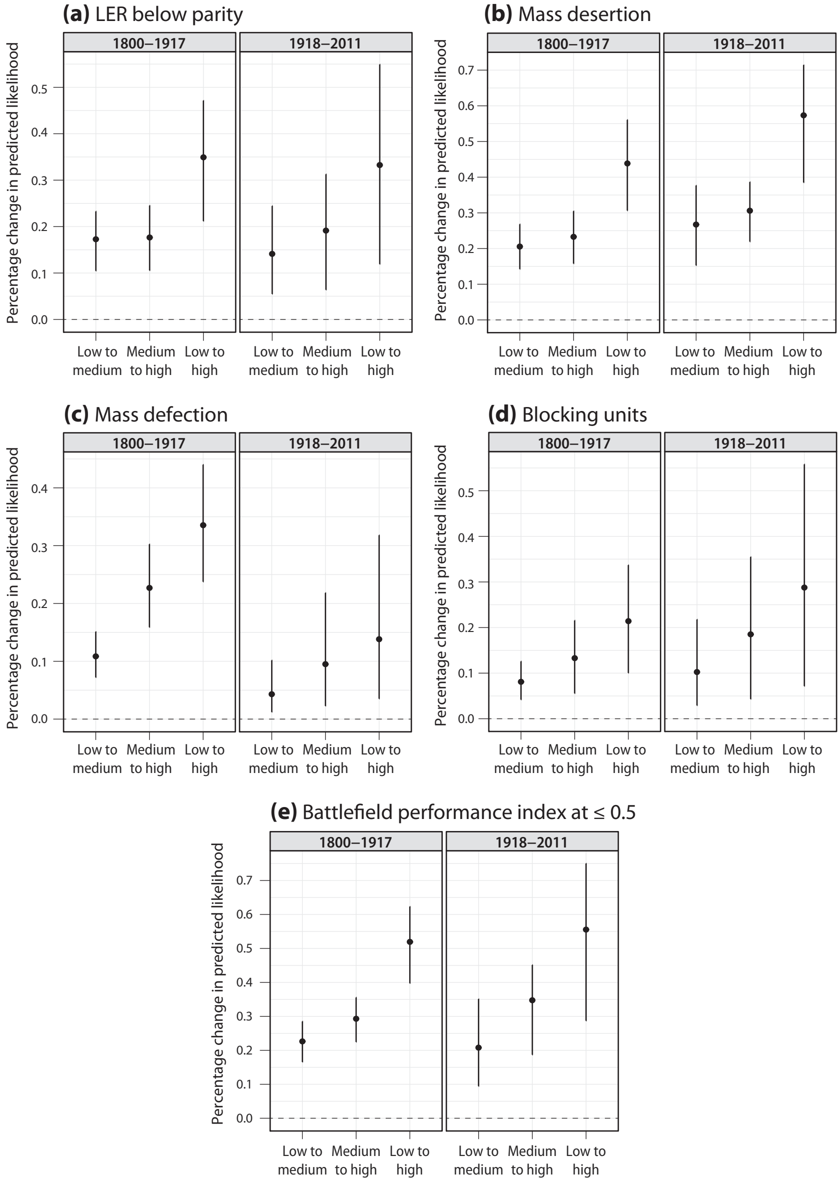

FIGURE 4.6. Changes in the Predicted Likelihood of Wartime Behaviors for Shifts in Bands of Military Inequality

when shifting from medium to high; and a 28.8 percentage point increase when moving from the low to high band of inequality (Figure 4.6(d)).

Finally, we can observe the effects of shifting bands on the aggregate battlefield performance index. Here I set the outcome of interest as the predicted likelihood that a belligerent will manage to score only a 0.50, the BPI's midpoint that denotes two of these four battlefield problems appeared in a belligerent's army. Starting with the early modern era, a shift from the low to medium band is associated with a 22.6 percentage point increase; from the medium to high band, a 29.3 percentage point increase; and a nearly 52 percentage point increase in the likelihood of a middling BPI score when moving from the low to high band. In the modern era, a shift from the low to medium band is associated with a 20.8 percentage point increase; a 34.7 percentage point increase when climbing from the medium to high band; and a massive 55.5 percentage point increase when moving from the low to high band of inequality. These are substantively very large effects. By comparison, sliding from a fully autocratic to fully democratic political regime, a truly enormous shift, is associated with only a paltry 7.7 percentage point increase in a belligerent's BPI score in the early modern era (and *Regime Type* does not reach statistical significance).

These findings collectively provide strong support for the claim that prewar military inequality sabotages a belligerent's subsequent battlefield performance. Both *Military Inequality* and *Bands* uncover the same relationship: rising levels of inequality are associated with the increased likelihood that each of these four problems will appear and that overall performance, as measured by the BPI, will correspondingly fall. These findings also confirm the trend initially noticed in the descriptive statistics: ethnic discrimination, best captured by the medium band of inequality, does exact a toll on battlefield performance. Ethnic marginalization is costly, jeopardizing a state's military fortunes even if collective violence is never wielded against these targeted populations. The expected stepwise pattern is also evident. Each new step up the rungs of the ladder of inequality is associated with a marked, in many cases neatly symmetrical, decrease in battlefield performance. The discontinuous jumps in the predicted probability of mass desertion and blocking detachment deployment in the modern era helps underscore how prewar violence can create incentives for targeted soldiers to escape and for commanders to turn to blocking detachments to force them to remain in place. Again, while low inequality belligerents are not entirely free of battlefield problems, their issues pale in comparison to belligerents that cluster at medium and (especially) high bands of inequality.

4.6. A Closer Look: Two-Control Group Comparison Using Matching

The convergence of findings using *military inequality* and *bands* is encouraging. Yet caution is warranted. It is possible that high inequality belligerents are systematically different than their medium or low inequality counterparts in ways that skew our conclusions about inequality's effects on battlefield performance. Perhaps these belligerents are far weaker or more autocratic than their more inclusive counterparts. If so, then these traits, rather than military inequality, might be driving observed differences on the battlefield. Moreover, if these imbalances are in fact present, then naive statistical analysis alone is likely insufficient to estimate the direction and magnitude of military inequality's effects. I therefore turn to matching to boost our confidence in the association between inequality and battlefield performance that was identified in the preceding analysis. Using a well-known approach called Coarsened Exact Matching, I constructed a balanced sample consisting of "treatment" and "control" belligerents that share similar attributes except for their prewar levels of military inequality. Here, "treatment" cases had high levels of inequality, while "controls" had either low or medium levels.[49] Belligerents with no match from the "control" pool are down-weighted or dropped from the analysis entirely. This procedure leaves behind a balanced sample in which differences in traits between treated and control cases are squeezed out, helping isolate the effects of military inequality.[50]

As an additional hurdle for my proposed argument, I check the robustness of prior findings by comparing high belligerents with two different control groups.[51] I first reestimate the full models in tables 4.2 and 4.3 using *Treatment*, which compares high belligerents to similar low inequality states (Control Group 1).[52] This comparison should illuminate the stark penalties associated with shifting from low to high inequality within the matched sample across all measures of battlefield performance.

I then reestimate the full models in tables 4.2 and 4.3 using medium inequality belligerents as the second control group. This comparison helps isolate the consequences of shifting from medium inequality, which most

49. I pool the small number of extreme inequality belligerents with those possessing high levels to create a single treatment group.

50. On matching, see Rubin 2006; Ho et al. 2007; Iacus, King and Porro 2012.

51. On two-control group comparisons, see Rosenbaum 2010, 332-39.

52. Specifically, I use CEM to adjust for imbalances across thirteen covariates from the full models in tables 4.2 and 4.3, with one exception. I swap regime type for the dichotomous democracy measure to facilitate matching; I also created a new interactive term for *Democratic Initiators*. All regressions use the CEM-generated weights.

belligerents achieve through some form of ethnic discrimination, and high inequality, a peak reached only by inflicting collective violence on targeted ethnic groups. We can also use this comparison to detect whether increasing the "dosage" of inequality at higher levels still erodes battlefield performance or, alternatively, if most of the causal action occurs in the comparison between low and high inequality belligerents (as represented as Control Group 1).

What, then, are the battlefield consequences of shifting from low to high inequality (Control Group 1), and from medium to high inequality (Control Group 2)? Figure 4.7 illustrates the average marginal effects that result from each of these comparisons for each battlefield measure across each historical era using only the matched data. Put simply, these results suggest that the average marginal effects of each shift are large and statistically significant across time and (nearly) all measures, indicating once again that military inequality is associated with decreased battlefield performance.

Take relative casualties, for example (figure 4.7(a)). We observe a large 33 percentage point increase in the likelihood that a belligerent will experience below-parity losses when moving from low to high inequality in the early modern era, holding all other variables at their mean. We observe another 15.5 percentage point increase in the same era when shifting from medium to high inequality, indicating that the shift from ethnic discrimination to collective violence imposes additional penalties. In the modern era, the low-to-high comparison is associated with a 27.1 percentage point increase in the likelihood of below-parity casualties, while a medium-to-high shift produces an additional 15.4 percentage point increase (though it dips slightly below conventional levels of statistical significance).

Mass desertion and other cohesion-related battlefield problems exhibit similar patterns. A shift to high from low inequality produces a 36.8 percentage point increase in the likelihood of mass desertion, while a shift from medium to high inequality creates another 20.5 percentage point increase (see figure 4.7(b)) in the pre-1918 era. The effects are even larger in the modern era: a comparison with Control Group 1 increases the likelihood of mass desertion by 46.2 percentage points, while Control Group 2 produces an additional 45.5 percentage point increase. Turning to mass defection, a shift from low inequality to high is associated with a 24.3 percentage point increase in the early modern era. A further 18.2 percentage point increase is associated with a jump from medium to high inequality in the same era. In the modern era, we observe smaller increases of 15.7 and 13.9 percentage points for Control Group 1 and 2 comparisons (figure 4.7(c)). Military inequality affects the likelihood of a belligerent using blocking detachments in similar fashion. Control Group 1 and 2 return near identical 15.5 and 13.3 percentage point increases in the odds of blocking detachments making a battlefield appearance in the

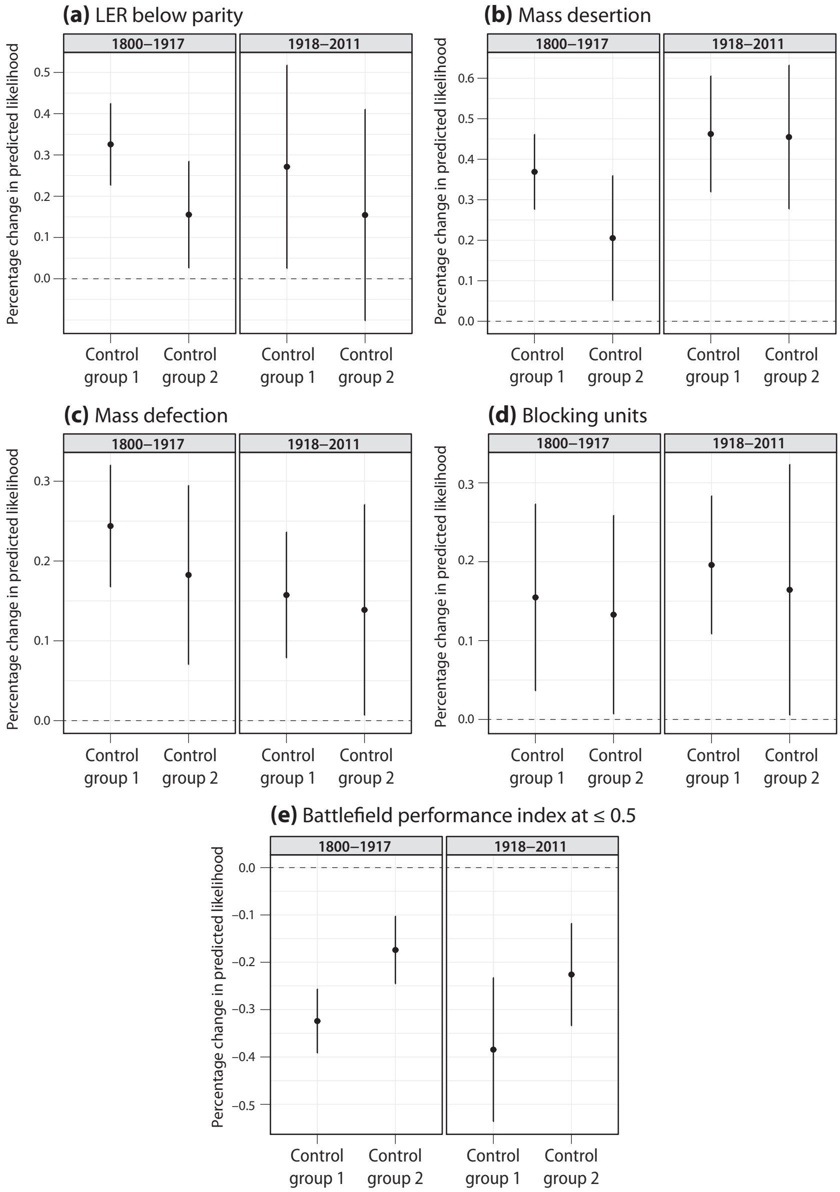

FIGURE 4.7. Average Marginal Effects of High Military Inequality Compared to Low (Control Group 1) and Medium (Control Group 2) Military Inequality Using Matching

early modern era, respectively. In the modern era, a shift from low to high is associated with a 19.6 percentage point increase, and from medium to high a further 16.4 percentage point increase (figure 4.7(d)). These same patterns hold in the aggregate battlefield performance index. In the early modern era, we observe downward shifts of −0.324 and −0.174 for Control Group 1 and 2 comparisons, respectively. In the modern era, the penalties are even more severe, with a −0.384 and a −0.226 decrease in BPI values for Control Group 1 and 2 (figure 4.7(e)), respectively.

In sum, the penalties levied against battlefield performance are highest with Control Group 1, where the disparity in military inequality is largest across belligerents. By contrast, the average marginal effects for Control Group 2 are typically reduced, as befitting the smaller difference in levels of prewar military inequality. Yet penalties still accrue when climbing from medium to high levels of military inequality. While ethnic discrimination imposes battlefield costs, the state's use of collective violence against its own citizens represents an additional distinct tax on its battlefield fortunes. This second set of analyses should be doubly reassuring: we not only recover the same basic relationship within the matched sample as we do the broader Project Mars dataset, but it survives the use of multiple comparison groups, indicating that we have correctly identified a durable empirical pattern.

4.7. Robustness Checks

This chapter has presented a raft of new data and findings. It is only natural to wonder, however, how robust these findings are to alternative model specifications and measurement strategies. While a full accounting can be found in the book's online appendix, I briefly discuss some of the most important robustness checks here. Drawing on the full models in table 4.2 and table 4.3 as benchmarks, a partial list of supplemental analyses includes: (1) reestimating all models with alternative measures for material preponderance, regime type, and military organization; (2) including fixed effects for decades, regions, specific belligerents, and multi-campaign wars;[53] (3) dropping all wars in which belligerents substituted air or naval power for ground forces, minimizing their exposure to casualties and opportunities for mass desertion and defection;[54] (4) reestimating all models with indicator variables for the confidence

53. Belligerents include the Ottoman Empire, United Kingdom, France, USA, Russia, and Germany. Multi-campaign wars include the Napoleonic Wars, World War I, and World War II.

54. Ten wars in total fit these criteria, including the 1999 Kosovo War and the 2001 Afghan War as well as older cases like the Netherlands during the Bombardment of Algiers (1816).

of our judgment in data quality for all battlefield performance measures; (5) sensitivity analysis using different thresholds (5 percent, 20 percent) for defining mass desertion and defection; and (6) replacing the binary variables for below-parity casualties, mass desertion, and mass defection, with mean estimates of their actual values. In all cases, military inequality retained its statistical significance and substantive importance.

Two robustness checks deserve special mention here. First, I conducted a placebo test in which, as a mad scientist of sorts, I randomly assigned new military inequality coefficients to belligerents for all 825 observations. Doing so permits investigation of whether we have correctly identified an association between inequality and battlefield outcomes or, alternatively, we have accidentally captured a spurious relationship due to some hidden process in the data. If an association is truly present, then military inequality should no longer be statistically significant once these new randomly generated values have replaced the original ones. Indeed, this is exactly what happens; military inequality loses statistical significance in nearly every model in table 4.2 and table 4.3, confirming that we have identified a genuine association at work.

Second, perhaps these findings are simply an artifact of the more expansive list of wars and belligerents of Project Mars. Nearly all existing quantitative work draws on the venerable, if smaller, Correlates of War data universe, and so what is needed is a direct examination of military inequality's effects within this more restrictive set of cases. I therefore reestimated the models in tables 4.2 and 4.3 twice, first using only COW-approved belligerents and then drawing only on COW-approved wars. This latter test is especially severe since it discards nearly two-thirds of the (hard-won) Project Mars dataset. In the COW-only belligerent analysis, military inequality remains statistically significant at $p \leq 0.002$ in every single model. Using COW's latest Version 4.0,[55] military inequality remains statistically significant at $p \leq= 0.01$ level in eight of ten models for both historical eras using only COW-approved wars. Only mass defection proves somewhat problematic: military inequality remains statistically significant at the $p \leq 0.05$ level for the early modern era but narrowly misses conventional levels of significance for the modern era. While Project Mars thus provides a more expansive view of conventional wars and their participants, the book's findings can nonetheless be replicated within the dominant COW universe, boosting our confidence in the robustness and generalizability of the relationship between military inequality and battlefield success.

55. Correlates of War 2010.

4.8. Assessing Alternative Explanations

The quantitative evidence marshaled by Project Mars unfortunately provides little support for leading alternative explanations. Material preponderance and regime type explanations in particular struggle to account for individual measures of battlefield performance and the combined BPI. Moreover, the decision to split the sample into two historical eras exposed inconsistencies in these arguments. It is commonplace, for example, to find that these variables are significant in one time period but not another, or that the direction of the proposed relationship in the early modern era reverses course once we move into the modern era.

Perhaps most surprising is how regime type appears largely irrelevant for explaining battlefield performance. A belligerent's prewar Polity2 scores, the standard measure for regime type, are almost never statistically significant, and the relationship between this variable and certain outcomes often changes direction depending on the historical era. The only time regime type reaches statistical significance is in the early modern era, when a shift from a full authoritarian to full democratic regime is associated with a 23.8 percentage point reduction in the likelihood that a belligerent's army will suffer below-parity casualties. Democracy, a binary variable capturing whether a belligerent had a prewar Polity2 score of ≥ 7, fares somewhat better when replacing regime type in these models. It is associated with a modest 7.8 percentage point reduction in the likelihood of a belligerent fielding blocking detachments and with a 6.8 percentage point increase in the composite BPI in the modern era. But it never reaches statistical significance for any other measure. Facing a democratic opponent is also associated with a 25.2 percentage point jump in the likelihood of suffering below-parity casualties in the early modern era and a similar 22.2 percentage point increase in the modern era. The notion that democracies can induce desertion and defection by credibly promising to treat enemy soldiers well given their commitment to human rights norms is not supported by these data, however.[56] Democratic opponent is not associated with an increased likelihood of mass desertion or defection, and belligerents appear no more likely to resort to blocking detachments to ward off the siren song of democracies. Finally, I find no evidence that democratic initiators are superior at any aspect of war-fighting in either era than their autocratic adversaries.[57]

56. See, for example, Wallace 2012; Reiter and Stam 2002, 79.

57. Democratic initiators are captured by the combined weight of regime type, initiator, and the interactive regime type*initiator term.

Standard indicators of material preponderance and local force ratios also fare poorly. Relative forces is only statistically significant once, where a preponderance of force is actually associated with an increase in the likelihood of suffering below-parity casualties. Shifting from the 25th to 75th percentile of initially deployed forces—imagine a belligerent army moving from being outnumbered 2:1 to outnumbering its opponent 2:1—results in a 16.6 percentage point increase in the odds of experiencing a poor loss-exchange ratio in the early modern era. Taken together, these non-findings confirm skepticism about both the importance of troop strength and of seeking decisive outcomes in the war's first battle through numerical preponderance.[58] Even Great Power status appears to have little connection to battlefield performance; great power is only significant in a single regression, where Great Power status reduces the likelihood of below-parity casualties by 9.3 percentage points in the early modern era. Casualties are, however, somewhat sensitive to how far a belligerent is fighting from its capital city. A 15 percentage point reduction in the likelihood of below-parity casualties is observed when shifting distance from the 10th to 90th percentile (roughly from 50 to 8,200 kilometers from the capital) in the early modern era. The same shift results in a modest 7.8 percentage point decrease in the modern era. If we treat distance from the capital as a proxy for power projection, then casualties do appear at least partly driven by material power. But cohesion problems appear divorced from a belligerent's power capabilities, with none of these measures ever shifting the probability of desertion, defection, or blocking detachments.[59]

New data leads to new insights, however. The traits of military organizations, often neglected in existing theories, emerge as important drivers of battlefield performance. Standing armies, for example, are associated with a 5.3 percentage point decrease in the likelihood of defection in the post-1917 era. These armies are also 15.7 percentage points less likely to deploy blocking detachments in the early modern era. These findings aside, non-standing armies largely hold their own with permanent standing forces, though with so few examples in the modern era we must take care to avoid sweeping generalizations. Composite armies, identified by their multiple recruitment streams, are associated with a small 3.8 percentage point increased likelihood of mass defection in the modern era, but are otherwise not especially vulnerable (or robust) compared to single-stream recruitment. Surprisingly, the

58. Nolan 2017.

59. *Non-COW belligerent* also jitters unpredictably across different battlefield measures and eras, rarely reaching conventional levels of statistical significance. These results are therefore not driven by the presumed relative weakness of the new belligerents added to Project Mars.

vaunted advantages of volunteer armies are not apparent in these data. *Volunteer* is only associated with a 6.2 percentage point reduction in the likelihood of using blocking detachments, and only for the early modern era. *War birth,* which denotes that a state was fighting and state-building simultaneously, also has only a weak relationship with these battlefield measures. These states sometimes have a greater likelihood of mass desertion and defection, but the results are sensitive to modeling choices and often inconsistent across eras. The only consistent finding is that war birth is positively associated with a 15.4 percentage point increase in the likelihood of blocking detachment use in the modern era.

One of the most powerful alternative explanations for battlefield performance lies in the political context of the war itself. Civil war is associated with a sharp 21.8 percentage point rise in the predicted likelihood of observing mass desertion in the early modern era. Compared with classical interstate war, civil wars also observe a greater probability of mass defection from belligerent armies. In the early modern era, civil war is associated with a 12.3 percentage point increase in the likelihood of mass defection, falling to an 8.4 percentage point increase in the modern era. As a result, the mean BPI score for belligerents enmeshed in civil wars is lower than belligerents fighting traditional interstate wars in both historical eras. Once again we observe the power of military inequality, which remains consistent in magnitude and statistical significance across both empirical domains.[60] While civil wars increase the risk of cohesion and discipline problems within armies, the proposed military inequality argument offers a unified account that bridges these two domains that have been studied in isolation for too long.

4.9. Conclusion

These statistical tests provide substantial evidence of the association between rising military inequality and declining battlefield performance. Both *Military inequality* and *Bands* remain statistically significant and substantively important across all four measures of battlefield performance as well as the combined BPI. Each new step up the ladder of inequality is associated with an increased likelihood of lopsided casualties, the outbreak of mass desertion and defection, and the fielding of blocking detachments designed to coerce soldiers into fighting. These results also hold across a difficult two-control group comparison test as well as a phalanx of robustness checks and

60. In a supplemental analysis, I split Project Mars into two samples (civil war and non-civil war) and reestimated all models. Military inequality remained statistically significant and substantively important in both samples.

alternative specifications. Together, the chapter has assembled a formidable set of challenges for the military inequality argument. Yet it remains the only explanation capable of explaining both individual battlefield outcomes and the overall pattern of performance across two different historical eras. Leading alternative explanations, including those emphasizing material power and regime type, struggle to stitch together compelling narratives for a single measure of performance, let alone the entire pattern. What these tests cannot do, however, is trace the causal processes linking military inequality through the proposed mechanisms to subsequent battlefield conduct. Nor can they capture the trade-offs between combat power and cohesion that commanders face once they field divided armies. As a result, we must turn to our paired historical comparisons for close-range investigation, a task I begin in the next chapter.