Fighting for Tyranny:
State Repression and Combat Motivation

Arturas Rozenas
New York University
arturas.rozenas@nyu.edu

Roya Talibova
University of Michigan
talibova@umich.edu

Yuri M. Zhukov
University of Michigan
zhukov@umich.edu

How hard will citizens fight to defend a state they see as tyrannical? Using a stylized model of soldier's choice, we show that exposure to state repression should increase effort by lower-motivated soldiers, but decrease effort by higher-motivated ones. We test this claim by utilizing over 100 million declassified Red Army personnel records from World War II. Our empirical strategy exploits plausibly exogenous variation in the scale of Stalin's repression prior to war due to explicit random targeting, logistical costs, and local administrative discretion. Consistent with our expectations, soldiers from places exposed to higher repression were more likely to fight until death and less likely to flee, but also received fewer decorations for personal bravery. Repression appears to have induced obedience at the expense of initiative and increased the human costs of war.

JEL Classification: D74, F51, H56, N44

September 2021

1. Introduction

The development and survival of states often hinges on their ability to extract resources for war-making from their populations.[1] In modern mass warfare, one such key resource is the effort that ordinary citizens exert on the battlefield. What drives individuals to risk their lives, resist the temptation to flee, and take personal initiative when fighting for their country? Since the conduct of individuals in battle affects states' perceived ability to impose costs on their opponents and, ulimately, their decisions to wage or end war,[2] a deeper understanding of these micro-level processes is essential.

Existing research has examined the effects of regime type,[3] institutional discrimination,[4] ideology,[5] culture,[6] group loyalties,[7] personal stakes,[8] pecuniary incentives,[9] and fear of punishment[10] on combat motivation. We propose a complementary perspective, which underscores the role of soldiers' prior interactions with the state outside their line of duty. If we accept the premise that military institutions do not evolve in isolation from broader social and political structures, as military theorists have traditionally argued,[11] then we must also acknowledge the importance of lived experiences vis-a-vis the state that soldiers had prior to their service. Soldiers for whom these experiences were mostly positive may approach their duties differently than those who have come to see the state for which they are fighting as unjust or tyrannical.

The military effort of the Soviet Union during World War II, which we study in this paper, is arguably the most paradigmatic case for the question at hand. Within the span of a few years, the Soviet state headed by Joseph Stalin went from inflicting mass terror

[1] Tilly, 1985.
[2] Bueno de Mesquita et al., 2005; Fearon, 1995; Slantchev, 2003.
[3] Reiter and Stam III, 1998; Talmadge, 2015.
[4] Lyall, 2020; Huff and Schub, 2021.
[5] Bartov, 1992; Barber IV and Miller, 2019.
[6] Pollack, 2018.
[7] Shils and Janowitz, 1948; Costa and Kahn, 2003.
[8] Hall, Huff and Kuriwaki, 2019.
[9] Grossman, 1991.
[10] Chen, 2017; Lyall, 2017.
[11] Clausewitz, 1832/1984, 592-593.

against its people to rallying those same people to fight in its name, in what became the world's deadliest-ever conflict. The Great Patriotic War — the eastern front of World War II — accounted for 93% of all European casualties and 18 of the 25 costliest battles on record.[12] The Soviet Union lost over 11.2 million military personnel and 17.9 million civilians.[13] Almost 40% of the battlefield losses comprised soldiers who were captured, surrendered, deserted, or went missing.[14] Historians have puzzled over these numbers and debated whether Stalin's prewar coercion alienated the population to the point where many did not want to defend their homeland.[15] Incentives to avoid fighting were indeed compelling. In the battle of Stalingrad, average life expectancy was 24 hours for enlisted Soviet soldiers and three days for officers.[16] Given these odds, it is remarkable that the Red Army managed to keep millions of its troops in line fighting while others fled.[17]

We conduct a quantitative study of how prewar mass violence by the state impacts the combat motivation of soldiers during war. We employ new detailed data on the Red Army in World War II, compiled from over 100 million declassified personnel records, to reconstruct the wartime trajectories of over 12 million soldiers. We also use micro-level data from over 2 million secret police case files on mass arrests prior to the war. By linking these records, we evaluate whether exposure to prewar repression could in part explain why some soldiers fought to death while others surrendered, deserted or went missing, and why some received decorations for valor and initiative in battle and others did not.

Our empirical approach exploits several features of Stalin's repression. The mass terror was locally arbitrary in that it targeted people not on the basis of individual behavior, but on the basis of high-order group-level characteristics like ethnicity, class, or geographic region. This permits us to estimate the effect of repression by comparing geographically proximate locations adjusting for the observables the regime used to select

[12] Overy, 1998, xvi.
[13] Surinov and Oksenoyt, 2015.
[14] Krivosheev, 1997.
[15] Edele, 2017; Thurston, 2000.
[16] Merridale, 2006.
[17] Reese, 2011.

victims. In addition, we exploit variation in repression induced by access through railways, which was the key logistical constraint in reaching and transporting repression victims to camps. Finally, motivated by the fact that local administrators had enormous discretion in the implementation of repression, we exploit discontinuous changes in arrest levels across administrative borders. Sensitivity analyses and falsification tests suggest that the implied assumptions behind these empirical strategies are plausible.

Across all three empirical approaches, we find consistent evidence that Red Army soldiers from places with high levels of prewar repression had systematically different battlefield outcomes compared to others. First, they were more likely to be killed or wounded in action. Second, they were less likely to flee the battlefield, go missing, or defy orders for which they would face punishment. Third, these soldiers showed fewer acts of bravery as far as we can judge from their award records. It appears that repression made soldiers more compliant, but also less willing to take initiative.

We conduct additional tests to rule out the possibility that our results reflect wartime discrimination against soldiers from heavily-repressed areas, their selective assignment to more dangerous parts of the front or to units engaged in deadlier tasks. The same systematic differences emerge when we compare soldiers serving concurrently in the same units, who were exposed to similar battlefield conditions, leadership, and group cohesion while deployed. We also show that repression affects soldiers' behavior through both direct personal exposure and the exposure of one's peers from the same unit.

We rationalize these empirical results using a stylized model of soldiers' decision-making process, which focuses on trade-offs between intrinsic and extrinsic motivations.[18] Individuals will perform a task either because they are intrinsically motivated to do so, or because the right extrinsic incentives (i.e. punishments, rewards) are in place. Two individuals may assign different intrinsic value to the same task, and may respond differently to the same performance incentives — depending, in part, on their past experiences.

[18] Bernheim, 1994; Benabou and Tirole, 2003.

In line with this reasoning, we argue that repression induces "perfunctory" rather than "consummate" compliance.[19] People who learn first hand about the repressive nature of their state will come to expect that they — or even their families, as in the Soviet case — may be punished for the slightest hint of disloyalty. These individuals are more likely to comply with formal orders and collective norms, even at their own peril, but they do so not to consummate victory, but to avoid punishment.

The link between repression and war-making is not unique to the Soviet case. Similar tensions between pre-war state violence and wartime combat mobilization have surfaced in the Iran-Iraq War,[20] the Second Congo War[21] and the civil war in Syria[22] among others. Yet several properties make the Soviet case especially suitable for empirical analysis. Highly bureaucratized Soviet military administration generated enormous amounts of granular data, permitting quantitative study of the largest-ever military effort at the level of individual soldiers. Furthermore, due to a nearly universal draft of the adult male population, we can avoid problems of self-selection into the military. Finally, the Soviet case allows us to partial out two factors central to earlier literature on combat motivation: unit cohesion and pecuniary incentives. Personnel turnover was too fast — due to conscription and combat losses — to secure the types of inter-personal bonds documented in other armies;[23] the Red Army offered no material inducements for combat performance.[24]

Most directly, this article contributes to political economy and international relations literature on the micro-foundations of military effort.[25] While recent studies have investigated whether wartime coercion can deter desertion and flight,[26] we focus on repression

[19] Brehm and Gates, 1999, 17.
[20] Pollack, 2004, 182.
[21] Lyall, 2020, 332.
[22] Heydemann, 2013.
[23] Merridale, 2006.
[24] Soviet veterans received temporary benefits to assist with reintegration into civilian life (e.g. easier access to higher education) during mass demobilization in 1945. Other veterans' benefits (e.g. interest-free loans for housing construction, travel discounts) came into effect in the 1970s (Edele, 2006).
[25] Hirshleifer, 1989; Skaperdas, 1996; Costa and Kahn, 2003; Berman, Shapiro and Felter, 2011; Lehmann and Zhukov, 2019; Lyall, 2020.
[26] Chen, 2017.

prior to war. Our results suggest that soldiers who had experienced this violence more intimately may be more responsive to coercive measures imposed by the state on the battlefield. This paper also relates to research on intrinsic motivations to fight for material or ideological reasons.[27] Our findings suggest that past experiences of repression can undermine intrinsic motivations, resulting in lower exerted effort by highly-motivated types.

More broadly, this article extends research on how states' exploitative, coercive, and violent practices impact later economic and human development,[28] social structure,[29] trust,[30] voting,[31] and identity.[32] We show how political repression can shape states' ability to provide the most basic public good — national security. This underscores a previously overlooked negative externality of repression: whether or not it helps rulers stay in office, repression may impede the state's ability to defend itself effectively from outside threats.

The rest of the paper proceeds as follows. We start with a stylized model of combat motivation to show how the shadow of state repression might incentivize conformity in battle. We introduce our data and empirical strategies to identify the causal effects of repression on soldiers' battlefield fortunes. We present our empirical estimates, and consider their robustness and limitations. Finally, we discuss a range of alternative interpretations of our findings and present auxiliary empirical tests to evaluate their plausibility.

2. Repression and Combat Motivation — Theoretical Expectations

To generate structured predictions of how prior experiences of repression may affect battlefield behavior, we use a stylized model of soldiers' choice. The model explores the conditions under which individuals with heterogeneous preferences will conform to a homogeneous standard of behavior,[33] like obeying orders from commanding officers. We

[27] Barber IV and Miller, 2019; Grossman, 1991; Hall, Huff and Kuriwaki, 2019.
[28] Dell, 2010.
[29] Acemoglu, Hassan and Robinson, 2011.
[30] Nunn and Wantchekon, 2011; Grosjean, 2014.
[31] Rozenas and Zhukov, 2019.
[32] Blaydes, 2018.
[33] Bernheim, 1994; Kreps, 1997.

assume that individuals fight for reasons both intrinsic (e.g. honor, patriotism, duty) and extrinsic (e.g. fear of punishment). Repression can affect individual choices through two channels — alienation (reducing intrinsic motivations) and deterrence (increasing extrinsic ones). We show how these countervailing forces increase obedience among lower-motivated soldiers, while reducing the willingness of higher-motivated soldiers to take initiative beyond the stipulated mandate.

Suppose that each soldier must choose an action $a \in \mathbb{R}$, which measures one's observed level of battlefield resolve — defined as a willingness to continue fighting despite temptations to back down. More resolve implies a higher risk of death or injury. Let $\overline{a} \in \mathbb{R}$ denote the action ordered by commanders (e.g., "charge!"). The soldier could obey the order by choosing $a = \overline{a}$, or he could shirk by, for example, hiding in the trenches ($a < \overline{a}$). For concreteness, let $\underline{a}$ denote a cutoff such that if $a < \underline{a}$, the soldier shows especially low resolve by surrendering or deserting. Alternatively, the soldier could take initiative by going beyond what one's orders require ($a > \overline{a}$), such as personally capturing an enemy officer or continuing to carry out one's mission after being wounded.

Past research suggests that exposure to state violence reduces trust in state institutions[34] and, more generally, incites negative sentiments toward perpetrators of violence.[35] In our setting, this means that individuals who experienced state violence more intimately — personally, or indirectly through their family or community — should be intrinsically less motivated to fight for the state. This is the "alienation effect" of repression.

Formally, let $\omega \in \mathbb{R}$ represent a soldier's baseline intrinsic motivation. In the population of soldiers, ω is drawn from a distribution F, which we assume to be continuous with full support and increasing everywhere. Prior experiences of repression reduce the soldier's baseline intrinsic motivation to $\omega - \alpha r$, where $r \geq 0$ denotes repression and $\alpha > 0$ is the alienating effect. Action a results in an intrinsic loss of $(a - (\omega - \alpha r))^2$. In the absence of other considerations, the soldier would choose $a = \omega - \alpha r$.

[34] Nunn and Wantchekon, 2011; Grosjean, 2014.
[35] Dell and Querubin, 2018.

Like any other organization, the military relies on a range of extrinsic incentives. A soldier who shirks by retreating or hiding when ordered to charge can be demoted, tried by military tribunal,[36] or shot by a blocking detachment.[37] In the Soviet case, desertion or even surrender may further result in the punishment of the soldier's family members. To the extent that repression creates "internalized expectations about [how] authority will respond punitively to challenging acts,"[38] a soldier exposed to state repression in the civilian domain will see the state's threat of punishment on the battlefield as more credible. This is the "deterrent effect" of repression.

To capture the deterrent effect, suppose that a soldier who shows *less* resolve than required by commanders suffers an extrinsic loss equal to $\mathbb{E}(\pi|r)(\overline{a} - a)^2$, where $\mathbb{E}(\pi|r)$ is the penalty that a soldier expects to receive by showing less resolve than asked. We assume that $\mathbb{E}(\pi|r) = \delta r$, where $\delta \geq 0$ is the "deterrence" parameter: a soldier exposed to more repression will infer that the incumbent's willingness and capacity to punish is higher and will be more reluctant to shirk ($\underline{a} < a < \overline{a}$), let alone defect ($a < \underline{a}$).

We are interested in how the alienating and deterrent effects of repression concurrently influence the distribution of battlefield resolve in a population of soldiers. Each soldier chooses an optimal action that minimizes the sum of his intrinsic and extrinsic losses:

$$a^* \in \arg\min_{a \in \mathbb{R}} (a - (\omega - \alpha r))^2 + \delta r (a - \overline{a})^2 \mathbb{1}\{a < \overline{a}\}, \tag{1}$$

which solves to

$$a^*(\omega, r) = \begin{cases} \dfrac{\omega + r(\delta\overline{a} - \alpha)}{1 + \delta r} & \text{if } \omega \leq \overline{a} + \alpha r; \\ \omega - \alpha r & \text{otherwise.} \end{cases} \tag{2}$$

The optimal action a^* is increasing in repression r for soldiers with low baseline mo-

[36] Chen, 2017.
[37] Statiev, 2010.
[38] Beissinger, 2002, 326.

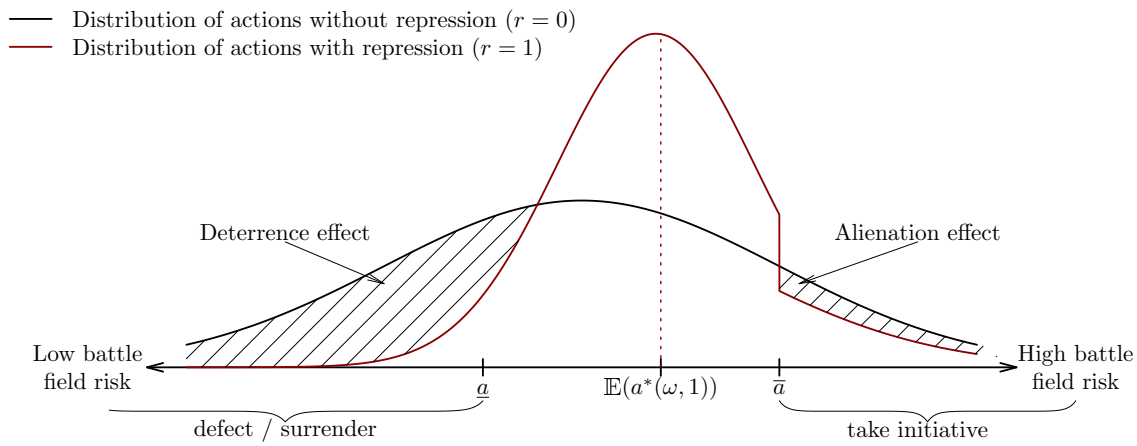


Figure 1: REPRESSION AND CONFORMITY ON THE BATTLEFIELD

tivation, $\omega \leq \overline{a} - \alpha/\delta$, and it is decreasing for soldiers with high baseline motivation $\omega > \overline{a} - \alpha/\delta$. The deterrent effect of repression dominates its alienating effect for soldiers with low intrinsic motivation (pushing a^* up), whereas for soldiers with high intrinsic motivation, the alienating effect dominates the deterrent one (pulling a^* down).

The following proposition stipulates that, as long as commanders order soldiers to show a sufficiently high level of resolve $\overline{a}$, repression leads to more uniform combat effort.

Proposition 1. *For each $\alpha > 0$ and $\delta > 0$, there is $\tilde{a}(\alpha, \delta)$ such that, if $\overline{a} \geq \tilde{a}(\alpha, \delta)$, then $\mathbb{E}(a^*(\omega, r))$ is increasing everywhere in r whereas $\Pr(a^*(\omega, r) < \underline{a})$ and $\Pr(a^*(\omega, r) > \overline{a})$ are decreasing everywhere in r.*

Figure 1 illustrates the mechanics behind the proposition (see Appendix A1 for the proof). First, the average level of resolve soldiers display on the battlefield, $\mathbb{E}(a^*(\omega, r))$, shifts to the right as a result of repression. To the extent that higher resolve induces more physical danger, repression should increase fatalities and injuries on average. Second, in addition to shifting the mean of the distribution of soldiers' resolve, repression also reduces its variance. Soldiers with low baseline motivation exhibit higher resolve, resulting in fewer cases of $a^* < \underline{a}$ (deterrent effect), and soldiers with high initial motivation exhibit lower effort than they would otherwise, with fewer cases of $a^* > \overline{a}$ (alienation effect).[39]

[39] While based on different micro-foundations, the latter result is similar to the well-known finding that

In conclusion, the net effect of repression is conformity: it increases extrinsic motivation to fight, but it saps the intrinsic motivation to take initiative above and beyond the formal mandate. Insofar as documented battle-field outcomes capture the relevant dimensions of combat motivation (which we show they do; see Section 4), we can draw the following testable predictions: as soldiers' exposure to repression increases, they become (1) more likely to be killed or wounded in action, (2) less likely to surrender, defect or go missing, but also (3) less likely to receive a medal for sacrifices beyond the call of duty.

3. The Context – Soviet Repression and War Effort

The Soviet legal code established a broad class of "counterrevolutionary" crimes, including treason, insurrection, espionage, contacts with foreign states, propaganda, agitation, and a failure to report any of the above. Between 1927 and 1953, around 3.8 million people were convicted of such crimes, most of them in the mid-late 1930's.[40]

State repression sought to eliminate "anti-Soviet elements," but the regime had few means to identify those who engaged in "counterrevolutionary" activities or held "anti-Soviet" views. Stalin ordered security officials to target broadly-defined segments of the population, like residents of particular provinces, minorities and peasants, without discriminating between individuals inside those demographic categories. Stalin directed his subordinates to cast a wide net: "because it is not easy to recognize the enemy, the goal is achieved even if only five percent of those killed are truly enemies."[41] Eventually, people from all demographic and professional groups, including members of the Communist Party, the military, and security agencies, "everybody from the Politburo member down to the street cleaner," became victims of collective targeting.[42]

Central authorities provided little concrete guidance as to who should be repressed. Moscow issued numerical quotas of persons to be executed or sent to camps in each re-

extrinsic incentives can undermine intrinsic motivation (Kreps, 1997; Benabou and Tirole, 2003).

[40] GARF, 1954.

[41] Gregory, 2009, 196.

[42] Ulam, 1973.

gion and "everything else depended on the ingenuity of Security operations personnel."[43] Local executives often engaged in "exceptional competition" to exceed their quotas and signal administrative or ideological zeal.[44] The hard constraints on this competition were largely circumstantial: the need to cover transportation costs for those condemned to the camps, and to find "a place [to] bury the corpses" for the rest.[45]

The blanket targeting of a wide cross-section of societal groups and arbitrary victimization of individuals within those groups created a perception that repression was largely random. Asked about how authorities decide whom to incarcerate or release, one NKVD officer explained, "Chance. People are always trying to explain things by fixed laws. When you've looked behind the scenes as I have you know that blind chance rules a man's life in this country of ours."[46]

After the German invasion on June 22, 1941, the Soviet Union drafted all military-age males – over 30 million civilians throughout the war – to support its 4.5 million-strong standing Red Army. Thus, the backbone of the Soviet defense against the Germans were ordinary citizens. The war became an "acid test" for Stalinism: would the people risk their lives for a regime that only recently had terrorized them?[47]

In some respects, Stalinism passed the test: millions of soldiers fought for the Soviet state, very often to death. Early in the war, the Red Army stumbled spectacularly due to prewar officer purges, politicized decision-making, and chronic mismanagement.[48] Ultimately, the Soviet Union won the war, and it did so largely by keeping its troops fighting *despite* organizational malaise and devastating human costs.[49] At the same time, many soldiers voted against Stalinism with their feet. Half of all personnel losses in the war's first year comprised soldiers missing in action or captured. Thousands were detained

[43] Solzhenitsyn, 1973, 69-70.
[44] Chukhin, 1999, 76.
[45] Jansen and Petrov, 2002, 86, 88.
[46] Conquest, 2008, 434.
[47] Thurston, 2000.
[48] Glantz, 1998.
[49] Reese, 2011.

for desertion, sabotage, or treason. While there were many reasons to flee the battlefield, widespread distaste for how the Soviet state treated its citizens clearly did not help.[50]

Moscow took draconian measures to hold its troops in line. On August 16, 1941, Stalin issued Order 270, stipulating that those "who surrender to the enemy shall be considered malicious deserters, whose families could be arrested."[51] Commanders were to prepare bi-weekly reports for the General Staff, listing captured soldiers and their families' addresses.[52] Among the first victims of this order was Stalin's own son, Yakov, whose wife was sent to a labor camp after his capture by the Germans.

Stalin issued another disciplinary measure, Order 227, on July 28, 1942. It required every front to organize "penal units" staffed by men accused of disciplinary problems, and send them to the most dangerous sectors to "atone for their crimes against the Motherland with their blood."[53] The order also mandated the creation of "blocking units" under NKVD command, authorized to detain or execute retreating personnel.

Given these incentives, soldiers who fought instead of fleeing might have done so because they were intrinsically committed to the cause or because they feared that they or their families would be punished. Soldiers who witnessed state repression prior to war may have been especially sensitive to coercive incentives. They may also have been less eager to defend the regime in the first place. To assess how these countervailing pressures affected battlefield choices, we must consider how these citizen-soldiers actually fought.

4. Data and Measures

We draw on data from military personnel records, NKVD arrest records, and contextual data from geo-referenced historical atlases and other sources.

[50] Edele, 2017.
[51] Zolotarev, 1997, 58-60.
[52] Kachuk, 2013.
[53] Statiev, 2010, 726.

4.1. Military records

Our source of information on battlefield outcomes is the Russian Ministry of Defense's "People's Memory" (*Pamyat' Naroda*) database, which contains over 106 million declassified Red Army personnel records. The database includes 21 million records on irrecoverable losses and discharges, 23 million records from military transit points, 10 million registration cards, 1.3 million POW records, 5 million burial and exhumation records, 27 million decoration records, and 425,000 combat logs and staff documents. Aside from basic biographical information, these data record combat unit details (recruiting station, enlistment date, unit, rank), decorations, and the reason and date of soldier's discharge.

Due to illegible handwriting, errors in optical character recognition, abbreviations, misspellings, incomplete or missing fields and other errors that are inevitable in archival data, these records required significant preprocessing. This included, among other things, homogenizing names of military ranks and thousands of military units, assigning tactical units to parent divisions, corps and armies, and standardizing geographic references.

The most challenging aspect of data preprocessing was record classification. The same soldier can appear in multiple records and there are no fields to match soldiers across them. This is an unsupervised classification problem, where each of 106 million records must be assigned to one of an unknown number of soldiers. We approached this problem with a probabilistic record linkage method,[54] which we tailored to be operable with our data. Appendix A2 details this procedure and the evidence validating its output.

Since we measure soldiers' exposure to repression through birth locations (see below), we excluded soldiers whose birthplaces were missing or could not be geocoded to the municipality level or lower. We also excluded soldiers born outside the territory of Soviet Russia (RSFSR), since some of these areas were not part of the Soviet Union prior to the war. Our final dataset contains 26,922,385 records for 11,680,930 soldiers.

[54] Enamorado, Fifield and Imai, 2019.

4.2. Measures of Combat Motivation

We construct several proxy measures of combat motivation based on how soldiers were discharged from the army and the decorations they received. Before the war ended, 46% of soldiers were discharged because they were killed or wounded in action (K/WIA), missing in action (MIA), became prisoner of war (POW), deserted, defected or committed treason (DDT), or were punished for misconduct (PUN). 25% had received at least one medal, including 17% who received decorations specifically for personal valor in combat.

We use K/WIA as a measure of soldiers' resolve to fight (a^* in the language of the model). The idea here is that soldiers who followed orders and fought instead of fleeing faced a mechanically greater risk of death or injury. In combat, performing one's duties means engaging in actions that can kill or physically harm others. The cumulative probability of receiving a severe traumatic injury — due to hostile action or accident — rises as one remains on the battlefield for longer periods of time.

The second outcome of interest is whether a soldier displayed low combat motivation by fleeing ($a^* < \underline{a}$). Soldiers who defected, deserted, or committed treason (DDT) fall into this category, as do those who were punished (PUN) for insubordination. Although less clear-cut, another indicator of flight is whether a soldier became a POW. Not all POWs had made an individual choice to surrender; some did so on the orders of commanders. In the Soviet system, however, orders to surrender were illegal and soldiers were instructed to disobey them, even if they lacked the physical means to resist detainment. Stalin's Order 270, which equated surrender with treason, stipulated that "every soldier is obliged ... to demand that their superiors, if part of their unit is surrounded, fight to the end."

To avoid being held personally responsible, Red Army officers were reluctant to report unaccounted-for soldiers as DDTs or POWs. The common wartime practice was to report them as MIA, as an official from Russia's Ministry of Defense recently acknowledged:

> By official reports, out of our five million-plus missing in action just 100 thou-

sand were reported as prisoners of war. In reality, there were 4.5 million. So the majority of those missing in action were prisoners of war. Everyone knew this. I'm certain that even Stalin knew.[55]

Accordingly, we treat MIA as another indicator of flight. This is a noisy measure, since some cases of MIA must have been KIA. However, this kind of measurement error is more likely to attenuate than inflate our estimates, and is unlikely to be numerically large — the quoted estimates suggest that $\Pr(\text{POW}|\text{MIA}) = 4.5\text{M}/5\text{M} = 0.9$.[56]

Finally, to capture cases of initiative above and beyond one's formal duty ($a^* > \underline{a}$), we consider whether a soldier received at least one decoration for acts of valor in combat. These could include — in order of prestige — the medals For Courage, For Battle Merit, Order of Glory, and Hero of the Soviet Union. In contrast to other Soviet medals, which were awarded en masse, these decorations recognized individual performance in situations involving a risk to life, and had to be justified with detailed descriptions of individual acts.[57] Appendix A3 provides examples.

To assess how well these measures map onto the theoretical concept of combat motivation, we examine their correlation with the Red Army's effectiveness across battles. Using official descriptions of 225 major battles from the Defense Ministry's "People's Memory" database, we classified each military operation as resulting in a territorial gain, loss or no change for Soviet forces. For each unit participating in a battle, we calculated monthly proportions of soldiers who were K/WIA, DDT, PUN, POW, MIA, or received one of the four valor decorations (Medal).[58] Using these unit-month level data, we estimated a regression equation where the dependent variable is a dummy equal to one if the battle resulted in territorial gain and the covariates are unit-month proportions of K/WIA, DDT,

[55] https://www.newsru.com/russia/04feb2011/stalin.html

[56] According to Krivosheev (1997)'s more conservative numbers, of 4.6M designated MIAs, 0.5M were "true" MIA's, 1M returned to the front, and 3.1M were POWs, implying $\Pr(\text{POW}|\text{MIA}) = 0.67$.

[57] We also exclude career service awards and hybrid medals like the Order of the Patriotic War, which was awarded on both an individual basis (e.g. for a high number of enemy kills) as well as collectively to units, towns, factories and entire categories of veterans.

[58] Appendix A5 explains the procedure we used to classify battles and match them to army units.

	Dependent variable: territorial gain
Proportion KIA	$0.05\ (0.02)^{*}$
Proportion MIA	$-0.16\ (0.04)^{**}$
Proportion POW	$-0.25\ (0.08)^{**}$
Proportion DDT	$-0.50\ (0.16)^{**}$
Proportion PUN	$-0.06\ (0.06)$
Proportion Medal	$0.05\ (0.02)^{*}$

OLS coefficients and standard errors (in parentheses) clustered by unit and battle. Unit of analysis is military unit by month (N = 46,240). Fixed effects for units, years, and months are included. Significance levels: $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$

Table 1: SOLDIERS' BATTLEFIELD OUTCOMES AND ARMY UNIT PERFORMANCE

PUN, POW, MIA, and Medal. The regression also includes fixed effects for units, years, and months, to partial out unit-level factors as well as temporal and seasonal trends.

The results in Table 1 suggest that the aggregate success of army units correlated positively with higher casualty rates and higher rates of medals, and negatively with all measures of flight — MIA, POW, DDT, and PUN (although the latter coefficient is not significant). To the extent that combat motivation contributes to higher operational effectiveness, Table 1 suggests that our proxy measures do capture the relevant latent quantity.

Figure 2 shows the geographic distribution of soldiers' birthplaces, with lighter colors representing higher proportions of soldiers who fought (K/WIA), and darker colors representing higher proportions that fled (MIA/POW/DDT/PUN). There is no macro-level spatial trend in the combat motivation as far as we can discern from these observables. Soldiers born near the western border — where much of the fighting took place — were no more likely to fight or flee than those born in Siberia. There is, however, substantial *local* variation in relative rates of fight and flight, which requires micro-level explanation.

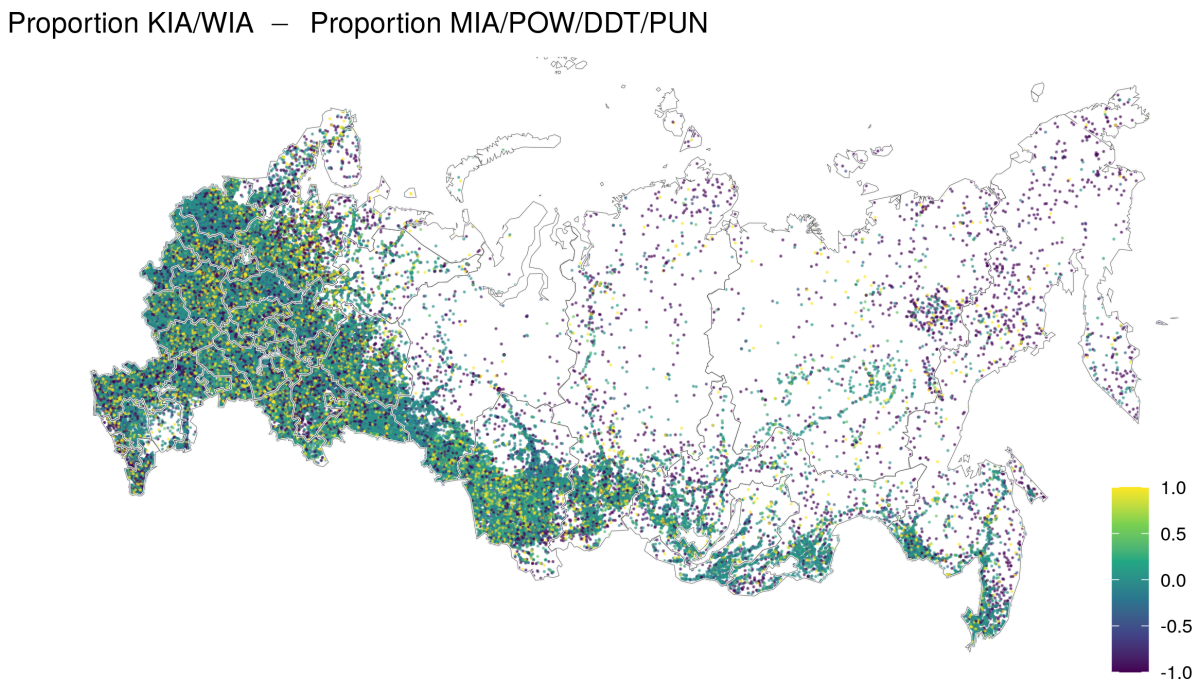


Figure 2: BATTLEFIELD OUTCOMES OF RED ARMY SOLDIERS, BY BIRTH LOCATION

4.3. Data on repression

Our data on repression come from the "Victims of Political Terror" archive, maintained by the Russian human rights organization Memorial.[59] This archive contains individual records for those arrested under Article 58 ("counterrevolutionary activity") of the Soviet penal code. It draws on declassified Ministry of Interior documents from federal, ministerial and regional archives, prosecutor's offices, and the Commission for the Rehabilitation of Victims of Political Repression, with supplementary information from newspapers, regional NGOs, "Memory Books," and survivors' families.

These data are limited in scope to individual prosecutions for alleged political dissent. This includes 70% of the 3.8 million convictions under Article 58, but excludes millions of victims of famines, deportations, and counterinsurgency operations. Using the same approach as with military data, we found geographic coordinates for 2.15 million pre-

[59] Memorial, 2014.

WWII arrests (74%), using victims' residential addresses (where available) or birthplaces.

We measure exposure to repression by counting the number of arrests in the vicinity of a soldier's birth location. Specifically, the variable we use in the analysis is:

$$\text{Repression} = \log(1 + \text{Arrests within 10 km of birth location}),$$

where the logarithmic transformation is used to reduce skewness.[60] Our geographic measure of repression rests on the idea that people are more aware of the repression in their home communities than in more distant locations. With the exception of elite show trials, political repression against ordinary citizens was not publicized, and people learned about the actions of the state mostly through family, neighbors, friends, or co-workers.

One concern in using birth locations is that some soldiers may have moved away before repression occurred. We can take stock of this issue by examining the distribution of travel distances between birth locations and the 1,869 military commissariats where soldiers were drafted. The median soldier was born 88km from his draft location; 20% were born within 1km, and 80% within 335km. If most soldiers remained in their areas of birth long enough to be drafted, they were likely also around for the terror of the 1930s.

We use absolute rather than per capita numbers of arrests because this is how narratives about state violence are typically framed and memorialized. 60 arrests (our sample median) from a town of 1,000 are unlikely to be perceived as two times more "repressive" than from a town of 2,000. Indeed, Soviet state security records, historical and autobiographical narratives measure the scale of repression exclusively in absolute numbers.

At the same time, we must ensure that our measures of repression are not conflated with local population density. The Soviet censuses of 1926, 1937, 1939 do not provide information on population below the district level.[61] However, we adjust for several proxies of local population density (distance to administrative center, road junctions, farmland).

[60] Analyses with alternative bandwidths (1-20km) do not produce major differences (Appendix A7.6).

[61] District-level geographic precision allows us to estimate population counts for small areas (e.g. grid cells, see footnote 74), but not point estimates for specific birth locations.

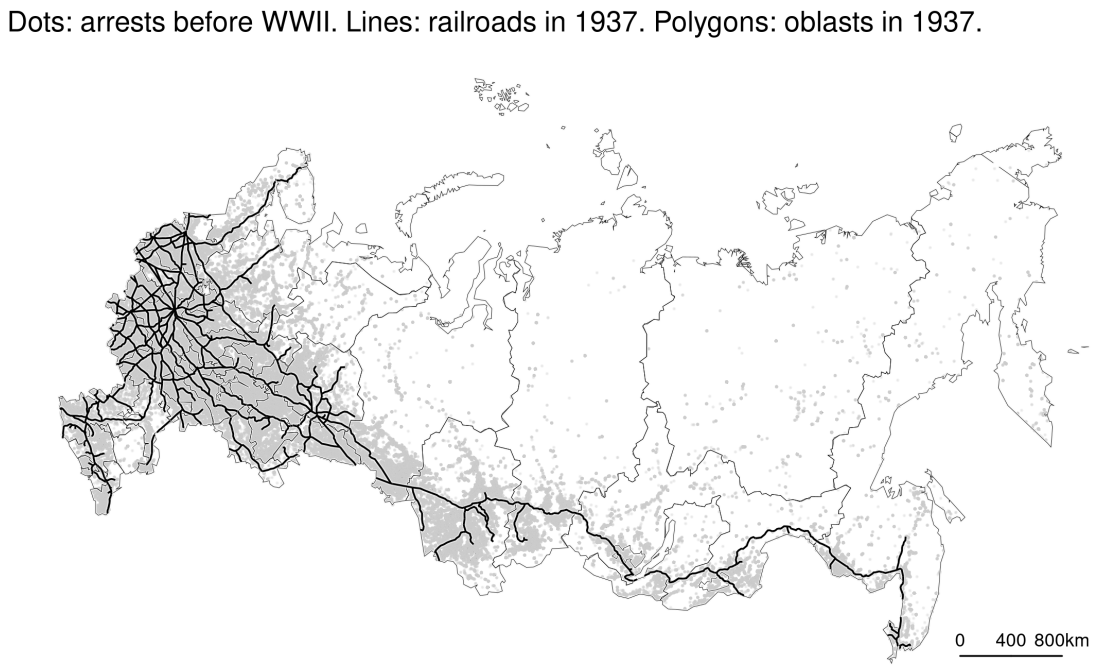


Figure 3: GEOGRAPHY OF PRE-WWII SOVIET REPRESSION

In addition, we implement a matched cluster sampling design that selects pairs of locations that are as similar to each other as possible on observables, including the number of soldiers drafted as a proxy for local population size (Appendix A7.1). Most conservatively, we replicate our results at an aggregate, district level, directly controlling for local population size and urbanization from the 1926 census (Appendix A7.4).

Figure 3 shows the geographic distribution of Soviet repression. Although not as widespread as the distribution of soldiers' birth locations, arrests affected every region of the country. Many arrests were concentrated around the main railroad network (black lines), although remote northern regions also did not fully escape the NKVD's reach.

4.4. Additional data

We collected additional data on local political economy, logistics and ethnicity. We measure state capacity using the distance in 1935 from each birthplace to the nearest dis-

trict administrative center, where local NKVD branches were based.[62] We geo-referenced maps of economic activity from the 1937 *Large Soviet Atlas of the World*,[63] which provided us with information on several important variables. To distinguish between urban and rural areas, we calculated hectares of cropland within 10km of each birthplace. To account for the targeting of peasants due to collectivization, we counted the number of state farms within 10km of each birthplace in 1937.[64] We also augmented these maps with georeferenced information on historical railway junctions, stations, and tracks,[65] which we use to construct an instrument for repression (see details in next section).

With few exceptions, the military records do not include information on soldiers' ethnicity. This is problematic, because ethnic minorities were more likely to face repression. To address this issue, we build a nationality classifier for soldiers' surnames. Using the Memorial archive, which contains nationality information for 916,675 arrestees with 163,284 unique surnames, we trained a Support Vector Machine (SVM) classifier to identify whether a surname represents Russian nationality. The algorithm achieved 96.5% out-of-sample predictive accuracy. We then assigned to each military personnel record a dummy variable equal to one if the surname is predicted to be of Russian nationality.[66]

5. Empirical Approach

We employ three empirical strategies: small area fixed effects, instrumental variables and regression discontinuity.

[62] TsIK, 1935.
[63] Gorkin et al., 1937, 155.
[64] Gorkin et al., 1937, 161.
[65] Afonina, 1996.
[66] For surnames that do not appear in the training data, we assigned the predicted nationality of the surname that is closest in Jarro-Winkler string distance. We compared oblast-level proportions of Russians against census data from 1939. Wilcoxian rank-sum tests suggest that our SVM-classified oblast-level proportions were drawn from the same distribution as oblast-level census proportions (Appendix A4).

5.1. Ordinary Least Squares with Grid Cell Fixed Effects

Our first empirical strategy is motivated by the observation that, at the local level, Stalin's terror was notoriously arbitrary in its selection of targets. Since the regime used group-level targeting, we can treat exposure to repression as plausibly exogenous given the observables that the regime itself used in selecting victims. One such observable was ethnicity: Soviet authorities often viewed national minorities as politically disloyal and subjected them to greater coercion. Another was socio-economic: the regime saw *kulaks* ("rich" peasants) as an obstacle to collectivization, but defined "kulak" so loosely that most rural residents faced a heightened risk of coercion. A third was geographic: regions in the western borderlands, the Far East, and areas with a history of peasant uprisings against collectivization faced higher arrest quotas.[67]

Let y_i denote a battlefield outcome for soldier i and let $\text{Repression}_{j[i]}$ denote repression around the birth location j of soldier i. We fit the following OLS regression:

$$y_i = \gamma \cdot \text{Repression}_{j[i]} + \boldsymbol{\beta}' \boldsymbol{X}_{ij} + s\big(\text{lon}_{j[i]}, \text{lat}_{j[i]}\big) + \text{Cell}_{k[i]} + \epsilon_i. \tag{3}$$

The vector $\boldsymbol{X}_{ij}$ contains individual-level covariates (ethnicity and year of birth) as well as location-level covariates, including hectares of cropland and the number of state farms within 10 km of soldier's birth location (to account for higher repression of peasants), distances to the nearest administrative district center and nearest road junction as proxies for local state capacity and population density. The term $s(lon, lat)$ represents a two-dimensional spatial spline, which we include to capture local geographic trends.

To account for higher targeting of certain administrative regions (*oblasts*), it would suffice to include fixed regional effects. But even within-regional comparisons would involve locations that potentially differ on unobserved background characteristics. To ensure more balanced comparisons, we partition 1937 Soviet Russia's administrative regions into

[67] Getty and Naumov, 1999.

a regular 25×25km grid, and include a fixed effect for the grid cell k in which soldier i was born. Because geographically proximate locations tend to share background characteristics like population density, ethnic and socio-economic composition, these small area fixed effects should balance the unmeasured confounders. They also ensure that our inferences are drawn by comparing birthplaces no more than $\sqrt{25^2 + 25^2} \approx 35$km apart.[68]

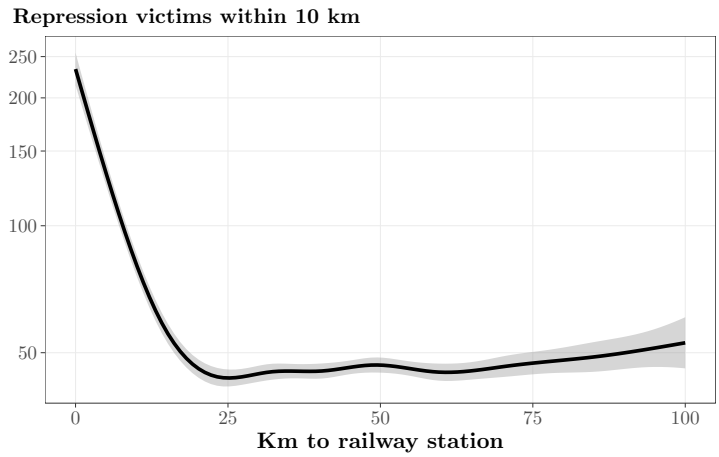
The underlying assumption behind this design is that exposure to repression is *locally* exogenous. We evaluate this assumption by testing whether the geographic distribution of arrest locations within grid cells was spatially random. Within each 25×25km cell, we tested the null hypothesis that arrest locations are the realization of a uniform Poisson point process. We were unable to reject this hypothesis in 87-96% of cells, depending on the test procedure (Appendix A6.1). Although arrest density varies between cells and regions, the *local* spatial distribution of repression appears quite arbitrary.

5.2. *Railway Access as Instrument*

Even if repression is exogenous on a small geographic scale, OLS estimates may suffer from attenuation bias due to errors in the measurement of repression through archival sources. To correct for such bias, we use two-stage least squares (2SLS). This approach exploits the fact that Stalin's repression was an industrial-level operation. Arrestees were transported to execution sites, prisons, and distant labor camps in large numbers and on short deadlines. The primary means for implementing these operations were railways.[69] A third of the Great Terror campaign's budget was earmarked for rail transport fees.[70]

Motivated by these facts, we use access to railways, measured as the distance from a birth location to the nearest railway station, as an instrument for repression. The idea here is that otherwise similar locations may be exposed to varying levels of repression due to differing costs of accessing and transporting arrestees. One concern with this instru-

[68] The median (maximum) distance between two birth locations in a grid cell is 8.9 km (16.9 km).
[69] Kokurin and Petrov, 2000, 525.
[70] Getty and Naumov, 1999, 478.



The estimated function $\hat{f}$ with 95% confidence bounds relating railway access to repression, adjusted for geographic covariates and grid cell fixed effects. Vertical axis is on logarithmic scale.

Figure 4: RAILWAY ACCESS AND REPRESSION

ment is that it may be capturing economic development and population density. All our 2SLS estimations include distance to the nearest administrative center and nearest road junction, which approximate local development and density more directly than railways. Indeed, the Soviet railway system was built not to help foster local economic development or connect population centers, but to help access resource-rich areas.[71]

To test whether birthplaces with better railway access saw more repression, all else equal, we fit the following semi-parametric regression:

$$\text{Repression}_j = f(\text{Raildist}_j) + \boldsymbol{\beta}'\boldsymbol{X}_j + \text{Cell}_{k[j]} + s(\text{lon}_j, \text{lat}_j) + \epsilon_j, \tag{4}$$

where j indexes birth locations, $\textit{Raildist}_j$ is distance from location j to the nearest railway station, and f is a smooth function approximated by cubic regression splines. As before, we add grid cell fixed effects, location level covariates, and a spatial spline. To ensure greater homogeneity, the 2SLS analyses use only locations within 100 km of rail stations.

Figure 4 shows a graph of the estimated function $\hat{f}$. The expected number of repression victims declines precipitously with distance to rail stations, even after accounting

[71] Hopper, 1930.

for local road density, distance to administrative centers, and other covariates. A 10km higher proximity to a railway station increases the number of victims by a factor of two.

Note that we estimate function f at the level of birth location, not individual soldier, because this is the level at which the relationship between railway access and repression operates. Specification (4) helps us find an optimal transformation f of $Raildist_j$ that yields the strongest *linear* first stage relationship. In the 2SLS regression specified at the level of a soldier, we use the variable $\hat{f}(\text{Raildist}_{j[i]})$ as the instrument. The first stage model is

$$\text{Repression}_{j[i]} = \alpha \cdot \hat{f}\left(\text{Raildist}_{j[i]}\right) + \boldsymbol{\beta}'\boldsymbol{X}_{ij} + \text{Cell}_{k[j]} + s(\text{lon}_j, \text{lat}_j) + \epsilon_i, \tag{5}$$

where $\boldsymbol{X}_{ij}$ includes both location-level and soldier-level covariates. In the second stage, we regress wartime individual outcomes on the predicted values of repression from (5).

The exclusion restriction behind our 2SLS strategy is that railway access impacted the future behavior of soldiers only through repression, and not some other channel outside the included covariates. One reason to doubt this assumption is that railways played a key role in the war effort: the front stretched thousands of kilometers from the Baltic to the Caspian Sea and motorized vehicles could only support operations up to 400 km.[72] However, only a small fraction of RSFSR's railroad network fell inside areas of active military operations or behind German lines: 3.7% in an average month, and 16% cumulatively at any point in the war. The railway structure also changed significantly in 1941-1945, as Soviet authorities built 6,700 km of new rail lines,[73] which are not part of the instrument.

Another potential violation of the exclusion restriction has to do with railroads' use in military mobilization. While almost all military-age males were drafted, it is possible that someone living near railways faced different battlefield conditions by virtue of being drafted in the chaotic early months of the war, when incentives to flee and the odds of being killed were highest. However, the proximity of one's birthplace to railroads does

[72] Davie, 2017.
[73] Zickel, 1989, 552.

not relate systematically to the timing of conscription (see Appendix A6.2). Furthermore, we show that our results hold even if we compare soldiers who served concurrently in the same military unit, and who therefore faced similar battlefield conditions.

5.3. Geographic discontinuities

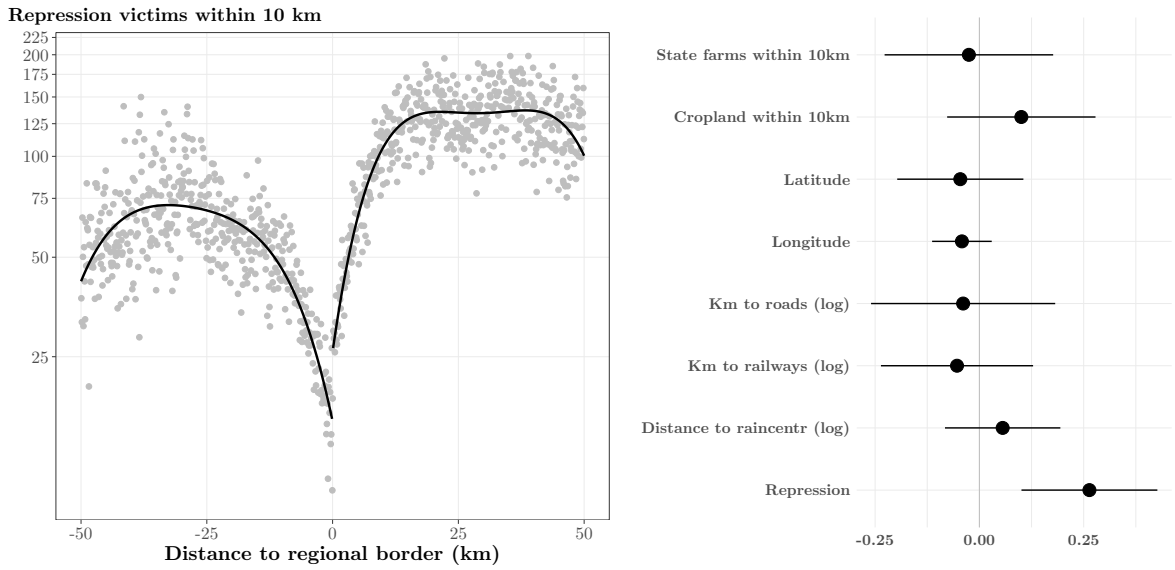
Our last empirical strategy utilizes the fact that regional state security officials practiced enormous discretion when implementing central orders. A town located in a region with a zealous NKVD chief could face significantly more repression than a nearby town from a different region with less ambitious or cruel security officials. The idiosyncratic qualities of local officials cannot be measured directly, but we can infer which Soviet administrative regions (*oblasti*) had lower or higher than expected numbers of victims. We first identify regions where the number of arrests fell below or above what one would expect conditional on local population size and urbanization using the following linear regression:

$$\text{Repression}_{rk} = \alpha + \beta_1 \cdot \ln\big(\text{Population}_{rk}\big) + \beta_2 \cdot \text{Urbanization}_{rk} + \epsilon_{rk}, \tag{6}$$

where r indexes region and k indexes grid cells.[74] After estimating the above regression, we calculate the average residual $\bar{\epsilon}_r$ for each region r. When $\bar{\epsilon}_r$ is positive (negative), repression in region r is above (below) what its background characteristics would predict.

We then select pairs of adjacent regions (r, r') such that $\text{sign}(\bar{\epsilon}_r) \neq \text{sign}(\bar{\epsilon}_{r'})$, that is, where one region has higher than expected and another has lower than expected levels of repression. Let d_{jr} denote the distance from birth location j in region r to the border of

[74] We use 1926 Soviet Census data on local population and urbanization. To disaggregate district-level data to smaller grid cells, we used dasymetric spatial interpolation, which employs ancillary data to obtain filtered area-weighted local estimates (Mennis, 2003). We used historical land cover maps (Gorkin et al., 1937) to exclude uninhabitable areas (water, deserts, glaciers) and distinguish built-up and rural areas.



(a) Local means with local polynomial fit.

(b) Discontinuity effects with 95% CI's (normalized outcomes).

Figure 5: DISCONTINUITY OF REPRESSION AT REGIONAL BORDERS

the nearest region. The forcing variable δ_{jr} is constructed as follows:

$$\delta_{jr} = \begin{cases} d_{jr} & \overline{\epsilon}_r > 0 \text{ and } \overline{\epsilon}_{r'} < 0, \\ -d_{jr} & \overline{\epsilon}_r < 0 \text{ and } \overline{\epsilon}_{r'} > 0. \end{cases}$$

For example, if $\delta_{jr} = -2$, then birthplace j is inside a low-repression region two kilometers away from a high-repression region. Had the administrative border between regions r and r' curved slightly differently to include birthplace j in r' instead of r, the level of repression in j would have been higher, in expectation. This is a plausible counterfactual: Soviet regions underwent a series of territorial reforms, in which authorities subdivided large regions into smaller, more "manageable" units (*razkrupnenie*). The first phase of these reforms concluded in 1936, just prior to the Great Purge.[75]

Figure 5a plots the relationship between forcing variable δ_{jr} and predicted levels of repression. To preclude comparisons of wildly different locations, we restrict the analysis to birthplaces within ±50 km of regional borders. The figure shows a clear discontinuous

[75] Shiryaev, 2011.

jump across regional borders. The bias-corrected local-polynomial estimate of the discontinuity effect[76] is 0.52 (S.E. clustered by grid cells is 0.14) on the logarithmic scale, or about 10 victims on the natural scale.

To check if observables other than repression also change discontinuously across borders, we conduct balance tests. Figure 5b displays point estimates and 95% confidence intervals of discontinuity effects for eight covariates, normalized to have a standard deviation of one for comparability. Only repression shows a discontinuous jump, suggesting that border discontinuities are a plausibly exogenous source of variation in repression.

We exploit the discontinuities in repression across administrative borders using a fuzzy regression discontinuity design (FRDD):

$$\begin{aligned} \text{Repression}_{j[i]} &= \alpha \cdot \mathbb{1}\{\delta_{jr[i]} > 0\} + g_1\big(\delta_{jr[i]}\big) + \boldsymbol{\beta}'\boldsymbol{X}_{ij} + s(\text{lon}_j, \text{lat}_j) + \epsilon_{1i}, \\ y_i &= \gamma \cdot \text{Repression}_{j[i]} + g_2\big(\delta_{jr[i]}\big) + \boldsymbol{\beta}'\boldsymbol{X}_{ij} + s(\text{lon}_j, \text{lat}_j) + \epsilon_{2i}, \end{aligned} \tag{7}$$

where g_1 and g_2 are smooth functions of forcing variable δ_{jr} estimated using regression splines, and indicator $\mathbb{1}\{\delta_{jr} > 0\}$ is the instrument. Both stages include covariates and spatial splines, but exclude grid cell fixed effects because, by construction (cells are nested within regions), the instrument cannot vary within cells.

5.4. Clustering and Weights

The outcome variables in our study are measured the level of individuals, but we observe exposure to repression at the level of birth locations. Due to the potential correlation of errors across individuals from the same location (cluster), the effective sample size is bound to be smaller than the number of individual soldiers in the data. To account for this correlation of errors, we cluster standard errors by birth location, which is the level at which the treatment varies. We also cluster standard errors by grid cells to account for spatial

[76] Calonico, Cattaneo and Titiunik, 2015.

autocorrelation. Finally, to incorporate the uncertainty inherent in our procedure of classifying military records, we weigh soldiers by the geometric mean of pairwise matching propensities of records assigned to them (see Appendix A2).

6. Results

We first analyze how repression changed soldiers' resolve to fight, as measured by their propensity to be killed or wounded. Table 2 reports coefficient estimates from OLS with fixed effects, 2SLS with a railway instrument, and FRDD with a spatial instrument.

	OLS	2SLS	FRDD
Coefficient	0.4 (0.1)**	3.3 (0.6)**	1 (0.3)**
Mean Y	21.2	20.5	19
First Stage $\mathcal{F}$		146.6	25.3
Gridcells	12,143	5,654	1,582
Birthplaces	182,225	147,803	44,154
Soldiers	11,427,895	9,725,085	2,316,908

Outcome = killed or wounded in action (KIA/WIA), measured on percentage scale (0 to 100). Standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record clustering probability. 2SLS analyses exclude birth locations > 100km from railroad. FRDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 2: Repression and Resolve to Fight

Estimates from all three designs suggest that soldiers from areas with more repression were more likely to die in battle. In the OLS specification, increasing repression in the soldier's birthplace from zero to 32 people (first quartile in sample) meant a $[\ln(32+1) - \ln(0+1)] \times 0.4 \approx 3.5 \times 0.4 = 1.4$ percentage point higher chance of death or injury. In the 2SLS and FRDD specifications, the changes are 11.6 and 3.5 percentage points.

Next, we analyze how repression shaped soldiers' proclivity to flee the battlefield. Table 3 shows coefficients for a range of observables that capture this outcome. In the first column, the outcome is an index *Flee*, indicating whether a soldier was reported as either

	Flee	MIA	POW	DDT	PUN
Model			OLS		
Coefficient	-0.2 (0.1)**	-0.2 (0.04)**	-0.03 (0.04)	0.003 (0.002)	-0.01 (0.005)**
Mean Y	24	17.4	5.8	0.2	0.8
Model			2SLS (First-stage $\mathcal{F} = 146.6$)		
Coefficient	-2.3 (0.4)**	-2 (0.3)**	-0.3 (0.2)	0.03 (0.01)**	-0.1 (0.02)**
Mean Y	24.4	17.6	5.9	0.2	0.8
Model			FRDD (First-stage $\mathcal{F} = 25.3$)		
Coefficient	-0.6 (0.2)*	-1 (0.2)**	0.5 (0.1)**	0.02 (0.01)*	-0.03 (0.02)$'$
Mean Y	23.9	17.6	5.3	0.2	0.9

Outcomes on percentage scale (0 to 100): missing in action (MIA), becoming prisoner of war (POW), defecting, deserting, committing treason (DDT), being punished for battlefield misconduct (PUN), or any of the above (Flee). See the note under Table 2 for the number of soldiers, birthplaces, grid cells, and other details. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 3: REPRESSION AND FLIGHT FROM THE BATTLEFIELD

missing, surrendered, deserted, defected, committed treason, or punished for misconduct. Outcomes in the remaining columns are the index's four constitutive measures.

Coefficients in the first column are consistently negative and statistically significant at 95% confidence across all designs. Substantively, increasing repression from zero to the first quartile (32) reduced a soldier's chances of flight by $3.5 \times -0.2 = -0.7$ to $3.5 \times -2.3 = -8.1$ percentage points, depending on the specification. Coefficients for MIA, the most frequent indicator of flight, are also consistently negative and significant.

Other indicators offer more mixed results, from negative in the case of PUN to negative, null and even positive for POW and DDT. Some of this variability may reflect the idiosyncrasies of Red Army reporting. Most POW's, as we noted, were officially misreported as MIA's and cases of DDT were very rare, which raises the possibility that these outcomes emerged under qualitatively different circumstances or reflect underlying heterogeneities in repression's effect. We also find that positive coefficients on POW and DDT do not survive robustness tests, whereas coefficients for MIA and PUN remain negative and mostly significant (Appendix A7.5).

	OLS	2SLS	FRDD
Coefficient	-0.2 (0.1)**	-2.1 (0.4)**	-1.2 (0.2)**
Mean Y	17.9	18	18.1
First Stage $\mathcal{F}$		146.6	25.3

Outcome = receiving at least one valor decoration (*For Battle Merit*, *For Courage*, *Order of Glory*, *Hero of Soviet Union*), measured on percentage scale (0 to 100). See the note under Table 2 for the number of soldiers, birthplaces, grid cells, and other details. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 4: REPRESSION AND BATTLEFIELD INITIATIVE

Table 4 reports the estimated effects of repression on battlefield initiative, measured through a soldier's receiving at least one decoration for acts of valor. The estimated coefficients are negative and significant in all models. Using the most conservative estimate (OLS), a soldier from a place with 32 repression victims was $3.5 \times -0.2 = -0.7$ percentage points less likely to receive a medal than a soldier from a place with no repression.

One potential concern with these results is that 2SLS estimates and, in some cases, FRDD estimates are substantively larger than the OLS estimates. The differences could be due to sample selection, since we restricted 2SLS analyses to locations within 100 km of railroads and FRDD to $\pm$ 50 km of regional borders. However, sensitivity analyses (Appendix A7.7) show that OLS coefficients are nearly identical when we restrict the sample to locations near railroads and borders.

The relatively large magnitude of the 2SLS results may also indicate possible violations of the exclusion restriction. In Appendix A7.8, we conduct sensitivity analyses to assess how large the violations of the exclusion restriction must be to invalidate the above results.[77] We find that, for example, to overturn the positive 2SLS effect of repression on KIA/WIA, moving one's birth location 38 kilometers closer to a railway station (i.e. from the median distance to zero) would need to increase the risk of K/WIA by at least 3% through a channel other than repression. The reported estimates appear to be robust to non-trivial violations of the exclusion restriction.

[77] Conley, Hansen and Rossi, 2012.

Two additional explanations of the differences between OLS and two-stage estimates are possible. OLS may suffer from attenuation bias due to errors in the measurement of repression. Since instruments should alleviate this attenuation bias, the larger magnitude of 2SLS and FRDD point estimates makes sense. Furthermore, the two-stage estimates represent the local effect of repression induced by railway access and proximity to regional borders. It is possible that repression induced by exogenous factors appeared more arbitrary, and as such, induced a stronger effect on combat behavior.

We conducted a battery of additional robustness tests. Instead of analyzing individual outcomes, we averaged them at the level of birthplace and then ran the same regressions on these aggregated outcomes (see Appendix A7.1). We did the same at the level of districts ($N = 333$), directly controlling for local population size and urbanization from the 1926 census (see Appendix A7.4). Our conclusions remain virtually identical. We also considered the possibility that our results are biased due to incompletely observed records. We observe discharge records for 46% of soldiers, and our analyses assumed that soldiers without such records continued their service until the end of the war. In Appendix A7.2, we replicate our earlier analyses while excluding individuals whose discharge reasons are not observed. Again, our conclusions remain robust. Our results also hold when we consider a more selective subset of medals (Appendix A7.3).

Finally, a well known problem with clustered treatment designs is bias due to unequal cluster size. In our case, because higher-population areas may, mechanically, see higher absolute numbers of arrests, the treatment level is correlated with cluster size. To evaluate these biases, we adopt a matched cluster sample design,[78] sampling pairs of birthplaces that are similar on observable pre-treatment covariates, are from the same grid cell and in the same quintile of cluster size. The procedure yields a matched sample of 43,118 clusters (23.5% of total), or 21,559 matched pairs. We ran our analyses on the matched sample, and found substantively consistent results (Appendix A7.1).

[78] Imai et al., 2009.

7. Interpretation

We now consider several potential interpretations of our findings, including conformity, as well as discrimination, selective assignment, battlefield conditions, and peer effects.

7.1. Conformity

Empirical patterns align closely with the logic of conformity in Proposition 1. Consistent with the comparative static that expected battlefield resolve $\mathbb{E}(a^*(\omega, r))$ is increasing in repression r, we find that soldiers exposed to higher levels of prewar repression were more likely to fight until death or injury. These soldiers were also less likely to shirk their duties to the point of fleeing the battlefield, consistent with the deterrent effect, where $\Pr(a^*(\omega, r) < \underline{a})$ is decreasing in r. Finally, the conformity logic holds that, due to the alienation effect — $\Pr(a^*(\omega, r) > \overline{a})$ is decreasing in r — repression should reduce incentives to exceed one's orders. Evidence that soldiers from places with more repression were less likely to receive valor decorations is consistent with this prediction.

The observation that prewar terror conditioned soldiers to signal greater resolve resonates with qualitative evidence from historical accounts. Merridale (2006, 46-47) tells the story of a soldier, Ilya, whose estranged father the NKVD arrested in 1937. After the arrest, Ilya changed his educational and professional plans "for fear of unwelcome inquiries." Yet in 1941, he volunteered to serve at the front, ultimately sustaining life-threatening injuries at Stalingrad. Those who survived or witnessed the terror were "bound together by shared awe, shared faith and shared dread... It was far easier, as even the doubters found, to join the collective and share the dream than to remain alone, condemned to isolation and the fear of death."[79] The same incentives that increased compliance also encouraged overly cautious decision-making. As Overy (1998, 32) writes, "the result [of Stalin's terror] was the triumph of military illiteracy over military science, of political conformity over military initiative."

[79] Merridale, 2006, 45-46.

While the prediction that repression induces conformity finds support in the data, there are alternative interpretations for some of these results, which we must evaluate.

7.2. Discrimination

We interpreted the negative relationship between repression and medals as indicating a lack of initiative. Another possibility, however, is that soldiers from heavily-repressed areas faced systematic discrimination. Unit commanders may have been more hesitant to recommend — and higher authorities less likely to approve — decorations for soldiers from "problematic" spots of the country. Archival evidence offers no indication that commanding officers were aware of aggregate levels of repression at subordinates' birth locations, which is necessary for the discrimination argument to explain our results.[80] Moreover, Table 3 shows that soldiers from high-repression locations were *less* likely to face punishment for real or presumed violations of the military code, which is the opposite of what we would expect to see if these soldiers were subject to higher scrutiny.

To assess more systematically this alternative explanation, we check whether a similarly negative relationship exists between repression and promotions. Rank advancement decisions followed a structurally similar bureaucratic process to medals, but were more weakly tied to individual performance in combat. Similar to medals, unit commanders were responsible for recommending individuals for promotion, with conferral authority residing with higher ministerial or party authorities (see Appendix A3 for details). Unlike medals — where specific combat actions were the main consideration — criteria for promotion were more varied, and included factors like length of service, the need to quickly fill higher-ranking billets, constraints due to ethnic or religious quotas, soldiers' disciplinary record, party membership and other indicators of political loyalty. As such, there were many more opportunities for discrimination to enter the promotion process than in the simpler conferral of merit-based awards.

[80]Officers did have information about the family and "political background" of individual soldiers, but this is clearly not the same as information about the geographic distribution of aggregate repression levels.

If discrimination drove our results on medals, we should see a similar negative effect for promotions. As Table 5 reports, this is not the case. Under our baseline specifications with wartime promotion (i.e. at least one rank advancement) as the outcome, none of the three estimates are negative in sign, and the two estimates that reach statistical significance are positive. Unless rank advancement was insulated from political pressure while the conferral of decorations was not (which seems implausible), discrimination cannot explain these results. It is more likely that the army's promotion system favored conformity over initiative, as past historical studies suggest.[81]

	OLS	2SLS	FRDD
Coefficient	0.04 (0.02)*	0.3 (0.1)**	0.04 (0.1)
Mean Y	7.2	7.2	7.3
First Stage $\mathcal{F}$		141.8	25.7
Gridcells	11,056	5,558	1,558
Birthplaces	161,817	132,134	39,528
Soldiers	7,226,470	6,161,836	1,467,297

Outcome = receiving at least one promotion to a higher rank over the course of the war, measured on percentage scale (0 to 100). Standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record clustering probability. 2SLS analyses exclude birth locations > 100km from railroad. FRDD analyses exclude locations in non-matched regions and > 50km from regional borders. This analysis further excludes soldiers for whom rank information is unavailable (18% of cases). Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 5: Repression and Advancement to Higher Rank

7.3. *Selective Assignment*

Did soldiers from repressed places die in larger numbers because they were selectively assigned to more dangerous parts of the front? We assess this possibility by checking whether exposure to prewar repression increases the likelihood of two outcomes that are strongly predictive of higher casualty rates. First is assignment to the infantry branch of

[81] Glantz and House, 2015, 10.

the army, where direct exposure to enemy fire was higher than in other branches, like artillery and aviation. Second is assignment to so-called "penal units," which were routinely ordered to charge through minefields and machine-gun fire.

	Infantry	Penal
Model	OLS	
Coefficient	-0.02 (0.04)	0.01 (0.002)**
Mean Y	86.1	0.1
Gridcells	10,466	10,466
Birthplaces	150,707	150,707
Soldiers	5,724,988	5,724,988
Model	2SLS (First-stage $\mathcal{F} = 144.2$)	
Coefficient	-0.3 (0.2)	0.03 (0.01)**
Mean Y	85.8	0.1
Gridcells	5,538	5,538
Birthplaces	123,449	123,449
Soldiers	4,788,716	4,788,716
Model	FRDD (First-stage $\mathcal{F} = 26.1$)	
Coefficient	0.1 (0.1)	0.01 (0.01)
Mean Y	85.2	0.1
Gridcells	1,551	1,551
Birthplaces	36,764	36,764
Soldiers	1,083,299	1,083,299

Coefficients represent estimated effect of repression on soldiers' chances of assignment to an infantry unit or penal unit on percentage scale (0 to 100). Robust standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell fixed effects, individual and birth location-level covariates. Because data on branch and unit assignments are available only for 50.3% of draftees, sample size is considerably smaller for these analyses. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 6: Repression and Military Unit Assignment

The evidence here is mixed. Table 6 reports the estimated effects of repression on assignment to these units. Conscripts from high-repression areas were no more likely to serve in the infantry branch, but there is some evidence that prewar repression correlates with assignment to penal units. A soldier from a location with 32 arrests (first quartile) was $3.5 \times 0.01 = 0.04$ to $3.5 \times 0.03 = 0.1$ percentage points more likely to serve in a penal

unit than a soldier from a location with no arrests. Given that just .08% of soldiers were assigned to penal units over the duration of the war, this accounts for only a tiny fraction of the estimated effect of repression on battlefield deaths and injuries.

7.4. Battlefield Conditions

A related concern is that battlefield conditions, rather than prewar experiences, are driving our results. Some soldiers may have faced different incentives by virtue of where and when they saw combat. Soldiers' incentives may have co-evolved with wartime improvements to the Red Army's strategy, tactics, force structure, organization, leadership, training, weaponry and logistical support. They may have also shifted with Soviet efforts to expose German cruelty and to improve morale through nationalist appeals.

To address this possibility, we expand out baseline models to include fixed effects for the unit in which each soldier served, and month of deployment. In the case of OLS:

$$\begin{aligned} y_i = \gamma \cdot \text{Repression}_{j[i]} + \boldsymbol{\beta}' \boldsymbol{X}_{ij} + s\big(\text{lon}_{j[i]}, \text{lat}_{j[i]}\big) + \text{Cell}_{k[i]} \\ + \text{Military unit}_{u[i]} + \text{Month of war}_{t[i]} + \epsilon_i. \end{aligned} \tag{8}$$

where u indexes military units and t indexes months from June 1941 to May 1945. We add the same set of fixed effects to our 2SLS and FRDD specifications. Because 26% of soldiers served in more than one unit during the war, we disaggregated soldiers' records by unit assignment for this analysis. The variable *Military unit* has almost 12,000 unique values, and identifies the smallest-echelon unit mentioned in each record.[82] Note that both unit assignment and month of service are "post-treatment" measures. If including these post-treatment measures washes away the effects of repression, then it would suggest that unit assignment and the evolution of the war drove these effects.

The results of these analyses are in Table 7. Coefficient estimates remain similar to

[82] Among records with non-missing unit information, we traced 53% to a specific division, 10% to a brigade, 28% to a regiment, 2.4% to a battalion and 7.4% to a company.

	KIA/WIA	Flee	Medal
Model		OLS	
Coefficient	0.3 (0.1)**	-0.2 (0.04)**	-0.2 (0.1)**
Mean Y	59.3	22.7	18.8
Gridcells	9,647	9,647	9,647
Birthplaces	128,888	128,888	128,888
Soldiers	4,431,766	4,431,766	4,431,766
Model		2SLS (First-stage $\mathcal{F} = 136.9$)	
Coefficient	2.8 (0.4)**	-1.2 (0.2)**	-2.7 (0.4)**
Mean Y	58	23.3	19.3
Gridcells	5,444	5,444	5,444
Birthplaces	106,016	106,016	106,016
Soldiers	3,716,848	3,716,848	3,716,848
Model		FRDD (First-stage $\mathcal{F} = 33.7$)	
Coefficient	1 (0.2)**	-0.2 (0.1)$^{'}$	-0.9 (0.2)**
Mean Y	55.8	24	20.5
Gridcells	1,525	1,525	1,525
Birthplaces	31,326	31,326	31,326
Soldiers	840,783	840,783	840,783

Outcomes on percentage scale (0 to 100). Standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell, unit and month fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability. Sample includes disaggregated personnel records, with non-missing unit assignment and date information. 2SLS analyses exclude birth locations > 100km from railroad. RDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 7: ESTIMATES ADJUSTING FOR MILITARY UNIT AND MONTH.

baseline specifications, although they slightly differ in magnitude. Depending on the estimator, increasing repression from zero to the first quartile (32 arrests) increased one's chances of death or injury by 1.1 to 9.8 percentage points, decreased flight by 0.7 to 4.2 points, and chances of receiving a medal by 0.7 to 9.5 points (as before, percentage change $\approx 3.5\hat{\gamma}$). Exposure to prewar repression continues to drive battlefield behavior, even when we compare soldiers serving concurrently within the same units, who thus fought in the same battles, under the same commanders, with the same comrades-in-arms.

7.5. *Peer Effects*

Soldiers do not make decisions independently, and their choices may reflect not only their own prewar experiences, but also the backgrounds and actions of others in their unit. If a majority of one's comrades flees the battlefield, a soldier may have few opportunities to deviate from this pattern. The scope of individual agency may be similarly limited if almost no one flees. Heroic acts by others may inspire one to take similar action, and so on. Due to this interdependence, repression may affect combat motivation not only through individual exposure, but also, indirectly, through its effect on one's peers.

To analyze peer effects, we follow the econometric approach in Carrell, Sacerdote and West (2013) and augment the regression equation in (8) with two additional terms:

$$\begin{aligned} y_{it} =& \gamma \cdot \text{Repression}_{j[i]} + \rho \cdot \overline{\text{y}}_{ut[-i]} + \zeta \cdot \overline{\text{Repression}}_{ut[-i]} + \boldsymbol{\beta}' \boldsymbol{X}_{ij} + s\big(\text{lon}_{j[i]}, \text{lat}_{j[i]}\big) \\ &+ \text{Cell}_{k[i]} + \text{Military unit}_{ut[i]} + \text{Month of war}_t + \epsilon_{it} \end{aligned} \tag{9}$$

where $\overline{y}_{ut[-i]}$ and $\overline{Repression}_{ut[-i]}$ are, respectively, the average outcome and the average level of repression for the peers of soldier i in unit u during month t (calculated excluding soldier i). In the terminology of Manski (1993), ρ is the endogeneous peer effect and ζ is the exogenous peer effect. Since state repression was a taboo topic and long-term bonds between soldiers could not crystallize due to high turnover within units, it is unlikely that soldiers could form correct expectations about the repression experienced by their peers. Given this lack of an empirically plausible mechanism, we assume that exogenous peer effects did not play an important role in combat motivation ($\zeta = 0$). Provided that $\rho \neq 1$ and $\gamma \neq 0$, we can solve for the reduced form equation

$$\begin{aligned} y_{it} =& \gamma \cdot \text{Repression}_{j[i]} + \psi \cdot \overline{\text{Repression}}_{ut[-i]} + \boldsymbol{\beta}' \boldsymbol{X}_{ij} + s\big(\text{lon}_{j[i]}, \text{lat}_{j[i]}\big) \\ &+ \text{Cell}_{k[i]} + \text{Military unit}_{ut[i]} + \text{Month of war}_t + \epsilon_{it}, \end{aligned} \tag{10}$$

where $\psi = \gamma\rho/(1-\rho)$ is the reduced form peer effect.

We estimate the above equation using OLS. The estimates are valid only if the assignment of soldiers to units with low versus high average levels of repression is exogenous. This assumption is plausible given the pressures of mass mobilization in the USSR. Soviet mobilization plans left little room for accommodating the individual preferences of 30 million military-age males (i.e. no self-selection) or organizing unit composition on a dimension as obscure as exposure to repression. Unit assignment had some systematic components — reservists (i.e. older individuals with prior training) were sent to the front more quickly than untrained conscripts, military commissariats responsible for implementing the draft were organized by regional military district (most covering tens of thousands of square kilometers), and specialized units existed for soldiers with both exceptional skills (e.g. special forces) and disciplinary problems (e.g. penal units). However, these specialized units represented a tiny share of the army, and we can address the correlation of individual abilities through unit fixed effects. We can similarly account for geographic sorting with fixed effects for the grid cell of a soldier's birth. Monthly fixed effects further account for common shocks due to seasonal variation and the changing dynamics of the war. In cases where the unit assignment was based on conscripts' observable characteristics (e.g. age, ethnicity, class), controlling for these variables should eliminate the potential upward bias in estimated group coefficients.

Table 8 reports the estimated reduced form parameters as well as the endogenous peer effects recovered from these estimates ($\hat{\rho} = \hat{\psi}/(\hat{\psi} + \hat{\gamma})$). The estimated coefficient on individual exposure to repression ($\hat{\gamma}$) remains consistent with our baseline estimates: after controlling for the repression of a soldier's peers from the same unit (and other covariates), a one-quartile increase in repression (from 0 to 32 arrests) increased one's chances of death or injury by 0.7 percentage points, decreased flight by 0.7 points, and reduced the probability of a medal by 0.4 points. The individual level estimates in our baseline models do not seem to be confounded by peer effects.

	KIA/WIA	Flee	Medal
Direct individual effect ($\hat{\gamma}$)	0.2 (0.03)**	-0.2 (0.03)**	-0.1 (0.03)**
Reduced form peer effect ($\hat{\psi}$)	0.3 (0.02)**	-0.03 (0.02)**	-0.3 (0.02)**
Endogenous effect ($\hat{\rho}$)	0.6 (0.04)**	0.1 (0.1)*	0.8 (0.1)**
Mean Y	63.3	24.2	14.6
Gridcells	9,658	9,658	9,658
Birthplaces	128,354	128,354	128,354
Soldiers	4,369,898	4,369,898	4,369,898

Outcomes on percentage scale (0 to 100). Bootstrapped standard errors in parentheses. All models include grid cell, unit and month fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability. Sample includes disaggregated personnel records, with non-missing unit assignment and date information. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 8: ESTIMATES ADJUSTING FOR PEER EFFECTS.

For all three outcomes, the endogenous peer effect estimate ($\hat{\rho}$) is positive and significant at the 95 percent confidence level, confirming that soldiers' fortunes were positively correlated with those of others in their unit. If one's unit took exceptionally high losses in a given month, an individual's own chances of death or injury were considerably higher. A similar pattern held for the probabilities of fleeing or receiving a medal. Soldiers' behavior — for better or worse — varied with the behavior of their comrades-in-arms.

8. CONCLUSION

Our analysis of Red Army personnel records suggests that soldiers with greater exposure to Stalin's terror were more likely to fight to death or injury than to flee the battlefields of the Second World War. They were also less likely to show personal initiative in battle, as far as that can be inferred from military decorations.

We adopted multiple estimation strategies to ascertain whether these patterns reflect a spurious correlation or a genuine, potentially causal relationship. We exploited local randomness in the selection of repression victims, exogenous variation due to logistical costs, and geographic discontinuities due to local administrative discretion. All three de-

signs indicate that soldiers from places with more prewar repression were more likely to fight until death and less likely to flee, but they also received fewer decorations for valor. We conducted a battery of robustness tests to consider inferential threats from clustered treatment assignment, measurement errors and the validity of the railway instrument. We also considered several alternative substantive explanations for our results, including discrimination, selective unit assignment, battlefield conditions and peer effects.

While we cannot exclude the possibility that other, unobserved factors are driving these statistical relationships, our analyses overwhelmingly suggest that the net effect of prewar repression was conformity. Soldiers from places with higher levels of repression obeyed orders and kept fighting not because repression turned them into zealous patriots willing to go beyond their call of duty, but because they were more cognizant of what the state might do if they did not comply. While past repression may have compelled less-motivated soldiers to more forcefully signal their resolve, our evidence also reveals that repression may have decreased effort by highly-motivated types. Together, the countervailing forces of deterrence and alienation may have helped resolve some principal-agent problems associated with fighting, but they did so more by inculcating "mindless obedience,"[83] than by incentivizing innovation or skill.

References

Acemoglu, Daron, Tarek A Hassan and James A Robinson. 2011. "Social Structure and Development: A Legacy of the Holocaust in Russia." *The Quarterly Journal of Economics* 126:895–946.

Afonina, G.M. 1996. *Kratskie svedeniya o razvitii otechestvennykh zheleznykh dorog s 1838 po 1990 g. [Summary of domestic railroad development from 1838 to 1990]*. Moscow: Ministry of Railways of the Russian Federation.

[83] Murray and Millett, 2000, 26.

Barber IV, Benjamin and Charles Miller. 2019. "Propaganda and Combat Motivation: Radio Broadcasts and German Soldiers' Performance in World War II." *World Politics* 71(3):457–502.

Bartov, Omer. 1992. *Hitler's Army: Soldiers, Nazis, and War in the Third Reich*. New York: Oxford University Press.

Beissinger, Mark R. 2002. *Nationalist mobilization and the collapse of the Soviet State*. Cambridge University Press.

Benabou, Roland and Jean Tirole. 2003. "Intrinsic and extrinsic motivation." *The review of economic studies* 70(3):489–520.

Berman, Eli, Jacob N. Shapiro and Joseph H. Felter. 2011. "Can hearts and minds be bought? The economics of counterinsurgency in Iraq." *Journal of Political Economy* 119(4):766–819.

Bernheim, B Douglas. 1994. "A theory of conformity." *Journal of political Economy* 102(5):841–877.

Blaydes, Lisa. 2018. *State of Repression: Iraq under Saddam Hussein*. Princeton University Press.

Brehm, John O and Scott Gates. 1999. *Working, shirking, and sabotage: Bureaucratic response to a democratic public*. University of Michigan Press.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2005. *The Logic of Political Survival*. Cambridge, Mass: MIT Press.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2015. "Optimal data-driven regression discontinuity plots." *Journal of the American Statistical Association* 14(4).

Carrell, Scott E., Bruce I. Sacerdote and James E. West. 2013. "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation." *Econometrica* 81(3):855–882.

Central Executive Committee of USSR. 1935. *SSSR. Administrativno-territorial'noe delenie soyuznykh respublik na 1 sentyabrya 1946 goda [USSR. Administrative-territorial division of union republics as of 1 September 1935].* Moscow: Izd-vo 'Vlast' Sovetov' pri presidiume VTsIK.

Central Statistical Directorate of USSR. 1928-1929. *Vsesoyuznaya perepis' naseleniya 1926 goda*. Moscow: Central Statistical Directorate of USSR.

Chen, Daniel L. 2017. "The deterrent effect of the death penalty? Evidence from British commutations during World War I." *NBER Working Paper* .

Chukhin, Ivan. 1999. *Kareliia-37: Ideologiia i praktika terrora [Karelia-37: Ideology and Practice of Terror].* Petrozavodsk: Izd-vo Petrozavodskogo gos. universiteta.

Clausewitz, Carl von. 1832/1984. *On War*. Princeton, NJ: Princeton University Press.

Conley, Timothy G, Christian B Hansen and Peter E Rossi. 2012. "Plausibly exogenous." *Review of Economics and Statistics* 94(1):260–272.

Conquest, Robert. 2008. *The Great Terror: A Reassessment*. Oxford University Press.

Costa, Dora L and Matthew E Kahn. 2003. "Cowards and heroes: Group loyalty in the American Civil War." *The Quarterly Journal of Economics* 118(2):519–548.

Davie, HGW. 2017. "The Influence of Railways on Military Operations in the Russo-German War 1941–1945." *The Journal of Slavic Military Studies* 30(2):321–346.

Dell, Melissa. 2010. "The persistent effects of Peru's mining mita." *Econometrica* 78(6):1863–1903.

Dell, Melissa and Pablo Querubin. 2018. "Nation building through foreign intervention: Evidence from discontinuities in military strategies." *The Quarterly Journal of Economics* 133(2):701–764.

Edele, Mark. 2006. "Soviet Veterans as an Entitlement Group, 1945-1955." *Slavic Review* pp. 111–137.

Edele, Mark. 2017. *Stalin's Defectors: How Red Army Soldiers Became Hitler's Collaborators, 1941-1945*. Oxford University Press.

Enamorado, Ted, Benjamin Fifield and Kosuke Imai. 2019. "Using a probabilistic model to assist merging of large-scale administrative records." *American Political Science Review* 113(2):353–371.

Fearon, James. 1995. "Rationalist explanations for war." *International Organization* 49(3):379–414.

GARF. 1954. "State Archive of Russian Federation, collection 9401, series 2, case 450, pp. 30-37.".
URL: *http://www.alexanderyakovlev.org/fond/issues-doc/1009140*

Getty, J Arch and Oleg V Naumov. 1999. *The Road to Terror: Stalin and the Self-destruction of the Bolsheviks, 1932-1939*. Yale University Press.

Glantz, David M. 1998. *Stumbling Colossus: The Red Army on the Eve of World War*. University Press of Kansas Lawrence, KS.

Glantz, David M and Jonathan M House. 2015. *When titans clashed: how the Red Army stopped Hitler*. University Press of Kansas.

Gorkin, A.F., O.Yu. Schmidt, V.E. Motylev, M.V. Nikitin and B.M. Shaposhnikov, eds. 1937. *Bol'shoy Sovetskiy Atlas Mira. Tom 1. [Large Soviet Atlas of the World]*. Vol. 1 Moscow: Kartografiya.

Gregory, Paul R. 2009. *Terror by Quota: State Security from Lenin to Stalin*. Yale University Press.

Grosjean, Pauline. 2014. "Conflict and social and political preferences: Evidence from World War II and civil conflict in 35 European countries." *Comparative Economic Studies* 56(3):424–451.

Grossman, H. 1991. "A General Equilibrium Model of Insurrections." *American Economic Review* (81).

Hall, Andrew B, Connor Huff and Shiro Kuriwaki. 2019. "Wealth, slaveownership, and fighting for the confederacy: An empirical study of the American civil war." *American Political Science Review* 113(3):658–673.

Heydemann, Steven. 2013. "Tracking the" Arab Spring": Syria and the Future of Authoritarianism." *Journal of Democracy* 24(4):59–73.

Hirshleifer, Jack. 1989. "Conflict and rent-seeking success functions: Ratio vs. difference models of relative success." *Public choice* 63(2):101–112.

Hopper, Bruce C. 1930. "Soviet Transport Plans: Winning the East." *Foreign Affairs* 8(4):652–657.

Huff, Connor and Robert Schub. 2021. "Segregation, Integration, and Death: Evidence from the Korean War." *International Organization* 75(3):858–879.

Imai, Kosuke, Gary King, Clayton Nall et al. 2009. "The essential role of pair matching in cluster-randomized experiments, with application to the Mexican universal health insurance evaluation." *Statistical Science* 24(1):29–53.

Jansen, Marc and Nikita Vasilevich Petrov. 2002. *Stalin's loyal executioner: people's commissar Nikolai Ezhov, 1895 - 1940*. Hoover Institution Press.

Kachuk, Nikolay. 2013. "Bez vesti ne propavshyi [Not missing in action]." *SB. Belarus' segodnia* (59).

Kokurin, Aleksandr and N.V. Petrov. 2000. *GULAG: Glavnoe upravlenie lagerey, 1917-1960 [GULAG: Main directorate of camps, 1917-1960]*. Mezhdunarodnyy fond Demokratiya.

Kreps, David M. 1997. "Intrinsic motivation and extrinsic incentives." *The American economic review* 87(2):359–364.

Krivosheev, Grigori F. 1997. *Soviet casualties and combat losses in the twentieth century*. Greenhill Books/Lionel Leventhal.

Lehmann, Todd C. and Yuri M. Zhukov. 2019. "Until the Bitter End? The Diffusion of Surrender Across Battles." *International Organization* (forthcoming):1–37.

Lyall, Jason. 2017. Forced to Fight: Coercion, Blocking Detachments, and Trade-offs in Military Effectiveness. In *The Sword's Other Edge: Trade-offs in the Pursuit of Military Effectiveness*, ed. Dan Reiter. Cambridge: Cambridge University Press pp. 88–125.

Lyall, Jason. 2020. *Divided Armies: Inequality and Battlefield Performance in Modern War*. Princeton University Press.

Manski, Charles F. 1993. "Identification of endogenous social effects: The reflection problem." *The review of economic studies* 60(3):531–542.

Memorial. 2014. "Zhertvy politicheskogo terrora v SSSR [Victims of political terror in the USSR].".
URL: *http://lists.memo.ru/*

Mennis, Jeremy. 2003. "Generating surface models of population using dasymetric mapping." *The Professional Geographer* 55(1):31–42.

Merridale, Catherine. 2006. *Ivan's War: Life and Death in the Red Army, 1939-1945*. Macmillan.

Murray, Williamson and Allan Reed Millett. 2000. *A war to be won: Fighting the Second World War*. Harvard University Press.

Nunn, Nathan and Leonard Wantchekon. 2011. "The slave trade and the origins of mistrust in Africa." *The American Economic Review* 101(7):3221–3252.

Overy, Richard. 1998. *Russia's War: A History of the Soviet Effort: 1941-1945*. Penguin.

Pollack, Kenneth M. 2004. *Arabs at war: military effectiveness, 1948-1991*. U of Nebraska Press.

Pollack, Kenneth M. 2018. *Armies of Sand: The Past, Present, and Future of Arab Military Effectiveness*. Oxford University Press.

Reese, Roger. 2011. *Why Stalin's Soldiers Fought. The Red Army's Military Effectiveness in World War II.* Lawrence: University Press of Kansas.

Reiter, Dan and Allan C Stam III. 1998. "Democracy and battlefield military effectiveness." *Journal of Conflict Resolution* 42(3):259–277.

Rozenas, Arturas and Yuri Zhukov. 2019. "Mass Repression and Political Loyalty: Evidence from Stalin's 'Terror by Hunger'." *American Political Science Review* 113(2):569–583.

Shils, Edward A. and Morris Janowitz. 1948. "Cohesion and disintegration in the Wehrmacht in World War II." *Public Opinion Quarterly* 12(2):280–315.

Shiryaev, E.A. 2011. "Reformirovanie administrativno-territorial'nogo ustroystva RSFSR v 1930-e gody [Administrative-territorial reforms of RSFSR in the 1930's]." *Gosudarstvennoe upravlenie. Elektronnyi vestnik* (26).

Skaperdas, Stergios. 1996. "Contest success functions." *Economic theory* 7(2):283–290.

Slantchev, Branislav L. 2003. "The Power to Hurt: Costly Conflict with Completely Informed States." *American Political Science Review* 97(1):123–135.

Solzhenitsyn, Aleksandr I. 1973. *The Gulag Archipelago 1918-1956*. NY: Harper & Row.

Statiev, Alex. 2010. "Penal units in the Red Army." *Europe-Asia Studies* 62(5):721–747.

Surinov, A.E. and G.K. Oksenoyt, eds. 2015. *Velikaya Otechestvennaya voyna. Yubileynyi statisticheskiy sbornik [Grate Patriotic war. Anniversary statistical digest].* Moscow: Rosstat.

Talmadge, Caitlin. 2015. *The Dictator's Army: Battlefield Effectiveness in Authoritarian Regimes*. Ithaca: Cornell University Press.

Thurston, Robert W. 2000. "Cauldrons of Loyalty and Betrayal: Soviet Soldiers' Behavior, 1941 and 1945." *The People's War. Responses to World War II in the Soviet Union* pp. 235–57.

Tilly, Charles. 1985. War making and state making as organized crime. In *Bringing the State Back In*, ed. P. Evans and T. Skocpol D. Rueschemeyer. Cambridge pp. 169–191.

Ulam, Adam B. 1973. *Stalin: the man and his era*. New York: Viking.

Zickel, Raymond E. 1989. Soviet Union: A Country Study. Federal Research Division.

Zolotarev, Vladimir Antonovich, ed. 1997. *Prikazy narodnogo komissara oborony SSSR. 22 iyunya 1941 g. - 1942 g. [Orders of the USSR state commissar of defense. 22 June 1941 - 1942.* Vol. 13 (2-2) of *Russkiy arkhiv: Velikaya Otechestvennaya* Moscow: Terra. ISBN 5-85255-7.

Appendices

For Online Publication

Contents

A1. Proof of Proposition 1

The expected risk is equal to

$$\begin{aligned}
\mathbb{E}_\omega(a^*(\omega,r)) &= \int_{-\infty}^{\infty} a^*(\omega,r)dF(\omega) \\
&= \int_{-\infty}^{\overline{a}+\alpha r} \frac{\omega+\delta r\overline{a}-\alpha r}{1+\delta r}dF(\omega) + \int_{\overline{a}+\alpha r}^{\infty} \omega-\alpha r dF(\omega) \\
&= \int_{-\infty}^{\overline{a}+\alpha r} \frac{\omega}{1+\delta r}dF(\omega) + \int_{\overline{a}-\alpha r}^{\infty} \omega dF(\omega) \\
&+ \int_{-\infty}^{\overline{a}+\alpha r} \frac{\delta r\overline{a}-\alpha r}{1+\delta r}dF(\omega) + \int_{\overline{a}+\alpha r}^{\infty} \alpha r dF(\omega) \\
&= \frac{\mathbb{E}(\omega)}{1+\delta r} + \frac{\delta r}{1+\delta r}\int_{\overline{a}+\alpha r}^{\infty} \omega dF(\omega) \\
&+ \frac{\delta r\overline{a}-\alpha r}{1+\delta r}F(\overline{a}+\alpha r) - \alpha r(1-F(\overline{a}+\alpha r)) \\
&= \frac{\mathbb{E}(\omega)}{1+\delta r} + \frac{\delta r}{1+\delta r}\left[\int_{\overline{a}+\alpha r}^{\infty} \omega dF(\omega) + F(\overline{a}+\alpha r)(\overline{a}+\alpha r)\right] - \alpha r.
\end{aligned}$$

Letting $z = \overline{a}+\alpha r$, we have

$$\begin{aligned}
\frac{\partial}{\partial r}\mathbb{E}(a^*(\omega,r)) &= -\frac{\delta\mathbb{E}(\omega)}{(1+\delta r)^2} + \frac{\delta}{(1+\delta r)^2}\left[\int_z^{\infty} \omega dF(\omega) + F(z)z\right] \\
&+ \frac{\delta r}{1+\delta r}\left[-f(z)z\alpha + \alpha F(z) + f(z)z\alpha\right] - \alpha \\
&= -\frac{\delta\mathbb{E}(\omega)}{(1+\delta r)^2} + \frac{\delta}{(1+\delta r)^2}\left[\int_z^{\infty} \omega dF(\omega) + F(z)z\right] + \frac{\delta r}{1+\delta r}\alpha F(z) - \alpha \\
&> -\frac{\delta\mathbb{E}(\omega)}{(1+\delta r)^2} + \frac{\delta z}{(1+\delta r)^2} + \frac{\delta r}{1+\delta r}\alpha F(z) - \alpha,
\end{aligned}$$

where the last inequality follows from the fact that $\int_z^{\infty} \omega dF(\omega) = \mathbb{E}(\omega|\omega > z)(1-F(z)) > z(1-(F(z))$. The above expression is positive if and only if

$$\overline{a} > g(r,\overline{a}) \equiv \frac{\alpha}{\delta}\left[(1-F(\overline{a}+\alpha r))(1+\delta r(1+\delta r))\right] + \mathbb{E}(w). \tag{1}$$

g is finite-valued, since $\lim_{r\to 0} g(r,\overline{a}) = \lim_{r\to\infty} g(r,\overline{a}) = 0$. Furthermore, g is decreasing in $\overline{a}$ (since F is increasing), and so for each r, there is an interior point $\hat{a}(r) = g(r,\hat{a}(r))$ such that $\overline{a} > g(r,\overline{a})$ if and only if $\overline{a} > \hat{a}(r)$. It follows that $\mathbb{E}_\omega(a^*(\omega,r))$ is everywhere increasing in r if and only if $\overline{a} > \max_r \hat{a}(r)$.

Since $\overline{a} > \underline{a}$, the probability that a soldier defects or surrenders is equal to

$$\begin{aligned}
\Pr(a^*(\omega, r) < \underline{a}) =& \Pr(a^*(\omega, r) < \underline{a}, w < \overline{a} + \alpha r) + \Pr(a^*(\omega, r) < \underline{a}, w > \overline{a} + \alpha r) \\
=& \Pr(\omega < \underline{a}(1+\delta r) - r(\delta\overline{a} - \alpha), w < \overline{a} + \alpha r) + \Pr(\omega < \underline{a} + \alpha r, w > \overline{a} + \alpha r) \\
=& \Pr(\omega < \underline{a}(1+\delta r) - r(\delta\overline{a} - \alpha)) \\
=& F\left(\underline{a} + r(\delta\underline{a} - \delta\overline{a} + \alpha)\right),
\end{aligned}$$

which is increasing for all r if and only if $\overline{a} > \underline{a} + \alpha/\delta$. Let

$$\tilde{a}(\delta, \alpha) \equiv \max\{\max_r \hat{a}(r), \underline{a} + \alpha/\delta\}. \tag{2}$$

It then follows that $\mathbb{E}_\omega(a^*(\omega, r))$ is increasing and $\Pr(a^*(\omega, r)$ is decreasing for all $r \geq 0$ if $\overline{a}(\delta, \alpha) > \tilde{a}$.

Finally, consider the probability that a soldier takes personal initiative above what the commander orders ($a^*(\omega, r) > \overline{a}$). Since $a^*(\omega, r) \leq \overline{a}$ for $\omega \leq \overline{a} + \alpha r$, we have

$$\begin{aligned}
\Pr(a^*(\omega, r) > \overline{a}) &= \Pr(a^*(\omega, r) > \overline{a}, w < \overline{a} + \alpha r) + \Pr(a^*(\omega, r) > \overline{a}, w > \overline{a} + \alpha r) \\
&= \Pr(a^*(\omega, r) > \overline{a}, w > \overline{a} + \alpha r) \\
&= \Pr(\omega - \alpha r > \overline{a}, w > \overline{a} + \alpha r) \\
&= \Pr(w > \overline{a} + \alpha r) = 1 - F(\overline{a} + \alpha r),
\end{aligned}$$

which is decreasing in r for $\alpha > 0$.

A2. Record classification

The Russian Ministry of Defense's *Pamyat' Naroda* database contains multiple records per soldier, but does not provide a unique ID (e.g. military card number) to automatically match all records to the appropriate individual. In the absence of this unique ID, each record r_i ($i = 1, ..., 106$ mln) must be assigned to a cluster in the set $\{c_1, ..., c_N\}$, where c_j is a soldier (cluster of records) and N stands for (unknown) number of soldiers for whom we have records. In our baseline analyses, we solve this unsupervised classification problem using a probabilistic record linkage approach. To evaluate the performance of this procedure, we later also apply an alternative, deterministic fuzzy matching approach.

A2.1. Probabilistic approach

Our baseline approach builds on the probabilistic record linkage method proposed by Fellegi and Sunter (1969) and further developed by Enamorado, Fifield and Imai (2019) and implemented in their *R* package `fastLink`. While we use the main engine of in the `fastLink` package, our record classification problem is a bit idiosyncratic and requires some extra steps, as we detail below.

A2.1.1 Blocking Since comparing each pair of 106 million records is computationally infeasible, we first partition the data into blocks of records that are maximally similar on some fields (e.g. surname, first name, patronymic). We then assign records to clusters only within each block, per standard procedure in record linkage with large datasets.

The `fastLink` package has a functionality to create blocks using k-means classification of alphabetically ordered text fields. However, we found that for our application, the package's blocking scheme returns highly imbalanced blocks with many containing only a single record and some having millions of records. To obtain more balanced blocks, we used the following hierarchical procedure:

1. Partition records by the first letter of the surname, creating a set of *initial blocks*.
2. Within each initial block, identify *frequent surnames*, which appear at least 500 times.
3. Calculate the alphabetic order distance between each pair of surnames within each block. Using a size-constrained k-means clustering algorithm (Higgins, Sävje and Sekhon, 2016), cluster surnames within each initial block using frequent surnames as primary data points, forcing each cluster to have at least 500 unique surnames.
4. Partition blocks with more than 25,000 records further, using size-constrained k-means clustering based on the first name.
5. Partition remaining blocks with more than 25,000 records again using the patronymic.

The blocking procedure is *hierarchical* because it partitions the records based on the first name only if the partition on the last name alone was too coarse, and so on. We found that the hierarchical approach combined with the use of frequent surnames as primary data points for k-means clustering was particularly effective in achieving more balanced blocks, because it avoided creating clusters around misspelled names or clusters around rare surname-first name combinations, both of which generate imbalanced clusters. The procedure yielded 12,997 blocks ranging from 1,014 records to 29,748 records per block.

A2.1.2 Computing linkage probabilities The next step is to compute the probabilities that any two records belong to the same soldier within each of the 12,997 blocks. In the dataset, there are nineteen fields that can potentially inform these linkage probabilities. However, we found that applying the `fastLink` procedure for all nineteen fields was computationally infeasible. Therefore, we adopted a stratified approach by splitting the nineteen fields into three strata and then calculating linkage probabilities for each stratum. The fields were stratified as follows:

1. (1) surnames, (2) first name, (3) patronymic, (4) date of birth;
2. (5) birth region, (6) birth region (*oblast*), (7) birth district (*rayon*), (8) birth town, (9) discharge year, (10), discharge month, (11) discharge day;
3. (12) enlistment year, (13) enlistment month, (14) enlistment day, (15) enlistment oblast, (16) enlistment committee, (17) outcome, (18) rank, (19) military unit.

Let π_{ij}^s denote the probability that records i and j are a match based on the fields in stratum s. To compute the degree of matching across all three strata, we need to aggregate the probabilities π_{ij}^1, π_{ij}^2, and π_{ij}^3 for each $i \neq j$ within a block. We impose a constraint that for any two pairs of records to be a match, it is necessary (but not sufficient) that they approximately match on the fields in the first stratum. Even if two records match exactly on the fields in the second and third strata, they cannot represent the same person if they do not have similar names and dates of birth.

We calculate *pairwise linkage weights* between records i and j across the three strata as

$$m_{ij} = \pi_{ij}^1(1 + \pi_{ij}^2 + \pi_{ij}^3).$$

Two records *can* (but don't have to) be a likely match even if the probabilities π_{ij}^2 and π_{ij}^3 are small (or zero). We found it important to allow for this possibility to reduce the false negative match rate, because the fields in the second and third strata have many missing values and the probabilistic linkage tends to assign vanishingly small match probabilities for fully or partially missing fields. On the other hand, records i and j cannot be a likely match if π_{ij}^1 is small because the fields in the first stratum have few missing values.

A2.1.3 Classification Having calculated the degree of matching m_{ij} for all pairs of records, we then assigned records into clusters. This classification problem is identical to that of finding a community structure in a non-binary directed network (Leicht and Newman, 2008), where each edge represents a degree of relationship between the nodes.

We solved this problem using Ward's hierarchical agglomerative clustering, which assigns nodes to the same cluster by minimizing the within-cluster variance of network edges (Murtagh and Legendre, 2014).

Similar to the problem we faced when creating blocks, a naive application of the clustering procedure results in highly imbalanced (and implausible) clusters, with some soldiers having hundreds or even thousands of records. To avoid this problem, we adopted a hierarchical approach: we start by assigning records into clusters using a low similarity threshold; and then we further partition only those clusters that have more than ten records using a higher similarity threshold. Experimentation with different parameters has shown that the results are affected very little by the chosen values of thresholds as long as they are not unreasonable (e.g., we could stop splitting clusters with fewer than 100 records, but this would be mean we are assuming that one soldier could have as many as 99 separate records in the dataset, which makes no sense).

Finally, for each cluster we calculated the *total linkage weight*, which measures how well all pairs of records assigned to a cluster link with each other. This weight is the geometric mean of pairwise linkage weights of all records assigned to a cluster k:

$$W_k = \left(\prod_{i<j} m_{ij} \right)^{1/n_k} = \left(\prod_{i<j} \pi_{ij}^1 (1 + \pi_{ij}^2 + \pi_{ij}^3) \right)^{1/n_k},$$

where n_k is the number of records assigned to cluster k. The theoretical range of the weight goes from from 0 (i.e. at least one pair of records within a cluster has zero degree of linkage) to 3 (i.e. all pairs in the cluster have pairwise linkage weights equal to one, $m_{ij} = 1$; matching probabilities are equal to one in all three strata of the matching fields). In our analyses with soldier-level data, we weight each observation by the total linkage weight W_k to give more weight to observations that are classified with greater certainty.[1]

A2.2. Record clustering via deterministic fuzzy matching

As a validation exercise, we also clustered the personnel records using deterministic fuzzy string matching. This procedure assigns records to clusters based on the string distances between a set of fields using preset thresholds. It entailed the following steps:

1. Select the same 19 record fields stratified into three groups, as outlined above.

2. Let d_{ij}^f denote string distance between records i and j on field f. After experiment-

[1] The entire probabilistic record classification procedure took about 60 hours on a high performance computing cluster with 32 cores.

ing with multiple distance measures, we settled on restricted Damerau-Levenshtein distance as it seems to produce the most face validity. Calculate the string distances d_{ij}^f for all the fields in the first strata (surname, first name, patronymic, year of birth).

3. Using the complete hierarchical clustering method, assign records *on field* f to the same cluster if the restricted Damerau-Levenshtein distance between each pair of strings within the cluster does not exceed two units for names and one unit for the year of birth. That is, we allow two field entries to belong to the same cluster if they are dissimilar on at most two characters for names and one character for the year of birth. We use different cutoffs because most birth years are assigned to the same cluster if we allow two mistakes in four-digit numbers, most of which start with 19.

4. Aggregate set of records i and j into the same cluster if they match within the bounds of an error on all four fields in the first stratum.

5. If a cluster contains more than two records, split each such cluster using the same procedure as above but now employing fields from the second stratum; if large clusters remain, split them again using fields from the third stratum.

In the above scheme, the distance between a string and a missing value (or a distance between two missing values) is assumed to be zero. This assumption is required since missing values cannot be modeled explicitly in this scheme, in contrast to the probabilistic approach. This assumption essentially means that whenever we do not observe evidence of two strings being different, we assume they are the same. For instance, a soldier may have a discharge record that lists his birth location and an award record, which does not list a birth location. We would fail to match these two records if we did not treat missing values as stated, which clearly would be in error.

A2.3. Evaluation I: marginal properties

We first evaluate the probabilistic clustering scheme by comparing the marginal properties of the soldier-level dataset generated by this scheme against the marginal properties of the dataset generated by the deterministic scheme. The two schemes differ on a number of dimensions, such as the distance metric, probabilistic vs deterministic assignment, and treatment of the missing values. If the marginal properties of the two datasets are reasonably similar, then the clustering scheme is robust with respect to the specific choices.

Table A2.1 shows that the two clustering schemes yield reasonably similar results. The deterministic scheme yields more clusters (soldiers) than the probabilistic scheme. Closer inspection shows that this is mostly because the deterministic procedure fails to match

	Probabilistic	Deterministic
Soldiers	11,680,930	12,441,132
K/WIA	22.23 %	21.61 %
MIA	17.47 %	17.04 %
POW	5.72 %	5.37 %
DDT	0.16 %	0.15 %
PUN	0.79 %	0.72 %
Glory medal	2.32 %	2.01 %
Promotion	9.35 %	8.2 %

Table A2.1: Marginal properties of datasets from probabilistic and deterministic clustering

many records with missing values. More important than the total number of clusters are the distributions of the key outcomes that we analyze. We see that the percentages of outcomes across the two datasets are very similar across all measures. This is a suggestive but nonetheless important indication that the clustering schemes worked "correctly" and are not greatly dependent on specific parametric choices.

A2.4. Evaluation II: comparison with ground truth

While typically, record linkage and clustering problems are unsupervised in the sense that we don't have the ground truth against which to compare the output of the algorithm, in this particular case, we have some *partial* access to the ground truth. About 11% of records (about 11.8M) contain a field named "ID card," which we believe denotes the identification number of a soldier's military card. The value of this ID is quite limited because it is only included in the award records and in some portion of enlistment records. This means we can only use it to cluster records with and between these types or records, but not others. But we can use this identifier to evaluate how well the probabilistic clustering scheme predicts these "ground truth" clusters for which we have data.

Within each block where records were clustered, we calculate the similarity between the ground truth clustering and the clustering generated by the probabilistic clustering scheme using three standard metrics: (1) true positive rate (TP), the proportion of records that belong to the same cluster that are also assigned to the same cluster by the algorithm; (2) true negative rate (TN), the proportion of records that belong to different clusters being assigned to different clusters; and (3) F_1 score defined as $\frac{TP}{TP + 1/2(FP + FN)}$.

Figure A2.1 shows the distributions of three measures. The rate of true positive predictions is high across all blocks, ranging from 0.85 to 1, with over 51% of blocks having a higher than 95% rate of true positives. The rate of true negative predictions is also

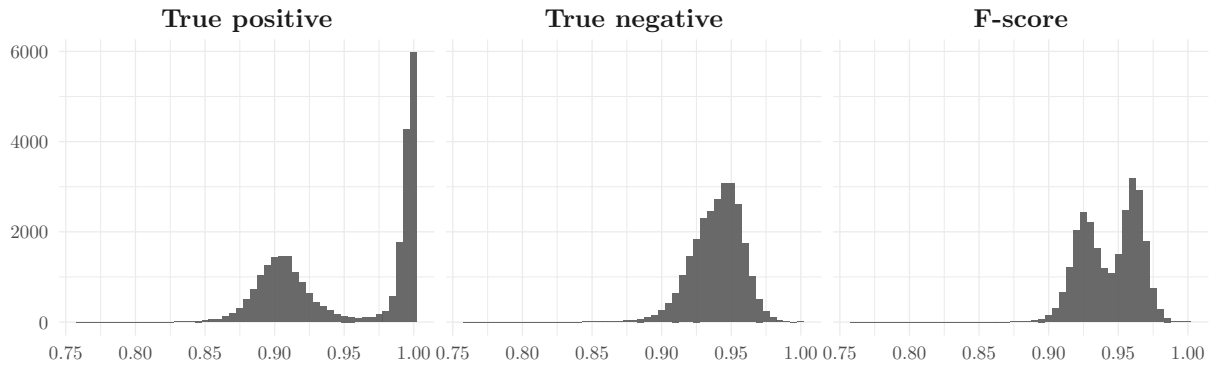


Figure A2.1: Performance of our clustering scheme against the ground truth.

high across all blocks, with an average of 94%. Finally, the F_1 score also indicates high predictive accuracy across most blocks.

A3. Valor decorations

A3.1. Categories of orders and medals

Soviet Army decorations and awards for WWII fell into multiple categories depending on their scope (individual, mass), target (civilian, military), merit (various classes of courage), timing (wartime, posthumous, commemorative), service and branch (aviation, infantry, armor, navy). Each category carried different parameters and standards for qualification. The USSR had a multi-tiered system of entities authorized to make award decisions, and this system itself was created based on award categories and their rank order.[2] The complex awarding system meant that any unique decoration might belong to one or multiple categories. As such, the qualification criteria and the decision-making authorities that oversaw the awarding process were unique to each award.

We focus on a particular set of decorations that were given specifically for individual initiative and valor. As a result, we exclude medals and orders awarded en masse to

[2]In general, unit commanders were responsible for the recommendation of individual assignment and promotion of enlisted men at times of war. Upon recommendation by unit commanders, different government agencies were responsible for the conferral of the award. The Main Administration of Personnel of the Commissariat of Defense had discretion over ranks up to lieutenant colonel. The Council of People's Commissars was responsible for rank advancement decisions between the ranks of lieutenant colonel and marshal (Bolin, 1946).

an entire unit (e.g. campaign medals)[3] or granted after 1945 as jubilee decorations.[4] We focus on decorations that were awarded only during WWII and for recognition of acts displayed on the battlefield. Filtering based on these criteria leads us to four military decoration categories: a) "For Courage," awarded to soldiers, borders and internal troops for personal courage and bravery displayed in defense of the Soviet Motherland and during the performance of military duties in circumstances involving a risk to life; b) "For Battle Merit," awarded for display of bravery during combat action resulting in a military success; c)"The Order of Glory," awarded to rank and file soldiers and non-commissioned officers of the Red Army for recognition of glorious feats of bravery, courage and fearlessness in combat for the Soviet Motherland; and d) "Hero of the Soviet Union" – the highest military distinction awarded for heroic feats in service to the Soviet Motherland.

A3.2. Medal "For Courage"

Established by a Decree of the Presidium of the Supreme Soviet on October 17, 1938, the Medal for Courage was intended for soldiers who provided active assistance to the success of military activities and for strengthening the combat readiness of troops. Soviet Army and Navy personnel, border and interior troops could receive the award. The description of the medal and the awarding regulations were amended by decrees of the Presidium of the Supreme Soviet of June 19, 1943 and on December 16, 1947.

"For Courage" was the second medal after "XX Years of the Red Army" to be established in the USSR. It was awarded mainly to rank and file soldiers and less often to junior officers. Senior officers and generals almost never received the Medal "For Courage". The first medals in this category were awarded two days after its establishment (62 soldiers). Approximately 26,000 servicemen received the medal before the start of the Great Patriotic War (we exclude these from our measure). Over 4,230,000 medals were awarded exclusively for feats performed during the war.

Awarding criteria

Criteria for recommending Medal "For Courage" included the following acts of bravery:

- For courage demonstrated in battles with the enemies of the Soviet Motherland;
- For courage demonstrated while protecting the state border of the Soviet Motherland;

[3]These decorations include medals awarded for the defense or capture of cities, such as "Medal for Defense of Leningrad", "Medal for Defense of the Caucasus", "Medal for Defense of Stalingrad", "Medal for the Capture of Berlin", "Medal for the Capture of Budapest", "Medal for the Victory over Japan", etc.

[4]There were significant changes to the awarding procedures and standards of all orders and medals in the postwar period that altered the definition of "merit" required for recognition.

- For courage demonstrated during performance of military duty in conditions associated with a risk to life.

A3.3. Medal "For Battle Merit"

The Medal "For Battle Merit" was established by a Decree of the Presidium of the Supreme Soviet on October 17, 1938 – on the same day as "For Courage". Subsequent changes to the description and awarding regulations took place on the same dates as "For Courage". Although, the Medal "For Battle Merit" was awarded to rank and file soldiers, civilians could also receive awards for wartime bravery. For instance, in summer 1941, a 15-year-old schoolboy Zhenya Nefedov received the Medal "For Battle Merit" in Moscow for his efforts against German incendiary bombs, with which Nazi bombers bombarded residential areas of Moscow. During one raid, the eighth-grader put out nine "lighters".

By the decree of June 4, 1944, the Presidium of the Supreme Soviet introduced a procedure for awarding orders and medals to servicemen of the Red Army for length of service. The only medal awarded to servicemen for 10 years of impeccable service was the Medal "For Battle Merit" (orders were awarded for 15, 20 and more years of service). This procedure of awarding "for length of service" was canceled only in 1958.

Our measure excludes the approximately 21,000 servicemen who received the medal before the start of the Great Patriotic War, and all those who received it after 1945.

Awarding criteria

Criteria for recommending the Medal for "For Battle Merit" included the following:

- For skillfull, proactive and courageous actions in battle that contributed to the successful fulfillment of combat missions by a military unit or subunit;
- For courage shown in defense of the state border of the Soviet Motherland;
- For excellent achievements in combat and training, mastering new military equipment and maintaining high combat readiness of military units and subunits during active military service.

A3.4. The Order of Glory

The Order of Glory was unique in that it could be awarded only for tactical-level combat valor and ranked among the most prestigious military decorations in Soviet history. It was reserved solely for enlisted personnel and non-commissioned officers (Empric, 2017).

Established by a decree of the Presidium of the Supreme Soviet on November 8, 1943, the Order of Glory comprised three distinct sequential classes, with the I class being the highest.[5] Soldiers received a right to be conferred a higher military rank and were referred to as a "Full Cavalier of the Order of Glory" if awarded all three classes of the order.[6] Only 2,656 Red Army soldiers received all three classes of the Order of Glory during and after WWII, and over 9% of those approved between 1944 and 1946 were posthumous recognitions (Empric, 2017). By 1945, approximately 1,500 Orders of Glory of I class, 17,000 II class, and 200,000 III class had been awarded.

According to official wartime military personnel records, Full Cavaliers of the Order of Glory included representatives of 41 distinct Soviet ethnicities, with Russian comprising the largest ethnic group (70%), followed by Ukrainians (17%) and Belorussians (2%). Half of all Full Cavaliers fought in one of two Red Army fronts: the 1st Belorussian Front, commanded by Marshal of the Soviet Union Georgiy Zhukov, and the 1st Ukrainian Front, commanded by Marshal of the Soviet Union Ivan Konev. More than 50% of Full Cavaliers came from infantry, followed by artillery (26.5%), combat engineers (11.45%), tank and mechanized forces (3.46%), aviation forces (2.03%), and miscellaneous and support troops (4.3%) (Empric, 2017). Ten of the Full Cavaliers of the Soviet Union earned their decorations while serving their respective sentences in penal units.

Unit commanders at the brigade level or higher had the right to award the Order of Glory of III class. Army (flotilla) commanders could award the II class. Only the Presidium of the Supreme Soviet of the USSR could award the I class. In these cases, the battalion or brigade commander would initiate the award recommendation, which would then have to be approved by the division, corps, army and front commanders, before being dispatched to Moscow for final vetting and approval.

Awarding criteria

Criteria for recommending Orders of Glory included the following acts of bravery:

- As the first to burst into the enemy's position, by personal bravery, contributed to the success of the common cause;
- While in a burning tank, continued to

[5]Until 1974, the Order of Glory remained the only order of the USSR, issued only for personal merits and never issued to entire military units, enterprises or organizations. The only exception to this rule occurred once in January 1945, when the entire contingent of a single unit was awarded the Order of Glory. In battles for the liberation of Poland, during a break-through of deep-echeloned German defenses on the left bank of the Vistula river, the soldiers of the 1st battalion of the Red Banner 215th Regiment of the Orders of the Red Banner, Lenin and Suvorov 77th Guards Chernigov Infantry Rifle Division captured three lines of enemy trenches in a swift assault and held their positions until the main forces arrived.

[6]The statute of the order provided for the rank promotion of those awarded all three classes, which was an exception to the Soviet decoration system.

carry out the combat mission;

- In a moment of danger, saved his unit's banner from enemy capture;
- With accurate fire from a personal weapon, destroyed from 10 to 50 enemy soldiers and officers;
- While in combat, using anti-tank rifle fire, knocked out at least two enemy tanks;
- Using hand grenades, destroyed from one to three tanks on the battlefield or in the enemy's rear area;
- Using artillery or machine gun fire, destroyed at least three enemy aircrafts;
- Defying clear danger, as the first to burst into an enemy bunker (trench or dugout), destroyed its garrison with decisive actions;
- As a result of personal reconnaissance, determined weak points in the enemy's defense and led forces into the enemy's rear area;
- Personally captured an enemy officer;
- At night, removed the guard post (patrol) of the enemy or captured him;
- With resourcefulness and courage, personally made his way to the enemy's position and destroyed his machine gun or mortar;
- While in night guard, destroyed the enemy's warehouse with military equipment;
- While risking his life, saved the commander in combat from imminent danger that threatened him;
- Defying personal danger, captured the enemy banner in combat;
- While wounded, returned to duty following immediate treatment;
- Using personal weaponry, shut down an enemy aircraft;
- By destroying enemy firepower with artillery or mortar fire, ensured the successful operation of his unit;
- Under enemy fire, made a passage into the enemy's wire fences for the advancing unit;
- Risking his life under enemy fire, assisted the wounded during a series of battles;
- Being in a destroyed tank, continued to carry out a combat mission from the tank's weapons;
- Rapidly crashing into the enemy column on his tank, crumpled it and continued to carry out the combat mission;
- Crushed one or several enemy weapons with his tank or destroyed at least two machine-gun nests;

- While in reconnaissance mission, obtained valuable information about the enemy;
- In an air battle, as a fighter pilot, destroyed from 2 to 4 enemy fighter aircrafts or from 3 to 6 bomber aircrafts;
- As a result of an assault raid, as an attack pilot, destroyed 2 to 5 enemy tanks or 3 to 6 steam locomotives, or detonated a train at a railway station, or destroyed at least two aircrafts at an enemy airfield;
- As a result of bold initiative, as an attack pilot in an air battle, destroyed 1 or 2 enemy aircrafts;
- As members of the crew of daylight bombers, destroyed trains, blew up bridges, the ammunition depot and fuel, destroyed the headquarters of enemy unit, destroyed the railway station, blew up the power station or dam, destroyed a military ship, transport, boat, or destroyed at least two enemy aircrafts;
- As a crew member on a light night bomber, blew up an ammunition depot or fuel dump; destroyed the enemy's headquarters; blew up a railroad train or bridge;
- As a crew member on a long-range night bomber, demolished a railroad station; blew up an ammunition depot or fuel dump; demolished a port facility; destroyed a sea transport or a railroad train; demolished or burned down an important factory or mill;
- As a crew member on a daylight bomber, as a result of courageous actions in aerial combat, show down 1 to 2 enemy aircrafts;
- As a crew member on a reconnaissance aircraft, for successfully accomplished reconnaissance, which resulted in valuable intelligence about the enemy.

Examples

Below are several examples of individuals who received Orders of Glory of each class.

Order of Glory III Class

From the award page of machine-gunner Egorov Dmitriy Nikolaevich (b. 1923), awarded the Order of Glory III Class on January 30, 1945:

"On January 13, 1945, while repelling counterattacks by numerically superior enemy infantry in the center of Budapest, Comrade Yegorov destroyed the enemy's machine gun point and 12 enemy soldiers with his personal machine gun. On January 14, 1945, while advancing to a bridge over the Danube River, Yegorov killed 6 enemy soldiers and took 2 Hungarian soldiers as prisoners."

Commander of the 200th Guards Rifle Regiment
Guard Major Panin

From the award page of Squad Commander Marchenko Anatoliy Andreevich (b. 1917), awarded the Order of Glory III Class on February 20, 1945:

"On 14 February, 1945, in an offensive battle against the German invaders in the area of of the city of Wanzen of the 1st Ukrainian Front, while performing a combat mission to destroy a group of machine gunners with his squad, and while entrenched in a cemetery, he showed himself to be a strong-willed, trained and staunch commander. He made his way through a break in the wall, chose a reliable shelter behind the stone, and with a long burst of machine gun destroyed the enemy's machine-gun crew of 3 people. The first to rise to the attack, he galvanized his squad and knocked out the entrenched machine gunners, in the meantime destroying 2 fascists with hand grenades. Clearing the houses of the city from the German machine gunners, he shot 3 Nazis and took the Hitler Banner of the military plant as war trophies. For the precise execution of a combat mission and decisive actions on the battlefield against the German invaders, comrade Marchenko is recognized with the Order of Glory III Class."

Commander of the 181st Infantry Regiment
Lieutenant Colonel Korkishko

From the award page of gunner Galyadinov Fayzirakhman Boltinovich (b. 1915), awarded the Order of Glory III Class on April 18, 1945:

"On April 18, 1945, during the hostilities in the Raygorod region, Comrade Galyadinov proved himself to be a courageous and staunch warrior. Comrade Galyautdinov's tank was destroyed on the battlefield. He ensured the exit of the entire tank crew covering them with his machine gun fire, and occupied a neighboring house with his crew to guard and defend the tank. During combat, Galyadinov destroyed a light machine gun and 2 soldiers of the enemy. Wounded in the chest, Comrade Galyadinov did not leave his place and remained in the cover of the tank until the infantry approached. His actions are recognized with the Order of Glory III Class."

Commander of the 78th Guards Heavy Tank Dnovsky Regiment
Guard Lieutenant Colonel Gerasimov

Order of Glory II Class

From the award page of cannon gunner, Guard Staff Sergeant Zolotikh Dmitriy Andreevich (b. 1924), awarded the Order of Glory II Class on August 27, 1944:

"On August 7, 1944, in a fierce battle during the liberation of Lesna station of Baranovichi region, using his 45-mm cannon in the infantry battle formations destroyed one German tank with a direct fire. The Germans, intensifying their onslaught and moving to a fierce counterattack with the support of 20 tanks, approached the firing position of his cannon at 100 meters. Wounded in the arm, he did not leave the battlefield and, not losing his composure in front of the enemy, opened a hurricane of fire on enemy tanks, and knocked out another tank, after which he destroyed up to 20 German soldiers and officers. After repeated orders from the fire platoon commander, he then left the battlefield. His actions are recognized with the Order of Glory II Class."

Commander of the 162th Guards Rifle Regiment
Guard Major Stepura

From the award page of foot reconnaissance platoon scout Shmonin Fyodor Vasilyevich (b. 1911), awarded the Order of Glory II Class on September 29, 1944:

"On August 21, 1944, in the battle for the village of Voinesti (Romania), Private Shmonin, showing fearlessness and courage, suddenly and carefully burst into a village and, having approached the house in which there were more than 30 German soldiers, he began to throw grenades at them and shoot the Germans running out of the house in a panic. In total, in this battle, Shmonin destroyed 12 German fascist invaders, and took 19 German soldiers as prisoners and brought them to the regiment headquarters. His actions are recognized with the Order of Glory II Class."

Commander of the 933th Rifle Regiment
Lieutenant Colonel Fimosin

From the award page of reconnaissance scout Dolgov Pyotr Nikolaevich (b. 1922), awarded the Order of Glory II Class on October 21, 1944:

"On October 2, 1944, during a night search for scouts in the area north-west of the city of Lomas, Comrade Dolgov was the first to cross the Narev River, and threw a cable rope to his comrades, thereby ensuring the crossing of the safe entire group. During the capture operation of the group, Comrade Dolgov silently crawled to the enemy trench and knocked down the German night guard. Having disarmed the enemy, Comrade Dolgov, with his comrades who arrived in time, delivered the prisoner to his destination. The mission was accomplished. His actions are recognized with the Order of Glory II Class."

Commander of the 444th Separate Reconnaissance Company
Senior lieutenant Pismorov

Order of Glory I Class

From the award page of SU-85 gunner, Sergeant Major Zaboev Vasiliy Andreevich (b. 1914), awarded the Order of Glory I Class on March 24, 1945:

"In battles near the village of Relsheersh, the vehicle commander was wounded during repeated attacks of the enemy. Comrade Zaboev assumed command and, in this battle, repelled 3 enemy attacks, destroyed 3 tanks, 2 guns, 2 mortarts, 1 machine-gun point, and up to 30 enemy soldiers and officers. In the same battle Comrade Zaboev was seriously wounded, but did not leave his combat post, and brought his car out in good working condition. His actions are recognized with the Order of Glory I Class."

Commander 1438th Self-propelled Artillery Red Banner Order of Suvorov Regiment
Colonel Zatylkin

From the award page of Soldier Semyonov Yegor Dmitrievich (b. 1906), awarded the Order of Glory I Class on May 31, 1945:

"On March 27, 1945, during the assault on height 60.6 for liquidation of the Alt-kyustrinskoensky bridgehead on the right bank of the Oder River, Comrade Semyonov showed examples of stamina and fearlessness in battle. At the signal for the start of the attack, Comrade Semyonov was the first to break into the enemy's location and knocked down five Nazis in hand-to-hand combat. When pursuing the retreating enemy, the first light machine gun went out of order. Comrade Semyonov quickly replaced it and, with his fire, destroyed the enemy light machine gun and 12 German soldiers, scattering the retreating Germans. Thus, he ensured the rapid advancement of the rifle company. His actions are recognized with the Order of Glory I Class."

Commander 487th Red Banner Infantry Regiment
Lieutenant Colonel Tarasov

A3.5. Hero of the Soviet Union

The Hero of the Soviet Union was the highest degree of distinction of the Soviet period and the most prestigious title in the Soviet hierarchy of awards.

Established by a Decree of the Presidium of the Supreme Soviet on April 16, 1934, title of Hero of the Soviet Union was given for personal or collective services to the Soviet state and society associated with the performance of a heroic deed. Along with this title, the awardee received a) the highest award of the USSR — the Order of Lenin; b) a badge of

special distinction — the Gold Star medal; and c) diploma of the Presidium of the USSR Supreme Soviet. The title also carried additional welfare privileges, such as medical, housing, entertainment benefits and a pension. The title of Hero of the Soviet Union was first conferred on April 20, 1934 to a number of Soviet aviators for rescuing the polar expedition and the crew of the Chelyuskin icebreaker.

On December 31, 1936, the title of Hero of the Soviet Union was for the first time awarded for military exploits. Eleven commanders of the Red Army — participants of the Spanish Civil War — became heroes. It is noteworthy that all of them were also pilots, and three of them were foreigners by origin: the Italian Primo Gibelli, the German Ernst Schacht and the Bulgarian Zakhari Zakhariev. Among the heroes was the lieutenant of the 61st fighter squadron, Chernykh S.A. In Spain, he was the first Soviet pilot to shoot down the latest Messerschmitt Bf 109B fighter. On June 22, 1941, he commanded the 9th Mixed Air Division. On the first day of the war, the division suffered huge losses (347 out of 409 aircraft of the division were destroyed). As a result, Chernykh was accused of criminal inaction and was executed on June 27, 1941.

In total, before the start of the Great Patriotic War, the title of Hero was awarded to 626 people (including 3 women), five of whom were twice heroes. 11,635 people (92% of the total number of heroes) were awarded the title during the Great Patriotic War. 101 were awarded twice and 3 were awarded thrice. In the first year of the war, only a few dozen people were awarded the title, all in the period from July to October 1941. By 1944, the number of Heroes of the Soviet Union increased by more than 3,000, mainly infantrymen. For the liberation of the Czechoslovakia, the title was awarded 88 times, for the liberation of Poland — 1667 times, for the Berlin operation — more than 600 times.

Among all the Heroes of the Soviet Union, 35% were enlisted, 61% were junior and field-grade officers and 3.3% (380 people) were generals, admirals and marshals. The youngest person to receive the title was 17-year-old partisan Lenya Golikov (posthumously). There were only two wartime cases when the title of Hero of the Soviet Union was awarded to all personnel in a unit, comprising 95 mostly posthumous decorations.

According to the ethnic composition, the majority of the Heroes were Russians — 7998 people, followed by 2,021 Ukrainians, 299 Belarusians, 161 Tatars, 107 Jews, 96 Kazakhs, 90 Georgians, 89 Armenians, 67 Uzbeks, 63 Mordvin, 45 Chuvashes, 43 Azerbaijanis, 38 Bashkirs, 31 Ossetians, 18 Mari, 16 Turkmen, 15 Lithuanians, 15 Tajiks, 12 Latvians, 12 Kyrgyz, 10 Komi, 10 Udmurts, 9 Estonians, 8 Karelians, 8 Kalmyks, 6 Kabardins, 6 Adygeis, 4 Abkhazians, 2 Yakuts, 2 Moldovans, and 1 Tuvinian.

Awarding criteria

The title could only be awarded by the Presidium of the Supreme Soviet of the USSR for exceptional heroic deeds. A Hero of the Soviet Union who performed a second heroic deed, no less than the one for which others who had performed a similar feat received the title of Hero of the Soviet Union, was awarded the Order of Lenin, a second Gold Star, and a commemorative bronze bust in his hometown. A Hero of the Soviet Union awarded two Gold Star medals could again receive the Order of Lenin and Gold Star for new heroic deeds similar to those previously committed.

Examples

From the award page of machine-gunner Bondarenko Pyotr Nikolaevich (b. 1921), awarded the title of the Hero of the Soviet Union on October 26, 1943:

"Guards gunner junior sergeant Bondarenko was among the first to cross with his weapon to the right bank of the Dnieper. On September 27, 1943, while fighting to repel enemy counterattacks, Bondarenko destroyed 4 firing points and up to 45 enemy soldiers with an open direct fire. In the battle on October 7, 1943, under heavy artillery fire and the attack of enemy aircraft, he fired at the enemy's counterattacking infantry, which was supported by 20 tanks. He was wounded by a sharpnel of a bomb, but despite the pain and severe bleeding, he continued to attack, setting fire to one T-4 tank and destroying up to 15 enemy soldiers. He was again wounded by sharpnel of another bomb, but despite being wounded, he continued to remain in the ranks. When repelling another counterattack, he was killed on the battlefield."

Commander of the 115th Krasnograd Guards Fighter Anti-Tank Artillery Regiment
Guard Lieutenant Colonel Kozyarenko

From the award page of junior lieutenant Marchenko Fyodr Illarionovich (b. 1919), awarded the title of the Hero of the Soviet Union on April 17, 1945:

"On April 14, 1945, in battles with the German invaders during the breakthrough of the heavily fortified enemy defenses on the West Bank of the Oder River, and during offensive operations, Comrade Marchenko, by his personal actions, inspired military deeds. In the battles for the village of Hardenberg on April 16, 1945, he showed exceptional courage and bravery. The Germans launched a counterattack. Comrade Marchenko personally led the unit, repelling the enemy's counterattack, with the slogan "Communists Forward For the Motherland", raising soldiers' spirits, and rushed to storm the enemy trenches. He was the first to break into the enemy trenches, where he destroyed 5 German soldiers and one officer from his personal weapons, taking 4 German soldiers as prisoners. Following his example, the soldiers knocked the enemy out of his trenches with

a swift blow and began to pursue. In the ensuing battle on April 17, 1945, Comrade Marchenko, showing courage and personal bravery, led the fighters forward. In the same battle, he was seriously wounded by a sharpnel as a result of the enemy shelling and died because of his wounds. For the courage and bravery shown, Comrade Marchenko deserves to be posthumously awarded the title of Hero of the Soviet Union."

Commander of the 180th Guards Rifle Regiment
Guard Major Kuzov

From the award page of senior sergeant Nemchikov Vladimir Ivanovich (b. 1925), awarded the title of the Hero of the Soviet Union on July 12, 1944:

"The regiment commander ordered to pick up 12 people from a group of brave soldiers to perform a particularly difficult and dangerous task. The first to voluntarily express a desire to perform any task was guard senior sergeant Nemchikov, stating that he was ready to complete any task for the sake of defeating the enemy and even sacrifice his life. Having received the order, this group, by swimming in special suits, was supposed to ferry 6 rafts with stuffed effigies to the enemy's shore in order to direct enemy fire on themselves, which was then detected and suppressed by our artillery. However, some rafts with effigies were destroyed by Finnish artillery while still on the shore and could not be lowered into the water. Comrade Nemchikov made an independent decision, threw himself into the water and swam to the enemy's shore, directing all the fire on himself. Having reached the opposite bank, Comrade Nemchikov began to fighting with the Finns with his machine gun and move towards the enemy's trenches. A group of 12 people provided the battalion with a crossing into the Svir River and the battalion completed its task successfully."

Commander of the 300th Guards Rifle Regiment
Guards Colonel Danilov

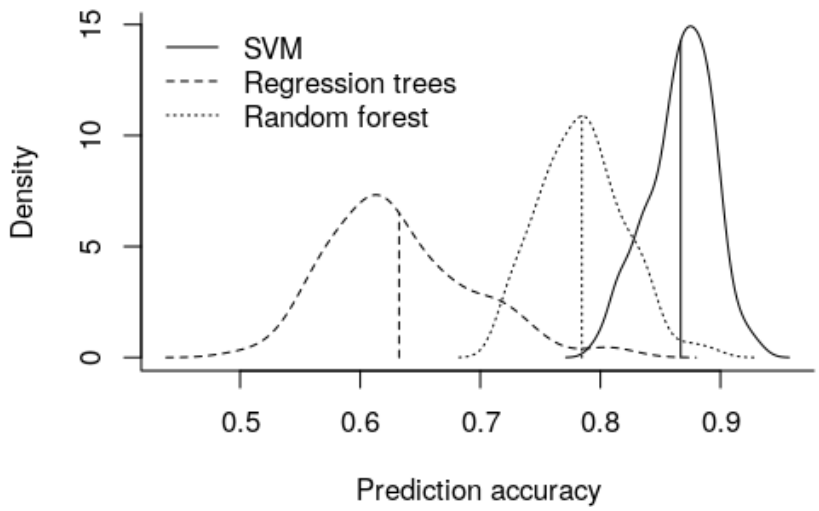
A4. Nationality classification

To develop a Soviet nationality classifier, we used the Memorial archive as a training set. The archive contains nationality information for 916,675 arrestees, with 163,284 unique surnames. The set of nationalities includes: Armenian, Belarussian, Chechen, Chinese, Estonian, Greek, Jewish, Kabardin, Kalmyk, Korean, Latvian, Lithuanian, Ossetian, Polish, Russian, Tatar and Ukrainian. Because the same surnames reappear multiple times in the archive, often with more than one nationality (due to intermarriage or other reasons), dictionary-based matching of each surname to its corresponding nationality is not feasible. To account for the uncertainty induced by this one-to-many match problem, we used

three supervised machine learning algorithms to create a classifier that matches each surname to its most-likely nationality. These classifiers are: Support Vector Machine (SVM), Regression Trees and Random Forest.

Due to the computational burden of fitting these models on a document-term matrix with 163,284 columns, we split the task into chunks of 1,000, and iterated over them. In each iteration, we created an $N \times 1000$ document-term matrix, where N is the number of individuals in Memorial who had one of the 1000 surnames in that chunk. We then fit a model, where the outcome is a $N \times 1$ vector of nationalities for each individual, and the explanatory variables are the 1000 unique surnames. We then calculated average classification accuracy for each algorithm (% of surnames correctly predicted). Because the set of surnames is fixed at 163,284, and extrapolation is not possible, we report only in-sample prediction accuracy below. Figure A4.2 reports the distribution of these accuracy scores for the three algorithms, with vertical lines showing the mean. SVM clearly outperforms the others, with regression trees faring the worst.

Figure A4.2: DISTRIBUTIONS OF CLASSIFIER ACCURACY SCORES.



Breaking these statistics down by nationality, we see that some groups (e.g. Russian) have fairly high accuracy scores (96.5% with SVM, 99.3% with Regression Trees, 94.7% with Random Forests). Others, like Ukrainians and Belarussians, don't score quite as high, even with SVM – likely due to intermarriage and similarity of surnames among the three biggest Slavic republics. In most erroneous cases, Belarusians and Ukrainians are typically mis-classified as Russians. For this reason, our analyses employ only a binary "ethnic Russian" variable, rather than using the full set of predicted ethnicities.

To assess the validity of our SVM classifications of soldiers' nationalities, we compared oblast-level proportions against census data from 1939. To do so, we spatially matched

the census data to 1937 oblasts, and calculated oblast-level proportions for each nationality listed above. We then calculated oblast-level proportions of soldiers' SVM-classified nationalities, and compared these proportions to those in the 1939 census.

Table A4.2 reports the distribution of test statistics and p-values from Wilcoxon signed-rank tests, conducted country-wide, and by each oblast. The null hypothesis is that the distributions of oblast-level proportions across nationalities (e.g. Russian = .70, Ukrainian = .05, etc.) are the same for census and SVM data. These results suggest that – for all regions except Dagestan, the most ethnically diverse one – we cannot reject the null, and therefore the oblast-level proportions were likely drawn from the same distribution.

Table A4.2: WILCOXON SIGNED-RANK TEST STATISTICS FOR NATIONALITY CLASSIFIER.

Oblast' (1937)	Wilcoxon Test	Oblast' (1937)	Wilcoxon Test
RSFSR	157	Leningradskaya Oblast'	162
Azovo-Chernomorskiy Kray	171	Mariyskaya ASSR	106
Bashkirskaya ASSR	113	Moskovskaya Oblast'	209
Buryato-Mongol'skaya ASSR	155	Saratovskaya Oblast'	155
Dagestanskaya ASSR	99*	Severnaya Oblast'	158
Gor'kovskaya Oblast'	124	Stalingradskaya Oblast'	150
Ivanovskaya Oblast'	124	Sverdlovskaya Oblast'	131
Kalininskaya Oblast'	144	Udmurtskaya ASSR	120
Kalmytskaya ASSR	121	Voronezhskaya Oblast'	121
Karel'skaya ASSR	162	Yakutskaya ASSR	157
Kirovskaya Oblast'	155	Yaroslavskaya Oblast'	153
Komi ASSR	157	Zapadnaya Oblast'	133
Kurskaya Oblast'	124	Zapadno-Sibirskiy Kray	148

Null hypothesis is that the distributions of oblast-level proportions across nationalities are the same for census and SVM data. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

A5. ERRORS IN THE MEASUREMENTS OF OUTCOMES

We measure battlefield outcomes using soldiers' reported discharge reasons, which are proxies rather than direct observations of theoretically relevant quantities. For example, we use K/WIA as a proxy for resolve, even if many soldiers' deaths had little to do their actual levels of resolve. Here we evaluate the direction and degree of bias introduced by this kind of measurement error. We use the example of K/WIA as a proxy measure for having battlefield resolve, but the argument equally applies to other outcomes.

Let $Y_i^* \in \{0, 1\}$ denote whether or not soldier i displayed battlefield resolve, which we cannot observe directly, and let $Y_i \in \{0, 1\}$ denote whether or not that soldier was

K/WIA, which we do observe. We formalize the measurement process that links observed outcome Y_i with the latent quantity Y_i^* as

$$\Pr(Y_i = 1 | Y_i^* = s, \boldsymbol{X}_i) = \varepsilon_s(\boldsymbol{X}), \tag{3}$$

for $s \in \{0, 1\}$. The vector $\boldsymbol{X}_i$ represents the covariates that potentially affect the probability of K/WIA independently of the soldier's resolve, and $\varepsilon_1(\boldsymbol{X})$ and $\varepsilon_0(\boldsymbol{X})$ are measurement errors for soldiers with and without resolve, respectively.

Let D denote the level of repression and let $p^*(D, \boldsymbol{X}_i) = \Pr(Y_i^* = 1 | \boldsymbol{X}_i, D)$ and $p(D, \boldsymbol{X}_i) = \Pr(Y_i = 1 | \boldsymbol{X}_i, D)$ denote the probability that a soldier, conditional on repression and other covariates, has battlefield resolve or that he is K/WIA, respectively. We can estimate only the latter, but are interested in the former. By the law of iterated expectations we get

$$p^*(D, \boldsymbol{X}_i) = \frac{p(D, \boldsymbol{X}_i) - \varepsilon_0(\boldsymbol{X}_i)}{\varepsilon_1(\boldsymbol{X}_i) - \varepsilon_0(\boldsymbol{X}_i)}, \tag{4}$$

and so the marginal change in the probability that the soldier has battlefield resolve when repression increases is equal to

$$\frac{\partial}{\partial D} p^*(D, \boldsymbol{X}_i) = \frac{\partial}{\partial D} p(D, \boldsymbol{X}_i) \frac{1}{\varepsilon_1(\boldsymbol{X}_i) - \varepsilon_0(\boldsymbol{X}_i)}. \tag{5}$$

The partial derivatives on both sides are *in the same direction* if and only if $\Pr(Y_i = 1 | Y_i^* = 1, \boldsymbol{X}_i) > \Pr(Y_i = 1 | Y_i^* = 0, \boldsymbol{X}_i)$. The second term in the above equation is always strictly larger than one, and so the marginal effect on p is always smaller in absolute value than the marginal effect on p^*. Thus, under the assumption that a soldier who is more willing to fight has a greater chance of being K/WIA, which seems highly plausible, the measurement error in the outcome results in attenuation bias.

A5.1. Measurement validation through unit-level operational performance

To further assess how well our measures map onto the theoretical concept of resolve, we assess whether these individual-level battlefield outcomes aggregate to the operational-level success or failure of army units. Specifically, we looked at how predictive these measures are of territorial gains by the Red Army.[7] To conduct this analysis, we matched sol-

[7] A potential alternative measure of military effectiveness is the loss-exchange ratio (LER) between Soviet and German forces (i.e. enemy losses divided by friendly losses). We do not consider the LER here because (1) the Russian MOD has not made these statistics available at the battle level, (2) there is little evidence that Soviet commanders cared about the LER or used it as a metric of success, and (3) such an analysis would be almost tautological, with Soviet casualty statistics appearing on both the left and right

diers' records to the 225 major battles listed in the "People's Memory" database (`pamyat-naroda.ru/ops/`), using information on the army units in which they served and their months of service in those units. Because these battles were large, army-level operations, this linkage procedure required first establishing the "parent" army for each division, regiment, battalion or company listed in the soldier's service history, and then filtering the records to include only those corresponding to the time of the battle. We then calculated the proportion of soldiers in each unit-month with each type of outcome (K/WIA, MIA, etc.).

To measure operational-level territorial gains, we conducted a text analysis of battle descriptions in "People's Memory", each approximately one paragraph in length. Rather than providing our own subjective assessment of battlefield success, this approach allows us to adopt the Russian MOD's own official characterization of events, which is more likely to reflect Soviet commanders' information set at the time. We read each description and classified it as denoting a territorial gain, loss, or no change in the status quo.

Because a small subset of descriptions were open to multiple interpretations (e.g. with Soviet troops advancing on one sector of the front, but retreating elsewhere), we accounted for measurement uncertainty by fitting a supervised machine learning model, with the manually-coded labels as a training set. Specifically, we employed a recurrent neural network (RNN) model with long short-term memory (LSTM) (Hochreiter and Schmidhuber, 1997). LSTMs are well-suited for learning problems related to sequential data, such as sequences of words of differential length, where the vocabulary is potentially large, and where context and dependencies between inputs are potentially informative for classification.[8] We employed a standard "vanilla LSTM" architecture (Graves and Schmidhuber, 2005), using the `keras` library in Python 3 (Chollet, 2015).[9]

Because our training set includes all 225 battles, we used 10 random subsets of these labels to train the model, setting aside the remainders for cross validation. This created 10

side of the equation. By contrast, there is ample evidence that Soviet authorities used territorial changes as measures of effectiveness, as illustrated by the fact that nearly all battle descriptions in "People's Memory" mention them, and by the fixation on this metric in Stalin's wartime orders (e.g. "Not one step back!").

[8]For an introduction to LSTMs, with applications to political science, see Chang and Masterson (2020).

[9]At the center of this architecture is a memory cell and non-linear gating units, which regulate information flow into and out of the cell. A "vanilla LSTM" block features three gates (input, forget, and output), block input, a single cell, and an output activation function. The block's output recurrently connects back to the block input and all gates. Greff et al. (2017) demonstrated that this architecture performs well on a variety of classification tasks, and that common modifications do not significantly improve performance. To preprocess the text, we mapped each of the 225 desciptions into a real vector domain, with each word represented as an embedding vector of length 100. The purpose of this step is to encode words as real-valued vectors in a high dimensional space, where words more similar in meaning appear closer in the vector space. We limited the total number of words used in modeling to the 5000 most frequent ones. We used an LSTM layer with 100 memory units, and a dense output layer with a sigmoid activation function for binary predictions. We fit the model using the efficient ADAM optimization algorithm, with binary cross-entropy as the loss function.

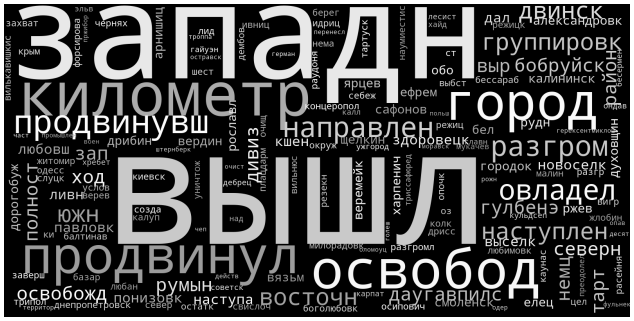


Figure A5.3: Word Cloud of Event Descriptions for Territorial Gains.

alternative sets of LSTM-classified battles, of which we retained the set with the highest out-of-sample predictive accuracy, as measured by the area under the Receiver Operator Characteristic (ROC) curve. In most instances, the network achieved convergence at < 100 epochs, with median predictive accuracy (area under ROC curve) of 0.94.

Figure A5.3 show word clouds for event descriptions corresponding to territorial gains. The font size is proportional to word frequencies in the LSTM-predicted test set, for events that LSTM predicted as being most likely to belong to this category (99th percentile). The word frequencies generally align with our qualitative understanding of territorial gains, with terms such as "advance" (the stems продвинул, вышл), "liberate" (освобод) and "to the west" (западн) featuring prominently.

After linking the 225 battles to our unit-month level data, we regressed territorial gains on the proportion of participating soldiers K/WIA, DDT, PUN, POW, MIA, and Medals, along with fixed effects for units, years, and months. The results in Table **??** (main text) correspond to the hand-coded version of these battle outcomes. However, estimates are numerically almost identical if we use the LSTM-predicted labels.

A6. PROBING IDENTIFYING ASSUMPTIONS

A6.1. Tests of Complete Spatial Randomness

A key identifying assumption for our OLS analyses is that the local geographic distribution of arrest locations is spatially random. This assumption does not preclude the existence of geographic clusters on a more macro scale, or require a uniform distribution of events across the country – more arrests will surely happen in densely populated areas than in the desert or the tundra. What it assumes is that, after accounting for differences between small geographic areas (e.g. by estimating a fixed effect for each 25×25 km cell), we can treat remaining geographic variation within each of these areas as random. To test

the validity of this assumption, we performed a series of tests of the null hypothesis that arrest locations are a realization of a uniform Poisson point process, including Quadrat Count Tests, Clark-Evans Tests, and Spatial Scan Tests.

The set of arrest locations within each grid cell represents a spatial point pattern, whose observed arrangement may be random (H_0) or the result of some non-random targeting process (H_A) (e.g. targeting of neighborhoods whom authorities suspect of disloyalty). Complete Spatial Randomness (CSR) requires that (a) events have an equal probability of occurring in any equally-sided subdivision of a region (i.e. if a grid cell is split into 4 tiles, an event has a $1/4$ chance of occurring in each tile), and (b) the locations of these events are independent of one another. If the CSR null hypothesis is true, and the point pattern is a realization of a random Poisson process, then the expected density of points (intensity of arrests) within grid cell j should be:

$$\lambda_j = \frac{n_j}{a_j} \tag{6}$$

where n_j is the total number of observed events within grid cell j and a_j is j's geographic area. If we divide j into K tiles of equal shape and area $(j_1, \ldots, j_K)$, then the expected number of points in any given tile j_k should depend only on the overall point density within j and the relative area of the tile:

$$E[N(j_k)] = \lambda_j a_{j_k} = n_j \frac{a_{j_k}}{a_j} \tag{7}$$

The *Quadrat Count Test* (Cressie and Read, 1984) tests the CSR hypothesis by partitioning grid cell j into rectangular tiles of equal area, and compares the observed tile count distribution (i.e. number of tiles with 0, 1, 2, ... events) against the distribution we would expect if these counts were independent random samples from a Poisson distribution with rate parameter λ_j. It then uses a Pearson's χ^2 goodness-of-fit test to quantify the difference between the observed and expected counts, with p-values calculated using Monte Carlo methods (i.e. generating 2000 random point patterns from $Poisson(\lambda_j)$ and comparing the χ^2 statistic for the observed point pattern against the simulated values).

We performed a series of Quadrat Count Tests for each of the 29,243 25×25 km grid cells in our study region, divided into $K \in \{1, 2^2 = 4, 3^2 = 9, 4^2 = 16, 5^2 = 25\}$ tiles of size 25×25 km, 12.5×12.5 km, 8.33×8.33 km, 6.25×6.25 km, and 5×5 km, respectively.

As Table A6.3 reports, we were unable to reject the null hypothesis (χ^2 test statistic p-value was greater than 0.05) in 91 percent of tests, including 100 percent of tests at $K = 1$ and 88 percent at $K = 25$. These results suggest that – for the vast majority of the grid cells

Table A6.3: QUADRAT COUNT TEST STATISTICS BY NUMBER OF TILES PER CELL.

Tiles per cell	Average χ^2 stat.	Average p value	$E[p > .05]$
1	0.00	0.98	(100%)
4	6.36	0.57	(91%)
9	14.95	0.49	(89%)
16	26.63	0.61	(89%)
25	38.21	0.62	(88%)
Overall	17.23	0.66	(91%)

Values represent average Pearson χ^2 test statistics and two-sided p values for Monte Carlo Quadrat Count Tests, calculated with each 25×25 km grid cell divided into different numbers of tiles: 1 tile (25 km across), 4 tiles (12.5 km), 9 (8.33 km), 16 (6.25 km) and 25 (5 km).

in our sample, across all partitions – there is no significant difference between observed and expected local event counts.

We supplemented these analyses with a pair of alternative approaches, which use distances between event locations (rather than fixed areal partitions) to calculate test statistics. Among these are the *Clark-Evans Test* (Clark and Evans, 1954), which uses a Normal approximation of the distribution of nearest-neighbor distances D within region j, with mean and variance

$$E[D_j] = \mu_j = \frac{1}{2\sqrt{\lambda_j}}, \quad var(D_j) = \sigma_j^2 = \frac{4-\pi}{4\pi\lambda_j} \tag{8}$$

where λ_j is the point density within grid cell j. To compare the observed distribution of distances to what we would expect under CSR, the test calculates a z-value:

$$z_j = \frac{\bar{d}_j - \mu_j}{\sigma_j} \tag{9}$$

where $\bar{d}_j$ is the sample mean of nearest-neighbor distances within grid cell j. Under the CSR null hypothesis, z_j should be a sample from $N(0,1)$. p-values are based on a two-tailed test, where significantly small values of $\bar{d}_j$ indicate spatial clustering and significantly large values indicate spatial dispersion.

Finally, we performed *Spatial Scan Tests* for clustering in spatial point patterns (Kulldorff, 1997). This test rejects the null CSR hypothesis if there exists a circle of radius r within grid cell j, which contains significantly more points than one would expect under a uniform Poisson process. The alternative hypothesis is that of an inhomogeneous Poisson process with different intensities $\beta_1\lambda_j$ within the circle, and $\beta_2\lambda_j$ outside the circle.

Table A6.4: TESTS OF COMPLETE SPATIAL RANDOMNESS.

Test	Average test stat.	Average p value	$E[p > .05]$
Quadrat	17.23	0.66	(91%)
Clark-Evans	0.55	0.55	(87%)
Spatial Scan	3.69	0.83	(96%)

Values represent test statistics and p values for Monte Carlo Quadrat Count Tests, two-tailed Clark-Evans Tests, and Spatial Scan Tests, averaged across all grid cells.

As Table A6.4 reports, the results of these additional CSR tests were consistent with those of the Quadrat Test. We were unable to reject the null hypothesis in 87% of grid cells with the Clark-Evans test and 96% with the Spatial Scan test. In the vast majority of grid locations, the spatial distribution of arrests does not significantly deviate from what we would expect under Complete Spatial Randomness.

In what types of grid cells is the local CSR assumption most likely to be violated? To explore this question, we regressed Quadrat Test p-values on several grid cell-level characteristics, including size of the local population in the 1926 Census, number of urban settlements, and distance to the nearest railroad line in 1937, while controlling for test parameters (number of tiles). These regressions suggest that the p-value is decreasing (i.e. CSR more likely to be violated) in urban locations with larger populations, located closer to the railroad. Doubling the population size within the a 25×25 km grid cell reduces the p-value, on average, by 0.016 (-0.017, -0.014); adding one urban settlement to the grid cell reduces the p-value by 0.028 (-0.036, -0.021). Doubling the distance to the railroad, meanwhile, increases the p-value by 0.024 (0.022, 0.026) on average. As Table A6.5 shows, all three sets of CSR tests suggest greater spatial clustering in grid cells with direct railroad access, with consistently lower average p-values in cells intersected by a railroad in 1937.

Table A6.5: RAILROAD ACCESS AND COMPLETE SPATIAL RANDOMNESS.

Test	$E[p\|\text{rail}]$	$E[p > .05\|\text{rail}]$	$E[p\|\text{no rail}]$	$E[p > .05\|\text{no rail}]$
Quadrat	0.54	(83%)	0.69	(94%)
Clark-Evans	0.40	(79%)	0.60	(90%)
Spatial Scan	0.39	(79%)	0.86	(97%)

Values represent average p values of CSR test statistics for grid cells with direct railroad access (rail) and without railroad access (no rail).

These patterns make logical sense, seeing as the population in urban, densely-populated

areas – and the pool of potential arrest targets – is likely to live more closely together, in apartment blocks and other residential clusters. As we detail elsewhere in the paper, arrests also tended to cluster around railroads for various logistical and economic reasons salient to the NKVD. We exploit this latter variation in our instrumental variable design.

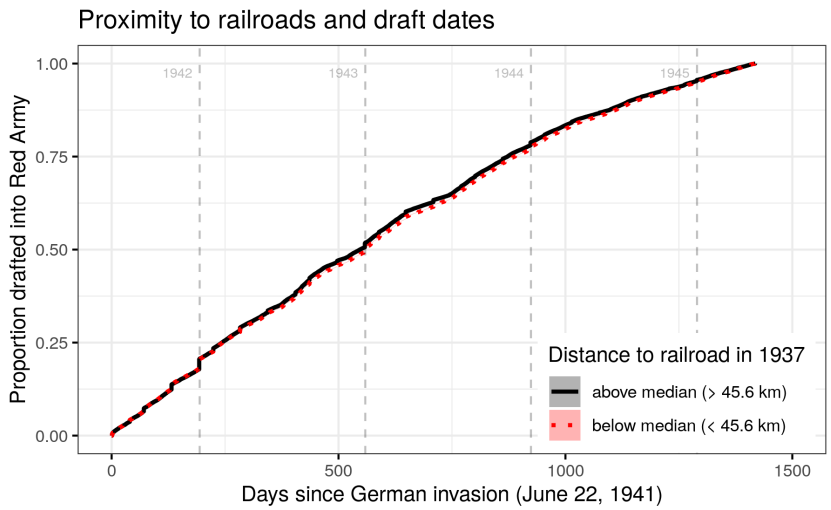
A6.2. Railroad access and draft dates

The validity of the railroad instrument depends in part on the assumption that railroad access at individuals' birth locations did not affect the battlefield conditions they faced upon being drafted. This assumption would be violated if, for instance, individuals living closer to railroads were drafted earlier in the war, during Germany's summer offensive of 1941 or before Stalin issued orders for stricter troop discipline.

To assess the plausibility of this scenario, Figure A6.4 reports non-parametric Kaplan-Meier estimates of the proportion of soldiers drafted by each day in the war.[10] The two curves correspond to soldiers born in locations with above- and below-median values of $\hat{f}\,(\text{Raildist}_j)$, corresponding to $\text{Raildist}_j = 45.6$km. The two curves overlap almost perfectly until about mid-1942, at which point they slightly diverge, with soldiers born closer to railroads being *less* likely to have been drafted by any given date. The two lines converge yet again in 1945. The median draft dates of soldiers in the two groups were just over a week apart – February 2, 1943 for soldiers born closer to railroads, and January 23, 1943 for those born further away. In sum, there is no evidence that the proximity of one's birth location to railroads systematically affected the timing of one's draft date.

[10]This analysis includes only individuals drafted between June 22, 1941 and May 9, 1945, for whom draft dates are available, with time precision at the daily level ($N = 5,924,878$, or 51% of full sample).

Figure A6.4: Kaplan-Meier Estimates of Cumulative Proportion Drafted



A7. Robustness checks

A7.1. Clustered treatment assignment

To address estimation problems due to clustered treatment assignment, we took three approaches: (1) aggregate analysis of cluster-level averages, (2) pair-matched cluster sampling, and (3) both. The first of these addresses the problem of correlated errors within clusters. The second corrects for biases due to over-weighting larger clusters. The third combines the two approaches for an even more conservative set of estimates.

A7.1.1 Cluster-level analysis We may worry that conventional standard errors are downwardly biased due to the presence of correlated errors within clusters. In our main OLS/2SLS/RDD analyses, we address this issue by reporting robust clustered standard errors (RCSE). Here, we go one step further, by conducting an aggregate, cluster-level analysis, which addresses the issue of correlated disturbances by eliminating within-cluster variation altogether (Green and Vavreck, 2008).

Our aggregate analyses adopt the same core specification as our main models (equations **??–??** in main text), replacing y_{ij} with $\bar{y}_j$ (average of individual outcomes for cluster j), and $\boldsymbol{X}_{ij}$ with $\bar{\boldsymbol{X}}_j$ (cluster-level averages of pre-treatment covariates). Because cluster-level averages are more precisely estimated for clusters containing more individuals, we weighted each observation by cluster size.

The results of the cluster-level analyses are in Table A7.6. Estimates are substantively consistent with the individual-level results in the main text.

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medals
Model	OLS						
Coefficient	0.5 (0.1)**	-0.3 (0.1)**	-0.3 (0.05)**	-0.02 (0.04)	0.003 (0.002)$^{'}$	-0.02 (0.005)**	-0.1 (0.04)**
Mean Y	17.4	32.2	17.1	14.1	0.2	0.8	15.8
Gridcells	12,143	12,143	12,143	12,143	12,143	12,143	12,143
Birthplaces	182,225	182,225	182,225	182,225	182,225	182,225	182,225
Model	2SLS (First-stage $\mathcal{F} = 147.6$)						
Coefficient	4.4 (0.6)**	-2.4 (0.4)**	-2.4 (0.4)**	-0.01 (0.3)	0.03 (0.01)**	-0.1 (0.03)$^{'}$	-1.7 (0.3)**
Mean Y	17.2	31.8	17.4	13.4	0.2	0.9	15.9
Gridcells	5,654	5,654	5,654	5,654	5,654	5,654	5,654
Birthplaces	147,803	147,803	147,803	147,803	147,803	147,803	147,803
Model	FRDD (First-stage $\mathcal{F} = 24.3$)						
Coefficient	1.5 (0.4)**	-1.2 (0.3)**	-1.4 (0.3)**	0.3 (0.2)$^{'}$	0.01 (0.01)	-0.04 (0.03)	-0.6 (0.2)*
Mean Y	16.5	31.3	17.8	12.5	0.2	0.8	16
Gridcells	1,582	1,582	1,582	1,582	1,582	1,582	1,582
Birthplaces	44,154	44,154	44,154	44,154	44,154	44,154	44,154

Coefficients represent estimated effect of repression on percent of locally-drafted soldiers with each type of battlefield outcome. All clusters. Outcomes on percentage scale (0 to 100). Robust standard errors in parentheses. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by number of draftees. 2SLS analyses exclude birth locations > 100km from railroad. FRDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.6: CLUSTER-LEVEL ANALYSIS (ALL CLUSTERS).

A7.1.2 Matched cluster sampling Our main individual-level OLS/2SLS/RDD analyses employ a the full sample of 11M+ soldiers, with cluster-level (birth location) exposure to repression. The clusters are geographic coordinates of soldiers' birth locations. The sample of 11M represents roughly a third of all soldiers who served in the Red Army during WWII, and exclused records with missing information on birth locations as well as those born in other Soviet republics outside the RSFSR. We assume that this missingness is random, and that we can treat the 11M individuals as a simple random sample. Under simple random sampling, however (e.g. take sample of 11M troops from across all clusters), individuals from larger clusters are more likely to appear in the sample than those from small clusters. This is a problem because (a) treatment is assigned at the cluster level, and (b) cluster size is potentially correlated with treatment (i.e. more arrests occurred in higher-population areas). One way to address this issue is to adopt a pair-matched clus-

ter sampling design, which selects pairs of clusters that are as similar to each other as possible on observable pre-treatment covariates, including cluster size (Imai et al., 2009).

Our sampling strategy is a variant of one-stage cluster sampling, where the primary sampling unit is the cluster, and the secondary sampling unit is the soldier.[11] Let $j \in \{1, \ldots, J\}$ index the $J = 183,354$ clusters (birth locations). Rather than sampling these clusters with equal probability, as in a standard cluster random sample, we select a subset $J^{(m)}$, where $J^{(m)}/2$ of the clusters are "treated" (i.e. high level of repression) and another $J^{(m)}/2$ are "control" clusters (low repression) that are well-balanced on all observable pre-treatment cluster-level covariates $\boldsymbol{X}_j$.[12] The covariates in $\boldsymbol{X}_j$ include the same birth location-level covariates we use in out main analysis (distance to the nearest district administrative center, the number of collective farms and hectares of arable land within 10 km), along with cluster-level averages of local soldiers' ethnicity (proportion Russian) and age (average birth date).[13] We also matched exactly on grid cell and cluster size (quantile of number of draft records from location j). This last step ensures that within-cluster sample average treatment effects are uncorrelated with differences in cluster sizes within each matched-pair (Imai et al., 2009, p. 36).

Table A7.7: CLUSTER SAMPLE SIZE.

Status	Number of Clusters	Treated	Control
pre-matching	183,354	80,256	103,098
post-matching	43,118	21,559	21,559

Treated (Control) clusters are ones where the number of arrests is above (below) the grid cell median.

Tables A7.7 and A7.8 report the number of clusters pre- and post-matching (J vs. $J^{(m)}$) and corresponding covariate balance statistics. The matching procedure yielded a sample of 43,118 clusters (23.5% of full sample), or 21,559 matched pairs. The procedure,

[11] In a one-stage cluster sample, all soldiers within the sampled clusters remain in the sample, regardless of cluster size. We also replicated our results with a two-stage cluster sampling design, in which soldiers within sampled clusters are sampled with equal probability. The two-stage approach ensures that cluster samples are of equal size, at the expense of a reduction in statistical power. Results were similar to one-stage cluster sampling, but more weakly powered, suggesting that the selection of clusters is much more consequential than secondary sampling of soldiers within clusters.

[12] Matching requires transforming our non-negative integer treatment variable (number of arrests) into a dichotomous indicator, where clusters above some threshold of arrests are "treated" and those below it are "control". Because we are interested in local variation in repression, we allowed this threshold to vary by grid cell, such that locations above their grid cell median are "treated" and the rest are "control." Changing this thresholding rule from "grid cell median" to "grid cell mean," "regional median/mean" or "national median/mean" did not substantively change the results, apart from reducing the matched sample size.

[13] We used Mahalanobis distance matching for these covariates.

Table A7.8: COVARIATE BALANCE STATISTICS, PRE- AND POST-MATCHING.

Covariate	Status	Mean Treated	Mean Control	Std. Diff.	KS Statistic
GRID ID	pre-matching	3297.754	4394.818	-0.307	0.08**
	post-matching	2725.942	2725.942	0	0
Population quantile	pre-matching	2.466	2.41	0.048	0.021**
	post-matching	2.829	2.829	0	0
Ethnic Russian	pre-matching	0.868	0.861	0.03	0.016**
	post-matching	0.942	0.943	-0.008	0.013*
Date of birth	pre-matching	1914.701	1914.733	-0.006	0.014**
	post-matching	1915.113	1915.106	0.002	0.007
Cropland within 10km	pre-matching	7768.185	7018.57	0.12	0.07**
	post-matching	7497.565	7619.092	-0.02	0.011
State farms within 10km	pre-matching	0.202	0.173	0.064	0.023**
	post-matching	0.198	0.203	-0.01	0.006
Distance to district center	pre-matching	20.969	36.047	-0.634	0.191**
	post-matching	17.859	19.82	-0.119	0.175**
Distance to road junction	pre-matching	46.296	61.569	-0.439	0.065**
	post-matching	41.985	42.798	-0.028	0.045**

Standardized difference (Std. Diff.) defined as $\frac{\text{mean}(T)-\text{mean}(C)}{\text{sd}(T)}$. Kolmogorov-Smirnov (KS) test statistics based on 1,000 bootstrapped samples. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

by design, achieves perfect balance on grid cells and cluster size (population quantile). Balance on remaining covariates is also greatly improved, with all standardized differences falling below the conventional .25 threshold (Ho et al., 2007). Although absolute and standardized differences are numerically small, bootstrapped Kolmogorov-Smirnov statistics indicate that there is some remaining imbalance on ethnicity and distance to district center. We address this imbalance by controlling for these and all other pre-treatment covariates in our analysis.

Table A7.9 reports individual-level analyses for the matched cluster sample. These results align closely with those we report in the main text.

A7.1.3 Matched cluster-level analysis Table A7.10 reports the most conservative set of estimates: an aggregate, cluster-level analysis that uses only the matched cluster pairs discussed above. Estimates align in direction and statistical with significance with those in Tables A7.6, A7.9, and the main text.

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medals
Model	OLS						
Coefficient	0.3 (0.1)**	-0.2 (0.1)**	-0.2 (0.1)**	0.01 (0.05)	0.003 (0.003)	-0.01 (0.01)$^{'}$	-0.1 (0.1)
Mean Y	20.9	23.3	17.2	5.3	0.1	0.8	18.2
Gridcells	4,095	4,095	4,095	4,095	4,095	4,095	4,095
Birthplaces	43,120	43,120	43,120	43,120	43,120	43,120	43,120
Soldiers	4,487,712	4,487,712	4,487,712	4,487,712	4,487,712	4,487,712	4,487,712
Model	2SLS (First-stage $\mathcal{F} = 68.2$)						
Coefficient	3.8 (0.8)**	-1.6 (0.7)*	-1.8 (0.5)**	0.2 (0.4)	0.03 (0.02)*	-0.1 (0.04)$^{'}$	-2.4 (0.5)**
Mean Y	20.5	23.6	17.3	5.5	0.1	0.9	18.2
Gridcells	3,343	3,343	3,343	3,343	3,343	3,343	3,343
Birthplaces	38,893	38,893	38,893	38,893	38,893	38,893	38,893
Soldiers	4,088,306	4,088,306	4,088,306	4,088,306	4,088,306	4,088,306	4,088,306
Model	FRDD (First-stage $\mathcal{F} = 36.5$)						
Coefficient	1.7 (0.4)**	-0.5 (0.3)	-1.1 (0.3)**	0.7 (0.2)**	0.02 (0.01)**	-0.03 (0.02)	-1.7 (0.4)**
Mean Y	20.5	23.5	17.3	5.3	0.1	0.9	17.7
Gridcells	1,066	1,066	1,066	1,066	1,066	1,066	1,066
Birthplaces	13,488	13,488	13,488	13,488	13,488	13,488	13,488
Soldiers	1,242,299	1,242,299	1,242,299	1,242,299	1,242,299	1,242,299	1,242,299

Outcomes on percentage scale (0 to 100): killed or wounded in action (KIA/WIA); missing in action (MIA), becoming prisoner of war (POW), defecting, deserting, committing treason (DDT), being punished for battlefield misconduct (PUN), or any of the above (Flee); receiving a valor decoration (Medal). Standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record clustering probability. 2SLS analyses exclude birth locations > 100km from railroad. FRDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.9: ANALYSIS OF MATCHED CLUSTER SAMPLE

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medals
Model	OLS						
Coefficient	0.4 (0.1)**	-0.3 (0.1)**	-0.3 (0.1)**	0.03 (0.1)	0.003 (0.003)	-0.02 (0.01)*	-0.1 (0.1)$^{'}$
Mean Y	16.5	29.4	18.5	9.9	0.1	0.8	17.8
Gridcells	4,095	4,095	4,095	4,095	4,095	4,095	4,095
Birthplaces	43,120	43,120	43,120	43,120	43,120	43,120	43,120
Model	2SLS (First-stage $\mathcal{F} = 65.5$)						
Coefficient	4.3 (0.9)**	-1.7 (0.7)*	-2.2 (0.6)**	0.5 (0.4)	0.03 (0.01)*	-0.1 (0.05)	-1.7 (0.4)**
Mean Y	16.4	29.5	18.4	10.2	0.1	0.8	17.6
Gridcells	3,343	3,343	3,343	3,343	3,343	3,343	3,343
Birthplaces	38,893	38,893	38,893	38,893	38,893	38,893	38,893
Model	FRDD (First-stage $\mathcal{F} = 30.7$)						
Coefficient	2.3 (0.5)**	-1.1 (0.4)**	-1.6 (0.4)**	0.6 (0.2)**	0.01 (0.01)	-0.1 (0.03)$^{'}$	-1 (0.3)**
Mean Y	15.9	29.3	18.4	9.9	0.2	0.8	17.8
Gridcells	1,066	1,066	1,066	1,066	1,066	1,066	1,066
Birthplaces	13,488	13,488	13,488	13,488	13,488	13,488	13,488

Coefficients represent estimated effect of repression on percent of locally-drafted soldiers with each type of battlefield outcome. Matched cluster sample. Outcomes on percentage scale (0 to 100). Robust standard errors in parentheses. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by number of draftees. 2SLS analyses exclude birth locations > 100km from railroad. FRDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.10: CLUSTER-LEVEL ANALYSIS (MATCHED CLUSTER PAIRS).

A7.2. Measurement error due to incomplete records

Another robustness check explores the possibility that measurement error due to incomplete records is driving our results. Tables A7.11-A7.12 replicates the earlier OLS, 2SLS and RDD analyses on the full and matched cluster samples, respectively, while excluding individuals whose reseasons for discharge are not observed. As the results show, after we drop the more ambiguous cases of draftees without observed terminal histories, the estimated coefficients are in the same direction as in our baseline specifications and they increase in absolute value, sometimes considerably. Results for wartime decorations and promotions are identical to those in the main paper because information for these variables comes from a separate set of archival materials, and does not require observing discharge records. These results indicate that – in most cases – measurement error is likely to bias our estimates downwards.

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medals
Model	OLS						
Coefficient	0.7 (0.1)**	-0.8 (0.1)**	-0.6 (0.1)**	-0.2 (0.1)*	0.005 (0.005)	-0.04 (0.01)**	-0.2 (0.1)**
Mean Y	46.7	53	38.3	12.8	0.4	1.7	17.9
Gridcells	11,514	11,514	11,514	11,514	11,514	11,514	12,143
Birthplaces	157,205	157,205	157,205	157,205	157,205	157,205	182,225
Soldiers	5,180,859	5,180,859	5,180,859	5,180,859	5,180,859	5,180,859	11,427,895
Model	2SLS (First-stage $\mathcal{F} = 141.6$)						
Coefficient	7.3 (0.9)**	-7.4 (0.9)**	-6.2 (0.7)**	-1.1 (0.6)$^{'}$	0.1 (0.03)**	-0.2 (0.1)**	-2.1 (0.4)**
Mean Y	45.5	54.2	39.2	13	0.4	1.8	18
Gridcells	5,580	5,580	5,580	5,580	5,580	5,580	5,654
Birthplaces	127,822	127,822	127,822	127,822	127,822	127,822	147,803
Soldiers	4,369,299	4,369,299	4,369,299	4,369,299	4,369,299	4,369,299	9,725,085
Model	FRDD (First-stage $\mathcal{F} = 29.3$)						
Coefficient	2.1 (0.5)**	-2.1 (0.5)**	-3 (0.5)**	1 (0.3)**	0.04 (0.02)*	-0.1 (0.05)*	-1.2 (0.2)**
Mean Y	44.2	55.5	40.9	12.3	0.4	2	18.1
Gridcells	1,561	1,561	1,561	1,561	1,561	1,561	1,582
Birthplaces	37,947	37,947	37,947	37,947	37,947	37,947	44,154
Soldiers	996,683	996,683	996,683	996,683	996,683	996,683	2,316,908

All records. OLS, 2SLS and FRDD estimates. Outcomes on percentage scale (0 to 100). Robust standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability. 2SLS analyses exclude birth locations > 100km from railroad. FRDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.11: ESTIMATES FOR SOLDIERS WITH OBSERVED DISCHARGE RECORDS.

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medals
Model				OLS			
Coefficient	0.7 (0.2)**	-0.7 (0.2)**	-0.6 (0.2)**	-0.1 (0.1)	0.01 (0.01)	-0.04 (0.02)$'$	-0.1 (0.1)
Mean Y	47.1	52.6	38.7	11.8	0.3	1.9	18.2
Gridcells	4,023	4,023	4,023	4,023	4,023	4,023	4,095
Birthplaces	38,638	38,638	38,638	38,638	38,638	38,638	43,120
Soldiers	1,990,881	1,990,881	1,990,881	1,990,881	1,990,881	1,990,881	4,487,712
Model				2SLS (First-stage $\mathcal{F} = 67.1$)			
Coefficient	6.8 (1.2)**	-6.9 (1.2)**	-6.4 (1.1)**	-0.3 (0.7)	0.1 (0.04)$'$	-0.3 (0.1)**	-2.4 (0.5)**
Mean Y	46.3	53.5	39	12.4	0.3	1.9	18.2
Gridcells	3,296	3,296	3,296	3,296	3,296	3,296	3,343
Birthplaces	34,846	34,846	34,846	34,846	34,846	34,846	38,893
Soldiers	1,808,425	1,808,425	1,808,425	1,808,425	1,808,425	1,808,425	4,088,306
Model				FRDD (First-stage $\mathcal{F} = 39.1$)			
Coefficient	2.9 (0.6)**	-3 (0.6)**	-4.2 (0.7)**	1.3 (0.4)**	0.1 (0.02)*	-0.1 (0.1)*	-1.7 (0.4)**
Mean Y	46.5	53.2	39.1	12	0.3	2	17.7
Gridcells	1,049	1,049	1,049	1,049	1,049	1,049	1,066
Birthplaces	12,006	12,006	12,006	12,006	12,006	12,006	13,488
Soldiers	548,765	548,765	548,765	548,765	548,765	548,765	1,242,299

Matched cluster sample. OLS, 2SLS and FRDD estimates. Outcomes on percentage scale (0 to 100). Robust standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability. 2SLS analyses exclude birth locations > 100km from railroad. FRDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.12: OBSERVED DISCHARGE RECORDS (MATCHED CLUSTER SAMPLE).

A7.3. Alternative measure of initiative

The results in Table **??** use a composite measure of intitiative, which takes the value of 1 if a soldier received at least one of four valor decorations. A potential concern with this measure is that, because 17.5% of soldiers received at least one such medal, this variable is insufficiently selective to faithfully capture battlefield initiative. By way of a robustness test, we replicated our analyses with a more selective subset of decorations, focusing on the *Order of Glory*. This medal has the distinction of being highly prestigious — just 2.3% ($N = 270,473$) of soldiers in our sample received one — but not so uncommon as to preclude credible estimation.[14] The results, in Table A7.13, are consistent with those in the main text: soldiers more exposed to repression were less likely to receive this decoration.

	OLS	2SLS	FRDD
Coefficient	-0.1 (0.01)**	-0.7 (0.1)**	-0.2 (0.05)**
Mean Y	2.4	2.4	2.3
First Stage $\mathcal{F}$		146.6	25.3

Outcome = receiving an *Order of Glory* (*Orden Slavy*) decoration of first, second or third class, measured on percentage scale (0 to 100). See the note under Table **??** for the number of soldiers, birthplaces, grid cells, and other details. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.13: REPRESSION AND ORDER OF GLORY DECORATIONS

A7.4. District-level aggregate analysis

To more directly account for local population size and urbanization, we conducted an aggregate analysis at the level of districts, which is the most fine-grained spatial unit for which 1926 Soviet census data are available ($N = 403$, including $N = 373$ within the RSRFR's 1937 borders). Our aggregate analyses adopt the same OLS and 2SLS specification as our main models (equations **??**–**??** in main text), replacing y_{ij} with $\bar{y}_j$ (average of individual outcomes for district j), and $\boldsymbol{X}_{ij}$ with $\bar{\boldsymbol{X}}_j$ (district-level averages of pre-treatment covariates). We further replaced grid-cell level fixed effects with regional (oblast-level) fixed effects, and added covariates for district population size (logged) and urbanization (percent of population residing in urban areas). Because district-level averages are more precisely estimated for areas containing more individuals, we weighted each observation by number of soldiers born in each district.

[14] By contrast, 9.3% of the soldiers in our sample ($N = 1.1M$) received the medal *For Courage*, 8.2% received *For Battle Merit* ($N = 960,734$), and 0.05% received *Hero of the Soviet Union* ($N = 5,815$). 2.5% received more than one decoration.

The results of the district-level analyses are in Table A7.14. Estimates are substantively consistent with the individual-level results in the main text.

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medals
Model				OLS			
Coefficient	1.6 (0.6)*	-0.5 (0.1)**	-0.6 (0.1)**	0.1 (0.1)	0.01 (0.002)**	0.04 (0.01)**	-0.6 (0.3)*
Mean Y	18.9	24.3	17.5	5.9	0.2	0.8	18.3
Soldiers	336	336	336	336	336	336	336
Model				2SLS (First-stage $\mathcal{F} = 348$)			
Coefficient	1.9 (0.8)*	-1 (0.2)**	-1 (0.2)**	-0.1 (0.2)	0.01 (0.004)*	0.1 (0.02)**	-0.7 (0.4)$'$
Mean Y	18.5	24	17.9	5.3	0.2	0.8	18.6
Soldiers	311	311	311	311	311	311	311

Outcomes on percentage scale (0 to 100): killed or wounded in action (KIA/WIA); missing in action (MIA), becoming prisoner of war (POW), defecting, deserting, committing treason (DDT), being punished for battlefield misconduct (PUN), or any of the above (Flee); receiving a valor decoration (Medal). Robust standard errors in parentheses. All models include regional fixed effects, and district averages of individual and birth location-level covariates. Observations weighted by number of soldiers. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.14: DISTRICT-LEVEL AGGREGATE ANALYSIS

A7.5. Estimates adjusting for unit and month fixed effects

Table A7.15 reports the full set of estimates for regressions that include fixed effects for the unit to which soldiers were assigned, and the month of the corresponding deployment.

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medal
Model				OLS			
Coefficient	0.3 (0.1)**	-0.2 (0.04)**	-0.1 (0.04)**	-0.1 (0.02)**	0.01 (0.005)′	-0.03 (0.01)**	-0.2 (0.1)**
Mean Y	59.3	22.7	18.7	2.8	0.2	1.2	18.8
Gridcells	9,647	9,647	9,647	9,647	9,647	9,647	9,647
Birthplaces	128,888	128,888	128,888	128,888	128,888	128,888	128,888
Soldiers	4,431,766	4,431,766	4,431,766	4,431,766	4,431,766	4,431,766	4,431,766
Model				2SLS (First-stage $\mathcal{F} = 136.9$)			
Coefficient	2.8 (0.4)**	-1.2 (0.2)**	-0.4 (0.2)*	-0.5 (0.1)**	0.04 (0.02)*	-0.3 (0.1)**	-2.7 (0.4)**
Mean Y	58	23.3	19	3	0.2	1.3	19.3
Gridcells	5,444	5,444	5,444	5,444	5,444	5,444	5,444
Birthplaces	106,016	106,016	106,016	106,016	106,016	106,016	106,016
Soldiers	3,716,848	3,716,848	3,716,848	3,716,848	3,716,848	3,716,848	3,716,848
Model				FRDD (First-stage $\mathcal{F} = 33.7$)			
Coefficient	1 (0.2)**	-0.2 (0.1)′	-0.3 (0.1)**	0.2 (0.1)**	0.03 (0.01)*	-0.1 (0.03)*	-0.9 (0.2)**
Mean Y	55.8	24	19.3	3.3	0.2	1.4	20.5
Gridcells	1,525	1,525	1,525	1,525	1,525	1,525	1,525
Birthplaces	31,326	31,326	31,326	31,326	31,326	31,326	31,326
Soldiers	840,783	840,783	840,783	840,783	840,783	840,783	840,783

Outcomes on percentage scale (0 to 100). Standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell, unit and month fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability. Sample includes disaggregated personnel records, with non-missing unit assignment and date information. 2SLS analyses exclude birth locations > 100km from railroad. RDD analyses exclude locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.15: ESTIMATES ADJUSTING FOR MILITARY UNIT AND MONTH.

A7.6. Estimates with alternative bandwidths

Our main analyses measure exposure to repression as the logged number of arrests within a 10 km bandwidth of a soldier's birth location. To assess the sensitivity of our results to this choice, Table A7.16 reports OLS coefficient estimates at alternative bandwidths from 1 to 20 km. For all bandwidths smaller than 10 km, estimates were consistent in sign and close in magnitude and precision to those at the 10 km baseline. For larger bandwidths, estimates remain mostly consistent in sign, but begin to attenuate and lose precision after 15 km. This attenuation pattern is not surprising, since larger bandwidths produce a

smoother map with less local variation in the repression measure.[15]

	KIA/WIA	Flee	MIA	POW	DDT	PUN	Medals
1km	0.2 (0.1)$'$	-0.3 (0.1)**	-0.3 (0.05)**	-0.1 (0.04)	0.003 (0.001)*	-0.01 (0.005)*	-0.1 (0.1)
2km	0.2 (0.1)**	-0.4 (0.1)**	-0.3 (0.03)**	-0.1 (0.04)$'$	0.002 (0.001)$'$	-0.01 (0.005)*	-0.1 (0.1)
3km	0.1 (0.1)*	-0.4 (0.1)**	-0.3 (0.03)**	-0.1 (0.04)*	0.002 (0.001)	-0.02 (0.005)**	-0.1 (0.1)
4km	0.2 (0.1)**	-0.4 (0.1)**	-0.3 (0.03)**	-0.1 (0.04)	0.002 (0.001)	-0.02 (0.005)**	-0.1 (0.1)
5km	0.2 (0.1)**	-0.3 (0.05)**	-0.3 (0.03)**	-0.1 (0.03)$'$	0.003 (0.001)$'$	-0.01 (0.003)**	-0.1 (0.1)*
6km	0.2 (0.1)*	-0.3 (0.05)**	-0.3 (0.03)**	-0.1 (0.04)$'$	0.002 (0.001)	-0.01 (0.003)**	-0.1 (0.1)$'$
7km	0.2 (0.1)*	-0.3 (0.1)**	-0.3 (0.03)**	-0.1 (0.04)	0.003 (0.002)$'$	-0.01 (0.004)**	-0.1 (0.1)$'$
8km	0.3 (0.1)**	-0.3 (0.1)**	-0.3 (0.04)**	-0.03 (0.04)	0.002 (0.002)	-0.01 (0.004)**	-0.2 (0.05)**
9km	0.3 (0.1)**	-0.3 (0.1)**	-0.2 (0.04)**	-0.04 (0.03)	0.003 (0.002)*	-0.02 (0.004)**	-0.2 (0.1)**
10km	0.4 (0.1)**	-0.2 (0.1)**	-0.2 (0.04)**	-0.03 (0.04)	0.003 (0.002)	-0.01 (0.005)**	-0.2 (0.1)**
11km	0.4 (0.1)**	-0.2 (0.1)**	-0.2 (0.04)**	-0.01 (0.04)	0.003 (0.002)	-0.01 (0.005)*	-0.2 (0.1)**
12km	0.3 (0.1)**	-0.2 (0.1)**	-0.1 (0.04)**	-0.01 (0.04)	0.002 (0.002)	-0.01 (0.005)*	-0.1 (0.1)*
13km	0.4 (0.2)*	-0.2 (0.1)*	-0.1 (0.1)$'$	-0.04 (0.04)	0.004 (0.002)$'$	-0.01 (0.01)	-0.2 (0.1)*
14km	0.4 (0.2)*	-0.1 (0.1)	-0.1 (0.1)	-0.04 (0.05)	0.003 (0.002)	-0.01 (0.01)	-0.2 (0.1)*
15km	0.3 (0.1)**	-0.1 (0.1)	-0.04 (0.1)	-0.02 (0.1)	0.0007 (0.002)	-0.01 (0.01)$'$	-0.2 (0.1)*
16km	0.2 (0.1)$'$	-0.03 (0.1)	0.02 (0.1)	-0.03 (0.1)	-0.002 (0.003)	-0.01 (0.01)	-0.1 (0.1)
17km	0.3 (0.1)*	-0.03 (0.1)	0.04 (0.1)	-0.1 (0.1)	-0.003 (0.003)	-0.01 (0.01)$'$	-0.1 (0.1)
18km	0.2 (0.1)$'$	0.01 (0.1)	0.1 (0.1)$'$	-0.1 (0.1)	-0.004 (0.003)	-0.01 (0.01)	-0.1 (0.1)
19km	0.2 (0.1)	-0.1 (0.1)	0.1 (0.1)	-0.1 (0.1)$'$	-0.004 (0.003)	-0.01 (0.01)	-0.04 (0.1)
20km	0.1 (0.1)	0.01 (0.1)	0.2 (0.1)*	-0.1 (0.1)$'$	-0.003 (0.003)	-0.01 (0.01)	-0.004 (0.1)

See the notes under Tables **??-??** for details. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.16: EFFECT OF REPRESSION AT ALTERNATIVE BANDWIDTHS (1–20 KM).

A7.7. Local effect heterogeneity

As noted in the main text, the magnitude of our OLS coefficient estimates is consistently smaller than their 2SLS/FRDD counterparts. Potential explanations for these differences include measurement error (e.g. underestimation of arrests in places with generally higher K/WIA rates), and local effect heterogeneity (i.e. the impact of repression was stronger in areas closer to railways and regional borders). While it is difficult to quantify the influence of measurement error on estimation in this case, we are able to rule out at least one source of local effect heterogeneity. The locality of 2SLS/FRDD effect estimates is driven by a combination of sample selection (we restricted 2SLS analyses to locations

[15] As bandwidths become so large that individuals born in the same grid cell have nearly-identical numbers of arrests, virtually all variation in repression becomes cross-grid cell (captured by fixed effects) rather than within grid cells (captured by the repression exposure measure). Larger bandwidths therefore necessitate changes to model specification, with fixed effects for grid cells of larger size.

within 100 km of railroads, and FRDD to locations within 50 km of regional borders) and differences in repression's effect on compliers versus the general population. In Table A7.17, we reestimate our OLS fixed effect models on the subsets of data used to fit our 2SLS and FRDD models. While these subset analyses do not directly address the compliance issue, they clearly show that sample selection cannot explain the differences in magnitude. Effect estimates are nearly identical in magnitude across the three samples, and do not approach anything resembling the almost tenfold differences we see between some of the OLS and 2SLS/FRDD coefficients.

	KIA/WIA	Flee	MIA	POW	DDT	PUN
Model	Full sample					
Coefficient	0.4 (0.1)**	-0.2 (0.1)**	-0.2 (0.04)**	-0.03 (0.04)	0.003 (0.002)	-0.01 (0.005)**
Mean Y	21.2	24	17.4	5.8	0.2	0.8
Gridcells	12,143	12,143	12,143	12,143	12,143	12,143
Birthplaces	182,225	182,225	182,225	182,225	182,225	182,225
Soldiers	11,427,895	11,427,895	11,427,895	11,427,895	11,427,895	11,427,895
Model	2SLS sample					
Coefficient	0.4 (0.1)**	-0.2 (0.1)**	-0.2 (0.05)**	-0.01 (0.03)	0.002 (0.002)	-0.02 (0.005)**
Mean Y	20.5	24.4	17.6	5.9	0.2	0.8
Gridcells	5,654	5,654	5,654	5,654	5,654	5,654
Birthplaces	147,803	147,803	147,803	147,803	147,803	147,803
Soldiers	9,725,085	9,725,085	9,725,085	9,725,085	9,725,085	9,725,085
Model	FRDD sample					
Coefficient	0.4 (0.1)**	-0.2 (0.1)*	-0.3 (0.1)**	0.02 (0.1)	0.005 (0.004)	-0.02 (0.01)
Mean Y	19	23.9	17.6	5.3	0.2	0.9
Gridcells	1,582	1,582	1,582	1,582	1,582	1,582
Birthplaces	44,154	44,154	44,154	44,154	44,154	44,154
Soldiers	2,316,908	2,316,908	2,316,908	2,316,908	2,316,908	2,316,908

Outcomes on percentage scale (0 to 100): killed or wounded in action (KIA/WIA); missing in action (MIA), becoming prisoner of war (POW), defecting, deserting, committing treason (DDT), being punished for battlefield misconduct (PUN), or any of the above (Flee); receiving a personal valor decoration (Medal). Standard errors in parentheses, clustered by birth location and grid cell. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability. 2SLS sample excludes birth locations > 100km from railroad. FRDD sample excludes locations in non-matched regions and > 50km from regional borders. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table A7.17: OLS ANALYSES ON RESTRICTED SAMPLES.

A7.8. Sensitivity analyses of the 2SLS exclusion restriction

A key identifying assumption of our instrumental variable analyses is the exclusion restriction, which requires that our instrument (distance to nearest railroad) influence individual battlefield outcomes only through its effect on treatment (arrests). An especially concerning violation of this assumption would be one where – for some unobserved socio-economic, cultural or other reason – people living near railroads in 1937 were systematically more likely to die in battle, less likely to surrender or flee, and less likely to receive decorations. We now conduct an additional set of analyses to assess how severe possible violations of the exclusion restriction would need to be in order to overturn our 2SLS results. Following Conley, Hansen and Rossi (2012), we model these potential violations with an extension of our main two-stage specification,

$$y_i = \zeta \cdot \hat{f}\left(\text{Raildist}_j\right) + \gamma \cdot \ln\left(\widehat{\text{Repression}}_{j[i]} + 1\right) + \boldsymbol{\beta}'\boldsymbol{X}_{ij} + \text{Cell}_{k[i]} + s\left(\text{lon}_{j[i]}, \text{lat}_{j[i]}\right) + \epsilon_i$$
$$\ln\left(\text{Repression}_j + 1\right) = \alpha \cdot \hat{f}\left(\text{Raildist}_j\right) + \boldsymbol{\lambda}'\boldsymbol{X}_j + \text{Cell}_{k[j]} + s\left(\text{lon}_{j[i]}, \text{lat}_{j[i]}\right) + \epsilon_j$$

where $\hat{f}\left(\text{Raildist}_j\right)$ is the excluded (linearized) instrument, and ζ is a parameter capturing the size and direction of exclusion restriction violations. If there are no violations, $\zeta \equiv 0$.

Our sensitivity analysis employs Conley, Hansen and Rossi (2012)'s union of confidence intervals approach, which estimates the maximum value ζ can take such that the γ coefficient estimate remains statistically significant at the 95% level. Given a support region for ζ, $\mathfrak{Z}$, we draw a value $\zeta_0 \in \mathfrak{Z}$ and subtract $\zeta_0 \cdot \hat{f}\left(\text{Raildist}_j\right)$ from both sides of the second-stage equation:

$$\left(y_i - \zeta_0 \cdot \hat{f}\left(\text{Raildist}_j\right)\right) = \gamma \cdot \ln\left(\widehat{\text{Repression}}_{j[i]} + 1\right) + \boldsymbol{\beta}'\boldsymbol{X}_{ij} + \text{Cell}_{k[i]} + s\left(\text{lon}_{j[i]}, \text{lat}_{j[i]}\right) + \epsilon_i$$

We then employ the usual asymptotic approximations to obtain a 95% confidence interval for $\hat{\gamma}$, assuming that $\zeta = \zeta_0$. We construct these intervals for all points in $\mathfrak{Z} = [-5, 5]$.

The sign of ζ determines whether violations of the exclusion restriction are more likely to attenuate or inflate estimates of γ. By construction, $\hat{f}\left(\text{Raildist}_j\right)$ is increasing in proximity to railroads (i.e. larger values indicate that a location is closer to the railway). Exclusion restriction violations are therefore more likely to attenuate $\hat{\gamma}$ if $\zeta > 0$ for KIA/WIA (meaning that individuals born closer to the railroad are more likely to die or become wounded), $\zeta < 0$ for MIA/POW/DDT/Punished and for medals (implying that those born closer to railroads are less likely to have these outcomes). If ζ takes the opposite signs, then standard 2SLS regression underestimates the true effect of repression.

Table A7.18: 2SLS SENSITIVITY ANALYSES.

Outcome	$\max(\zeta)$	$\hat{\gamma}$ at $\max(\zeta)$	95% CI	$\zeta \cdot \hat{f}\left(\frac{sd(Z)}{2}\right)$	Mean Y
KIA/WIA	1.7	1.004	(0.001, 2.006)	2.97	21.476
Flee	-1.174	-0.744	(-1.489, -0.0003)	-2.05	24.375
MIA	-1.065	-0.556	(-1.111, -0.002)	-1.86	17.712
DDT	0.008	0.02	(0.001, 0.039)	0.014	0.164
PUN	-0.012	-0.049	(-0.096, -0.001)	-0.021	0.817
Medal	-1.035	-0.7	(-1.4, -0.001)	-1.809	17.651

2SLS estimates of repression's effect (includes only results significant at 95% level in main analysis). Outcomes on percentage scale (0 to 100). Confidence intervals based on robust standard errors, clustered by birth location. All models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability.

Table A7.18 reports the results of these sensitivity analyses, including the maximum size ζ can take while maintaining a significant estimate of γ, along with the corresponding γ estimate and its 95% confidence region. Note that the table includes only those results, which we originally found to be significant at the 95% level in the main analyses. We also report the implied effect that a median-to-zero decrease in distance to railroad (38km to 0km) would have on y at each critical value of ζ.

In the case of KIA/WIA, for example, the critical value of ζ is 1.7. In order to overturn the positive effect of repression on this outcome, a median-to-zero decrease in distance from the railroad would need to increase one's chances of dying or becoming wounded by at least $1.7 \cdot \hat{f}(-38 \text{ km}) = 3$ percent. The magnitude of this violation would therefore need to be quite substantial, considering that the mean value of KIA/WIA is 21.5 percent. These results suggest that – for most battlefield outcomes, and especially KIA/WIA, MIA and Glory Medals – the effect of repression is robust to reasonably-sized violations of the exclusion restriction. Other results, such as the odd positive coefficient for DDT, appear to be highly sensitive to these violations, with $\zeta < .01$.

References

Bolin, Robert, L. 1946. "Handbook on USSR Military Forces: Chapter II–Personnel Administration.".

Chang, Charles and Michael Masterson. 2020. "Using Word Order in Political Text Classification with Long Short-term Memory Models." *Political Analysis* 28(3):395–411.

Chollet, François. 2015. "keras.".
URL: *https://github.com/fchollet/keras*

Clark, Philip J and Francis C Evans. 1954. "Distance to nearest neighbor as a measure of spatial relationships in populations." *Ecology* 35(4):445–453.

Conley, Timothy G, Christian B Hansen and Peter E Rossi. 2012. "Plausibly exogenous." *Review of Economics and Statistics* 94(1):260–272.

Cressie, Noel and Timothy RC Read. 1984. "Multinomial goodness-of-fit tests." *Journal of the Royal Statistical Society: Series B (Methodological)* 46(3):440–464.

Empric, E. Bruce. 2017. *Onward to Berlin! Red Army Valor in World War II - The Full Cavaliers of the Soviet Order of Glory*. Teufelsberg Press.

Enamorado, Ted, Benjamin Fifield and Kosuke Imai. 2019. "Using a probabilistic model to assist merging of large-scale administrative records." *American Political Science Review* 113(2):353–371.

Fellegi, Ivan P and Alan B Sunter. 1969. "A theory for record linkage." *Journal of the American Statistical Association* 64(328):1183–1210.

Graves, Alex and Jürgen Schmidhuber. 2005. "Framewise phoneme classification with bidirectional LSTM and other neural network architectures." *Neural networks* 18(5-6):602–610.

Green, Donald P and Lynn Vavreck. 2008. "Analysis of cluster-randomized experiments: A comparison of alternative estimation approaches." *Political Analysis* pp. 138–152.

Greff, K., R. K. Srivastava, J. Koutník, B. R. Steunebrink and J. Schmidhuber. 2017. "LSTM: A Search Space Odyssey." *IEEE Transactions on Neural Networks and Learning Systems* 28(10):2222–2232.

Higgins, Michael J, Fredrik Sävje and Jasjeet S Sekhon. 2016. "Improving massive experiments with threshold blocking." *Proceedings of the National Academy of Sciences* 113(27):7369–7376.

Ho, Daniel E., Kosuke Imai, Gary King and Elizabeth A. Stuart. 2007. "Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference." *Political Analysis* 15(3):199–236.

Hochreiter, Sepp and Jürgen Schmidhuber. 1997. "Long short-term memory." *Neural computation* 9(8):1735–1780.

Imai, Kosuke, Gary King, Clayton Nall et al. 2009. "The essential role of pair matching in cluster-randomized experiments, with application to the Mexican universal health insurance evaluation." *Statistical Science* 24(1):29–53.

Kulldorff, Martin. 1997. "A spatial scan statistic." *Communications in Statistics-Theory and methods* 26(6):1481–1496.

Leicht, Elizabeth A and Mark EJ Newman. 2008. "Community structure in directed networks." *Physical review letters* 100(11):118703.

Murtagh, Fionn and Pierre Legendre. 2014. "Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion?" *Journal of classification* 31(3):274–295.