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

We utilize over 100 million declassified Red Army personnel records from World War II to study how state repression shapes soldiers' motivation to exert effort in fighting. Exploiting multiple complementary identification strategies, we find that soldiers from places with higher levels of pre-war repression under Stalin's rule were more likely to fight until death and less likely to shirk their duties, but they also received fewer decorations for personal bravery. The coercive incentives created by repression appear to have induced obedience at the expense of initiative and increased the human costs of war.

JEL Classification: D74, F51, H56, N44

December 2022

The development and survival of states often hinges on their ability to extract resources for war-making from their populations (Tilly, 1985). 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? Because the conduct of individuals in war has far-reaching economic and political consequences (Acemoglu, Autor and Lyle, 2004; Scheidel, 2018), a deeper understanding of combat motivation at the micro-level is essential.

Existing research has examined the role of pecuniary incentives (Grossman, 1991; Hall, Huff and Kuriwaki, 2019), ideology (Barber IV and Miller, 2019), group loyalties (Shils and Janowitz, 1948; Costa and Kahn, 2003), status competition (Ager, Bursztyn and Voth, 2022), fear of punishment (Chen, 2017), and discrimination (Lyall, 2020) in motivating individuals to fight. We propose a complementary perspective, which underscores soldiers' prior interactions with the state outside their line of duty. If we accept the long-held premise that military institutions do not evolve in isolation from broader society and politics (Clausewitz, 1832/1984, 592-593), then we must also acknowledge the importance of lived experiences vis-a-vis the state that soldiers had before their service. Soldiers for whom these experiences were mostly positive may approach their duties differently than ones who have come to see their state as unjust or tyrannical.

We conduct a quantitative individual-level study of the military effort of the Soviet Union during World War II, which is arguably the most paradigmatic case for the question at hand. Within the span of a few years, the Soviet state headed by Josef Stalin went from inflicting mass terror 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 European casualties in the war and 18 of the 25 costliest battles on record (Overy, 1998, xvi). The Soviet Union lost over 11.2 million military personnel and 17.9 million civilians (Surinov and Oksenoyt, 2015). Almost 40% of the battlefield losses comprised soldiers who surrendered, deserted, or went missing (Krivosheev, 1997). 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 (Edele, 2017; Thurston, 2000). Incentives to avoid fighting were compelling: in the battle of Stalingrad, for example, the average life expectancy was 24 hours for enlisted Soviet soldiers and three days for officers (Merridale, 2006). Given these odds, it is remarkable that the Red Army managed to keep millions of its troops in line fighting while others fled (Reese, 2011).

The study's goal is to explain how prewar mass violence by the state impacts the combat motivation of soldiers during war. To this end, we collected detailed data on the Red Army in World War II. Using 106 million declassified personnel records, we reconstruct the wartime trajectories of almost 12 million soldiers. We link these wartime records with micro-level data from over 2 million secret police case files on state repression prior to the war. We then evaluate whether exposure to prewar repression could in part explain why some soldiers fought to the death while others surrendered, deserted, or went missing, and why some received decorations for valor in battle but others did not.

To validate our measures of soldier-level outcomes, we show that these outcomes map consistently onto the aggregate performance of the units in which those soldiers served. The Red Army was more likely to gain territory in battles where a higher share of its soldiers was killed or wounded, a lower share went missing, surrendered, deserted, or showed insubordination, and a higher share received awards for bravery. Because territorial gains are standard measures of aggregate combat performance in ground warfare (Biddle, 2004), these correlations suggest that we can plausibly interpret a soldier's death in battle or award for bravery as indicators of high combat resolve, and going missing, deserting, surrendering, or insubordination as indicators of low combat resolve.

An obvious limitation of this study is that repression was not completely random. We exploit several features of Stalin's repressive apparatus that help with causal identification. First, mass terror targeted particular groups of people based on their ethnicity, class, and region of residence. However, there was little differentiation between the individuals or communities within those demographic categories. We compare the outcomes of soldiers who originate from geographically proximate communities with differential levels of repression, while accounting for the same observables that the regime used when selecting group-level targets. Our identifying assumption here is that such controlled comparisons remove or at least minimize the "systematic component" in the allocation of repression.

Our second empirical approach exploits the fact that local administrators had wide discretion in implementing repression. We utilize discontinuous changes in arrest levels across administrative borders as a plausibly exogenous source of variation in repression. Because the Soviets redrew administrative borders just before the Great Terror campaign, a common problem with geographic regression discontinuity designs — that many things besides treatment vary across borders — is less formidable in our context. Balance tests indicate that communities residing near administrative borders were identical across many socio-economic characteristics and differed only in their exposure to repression, likely due to idiosyncratic differences in the local coercive apparatus.

Across both empirical approaches, we find that soldiers from places with high levels of prewar repression had systematically different battlefield outcomes compared to others: they were more likely to be killed or wounded in action; they were less likely to go missing, desert, surrender or defy orders; yet they also showed fewer personal acts of bravery as far as we can judge from their award records. The results remain consistent across a wide range of robustness checks, including a matched cluster design that compares soldiers from birth locations with similar socio-demographic characteristics, including estimated population size; regressions with aggregate data at the level of soldiers' birthplaces or home districts; analyses that adjust for the interdependence of soldiers' decisions within units; an instrumental variable design that exploits the logistical constraints faced by the state in reaching and transporting victims; and an expansion of the sample to include soldiers born in Soviet Ukraine (for which the data have lower quality).

We rationalize these empirical patterns using a stylized model of soldiers' decision-making process, which focuses on trade-offs between intrinsic and extrinsic motivation (Bernheim, 1994; Benabou and Tirole, 2003). Soldiers who learn firsthand about the repressive nature of their state might become intrinsically less willing to fight for it. At the same time, repression might increase soldiers' extrinsic incentive to fight, as exposed individuals come to expect that they — or even their families — may be punished for the slightest hint of disloyalty. Due to these countervailing forces, repression leads to "perfunctory" rather than "consummate" compliance (Brehm and Gates, 1999), where soldiers are willing to obey orders but are hesitant to take personal initiative. The model predicts that repression increases combat resolve at the low end of the distribution (gains in compliance), but reduces resolve at the high end of the distribution (loss of initiative).

To conduct a more direct test of this distributional effect of repression, we construct an index of combat resolve using aggregated data on Red Army units. This index is the predicted probability that a unit participated in a battle that resulted in a territorial gain, conditional on average soldier-level outcomes in that unit. Consistent with our theoretical logic, we show that the distribution of this index has a higher mean and lower variance in units whose soldiers came from locations with higher average levels of repression.

An alternative interpretation of our results is that they reflect wartime discrimination against soldiers from heavily-repressed locations, rather than the choices soldiers made in battle. These soldiers may have died at higher rates because they were selectively assigned to more dangerous parts of the front, or to units with worse leadership and equipment.

Officials may also have denied these soldiers awards for bravery, regardless of merit.

While we cannot conclusively rule out these possibilities, there are several important patterns in our data that discrimination cannot fully explain. For example, we find no evidence that soldiers from heavily-repressed areas were overlooked for promotions, or were more likely to be assigned to units with more exposure to enemy fire. We find the same systematic differences when we compare soldiers serving concurrently in the same units, who were exposed to similar battlefield conditions and unit leadership, and had similar kits. We also estimate regressions with time-varying effects and find that the variation in estimates is consistent with our proposed interpretation.

The tension between pre-war state violence and wartime combat mobilization has surfaced in many conflicts, including the Iran-Iraq War (Pollack, 2004, 182), the Second Congo War (Lyall, 2020, 332), and the civil war in Syria (Heydemann, 2013). 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. 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 (Merridale, 2006); the Red Army offered no material inducements for combat performance.[1]

Our study contributes to several strands of literature in empirical political economy. First is research on the micro-foundations of military effort. We show that soldiers' prior experiences of repression can undermine their intrinsic motivations to fight for material, ideological, or other reasons (Barber IV and Miller, 2019; Grossman, 1991; Hall, Huff and Kuriwaki, 2019). By focusing on repression prior to war, our study complements recent scholarship on the coercive incentives that soldiers face during war (Chen, 2017; Lyall, 2020). Our results suggest that soldiers who had experienced state violence more intimately as civilians may be more responsive to coercive measures on the battlefield.

More broadly, this article extends research on how states' exploitative and coercive practices impact development (Dell, 2010), social structure (Acemoglu, Hassan and Robin-

[1]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 only in the 1970s (Edele, 2006).

son, 2011), trust (Grosjean, 2014), behavior (Rozenas and Zhukov, 2019), and identity (Blaydes, 2018). Our findings highlight a previously overlooked negative externality from repression on states' ability to provide the most basic public good: national security.

1. Repression and Combat Resolve

Military institutions do not exist in a social vacuum. Whether they are conscripts or professionals, soldiers bring a variety of life experiences from the civilian domain to the battlefield. Depending on their prior interactions with the state, soldiers may differ in their intrinsic willingness to fight in the name of their state. They may also have variable expectations about how the state will react if they shirk or defect in the line of duty.

How does differential exposure to state violence in civilian life impact the resolve of soldiers in combat? Existing research does not provide a clear answer. There is systematic evidence that violence provokes negative sentiments toward the perpetrator (Blaydes, 2018; Dell and Querubin, 2018; Grosjean, 2014; Rozenas and Zhukov, 2019). In the military context, ethnic groups that face discrimination in the civilian domain are less willing to fight for the state (Lyall, 2020) and more willing to rebel against it (Cederman, Wimmer and Min, 2010). This implies that soldiers who are more intimately aware that they are fighting for a repressive, unjust state should exhibit lower combat resolve. We refer to this as the alienating effect of repression.

The intrinsic willingness to fight is not the only driver of a soldier's behavior. Like any organization, the military relies on extrinsic, sometimes coercive incentives to overcome low intrinsic motivation among its personnel. The costs of shirking — retreating or hiding when ordered to charge — can include demotion, court-martial (Chen, 2017), execution (Statiev, 2010), even retaliation against family members (Reese, 2011). As Beissinger (2002, 326) writes, repression works by creating "internalized expectations about [how] authority will respond punitively to challenging acts." A soldier exposed to repression in civilian life may see the state's threat of punishment on the battlefield as more credible and, therefore, may exhibit higher resolve than they would prefer based solely on their intrinsic motivation. This is the deterrent effect of repression.

Because the two effects of repression are in competition with each other, it is not obvious how their interaction should ultimately impact the behavior of soldiers on the battlefield. To untangle this problem, we employ a stylized model of soldiers' choices.

1.1. A Stylized Model

There is a continuum of soldiers indexed by type $\omega \in \mathbb{R}$, representing their *intrinsic* combat motivation, drawn from the distribution F.[2] Each soldier chooses an action $a \in \mathbb{R}$, which denotes the level of *combat effort* they exert on the battlefield at the risk of injury or death. For concreteness, let $\overline{a}$ denote the action ordered by commanders (e.g., "charge!") and let $\underline{a}$ denote a cutoff such that if $a < \underline{a}$, the soldier shows especially low resolve that can be considered shirking (i.e., minimizing the risk of death or injury by deserting, surrendering, or going missing). As a soldier's combat effort approaches $\overline{a}$, they are showing increasing compliance. If the combat effort exceeds $\overline{a}$, the soldier is taking personal initiative by going above and beyond what they were ordered to do: being the first to storm enemy positions, personally capturing an enemy officer, assisting the wounded under enemy fire, or continuing to carry out one's mission in a burning tank.

To model the alienating effect of repression, suppose that the intrinsic motivation to fight for a soldier exposed to repression level $r \geq 0$ is $\omega - \alpha r$, where α is the "alienation" parameter. Action a results in an intrinsic loss of $(a - (\omega - \alpha r))^2$ for the soldier. Absent other considerations, a soldier would minimize the loss by choosing $a = \omega - \alpha r$.

To model the deterrent effect, suppose that a soldier who shows *less* resolve than required by commanders suffers an extrinsic loss of $\mathbb{E}(\pi|r)(\overline{a} - a)^2$, where $\mathbb{E}(\pi|r)$ is the penalty that a soldier expects to receive for under-performance. Let $\mathbb{E}(\pi|r) = \delta r$, where $\delta \geq 0$ is the "deterrence" parameter.[3] A soldier exposed to more repression will infer that the state's ability to punish is higher and will be more reluctant to shirk.

Each soldier chooses an optimal level of combat resolve a^* to minimize the sum of their 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

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

[2] F has full support, is increasing everywhere, and has a density f.

[3] The setup could be micro-founded with a signaling model, where r depends on the state's type ("repressive capacity").

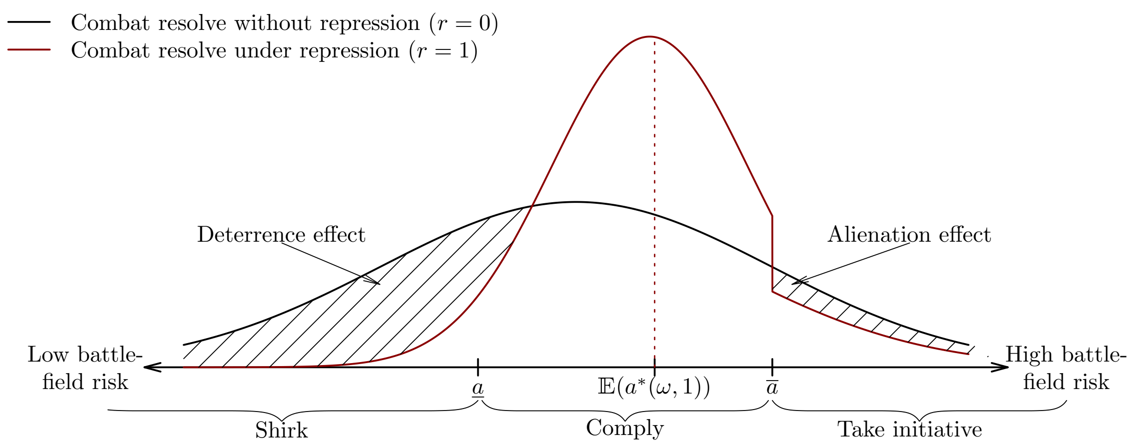


Figure 1: How repression shapes combat resolve

The optimal combat resolve a^* is increasing in repression r for soldiers with low intrinsic motivation, $\omega \leq \overline{a} - \alpha/\delta$, and decreasing for soldiers with high intrinsic motivation $\omega > \overline{a} - \alpha/\delta$. In the former case, the deterrent effect of repression dominates the alienating effect (pushing a^* up); in the latter, the alienating effect dominates the deterrent effect (pulling a^* down). The following proposition specifies how these two countervailing effects of repression alter the overall distribution of combat resolve.

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

Figure 1 illustrates the logic behind this proposition (see Appendix A1 for the proof). The black curve is the density of combat resolve without repression where $a^* = \omega$. The red curve is the density of combat resolve under repression ($r = 1$). The density under repression has a higher mean $\mathbb{E}(a^*(\omega, r))$, but it also has lower variance as the tails of the distribution are squeezed inwards. Repression increases the extrinsic incentive to obey orders (soldiers with low intrinsic motivation ω show more resolve than they would otherwise), but it saps the intrinsic motivation to take initiative beyond the formal mandate (soldiers with high intrinsic motivation ω show less resolve than they would otherwise).

1.2. Observable Implications

We cannot observe soldiers' intrinsic motivation (ω) or combat resolve (a^*), but only discrete outcomes like whether a soldier died or survived, received a medal, deserted, surrendered, or shirked their duties in some other manner. To test our proposition, we need to map these observable outcomes, in a theoretically plausible way, to latent combat resolve. In doing so, we follow the precedents set in earlier literature.

To measure low combat resolve ("shirking"), we use desertion, surrender, absence without official leave, and disciplinary transgressions — following the approach by Costa and Kahn (2003) in their study of the U.S. Civil War, and cross-national research on conventional warfare (Lehmann and Zhukov, 2019; Lyall, 2020). In our model, an observable instance of such actions could indicate that the soldier had shown low combat resolve ($a^* < \underline{a}$). To measure high combat resolve ("initiative"), we follow Barber IV and Miller (2019) and use medals for valor. Since valor decorations recognize acts of courage beyond what is normally expected of soldiers, a reasonable inference is that a soldier who received such an award showed combat resolve above what their orders dictated ($a^* > \overline{a}$).

Finally, how should we interpret, in our theoretical framework, a soldier's death on the battlefield? Although a soldier's survival depends on many factors beyond their control (i.e., battle intensity, competence of commanders, peer behavior), extant literature often uses battle deaths to assess actors' resolve in fighting wars. Because a higher level of effort typically carries a greater risk of physical harm, commitment to one's combat mission implies a tacit willingness to sacrifice oneself for the cause. This is true not only for suicide terrorism, where armed groups routinely screen for "reliable martyrs" who are unlikely to defect (Berman and Laitin, 2008), but in conventional war as well. Ager, Bursztyn and Voth (2022), for example, use the deaths of German pilots in World War II as a proxy for effort and/or risk-taking. Between two soldiers exposed to similar battlefield conditions while serving at comparable ranks, we assume that the soldier who died fighting likely showed higher resolve a^* than the one who survived — perhaps not as high as a soldier who received a decoration for valor, but certainly higher than soldiers who fled the battle.

Consistent with how we interpret these outcomes, a validation analysis in Section 3.2 shows that in battles where more soldiers died, fewer shirked, and more received awards for valor, the Red Army tended to gain territory rather than lose it.

Three testable predictions follow from Proposition 1:

Predictions. *As soldiers' exposure to state repression increases, they become (1) more likely to be killed in action (higher average a^*), (2) less likely to shirk by deserting, surrendering, going missing, or being punished for disciplinary transgressions (less likely that $a^* < \underline{a}$), and (3) less likely to receive an award for valor (less likely that $a^* > \overline{a}$).*

The observable implications of our theoretical logic are distinct from some alternative explanations. One could argue that exposure to repression motivates soldiers to "signal their loyalty" by showing consistently higher resolve and initiative. But if that is true, then repression should lead to more deaths, less shirking, and *more* awards for valor. Alternatively, if repression had only an alienating effect, we should see *fewer* deaths, more shirking, and fewer medals among soldiers exposed to repression.

The most challenging alternative theory is that repression drives soldier-level outcomes not through soldiers' choices but through the discriminatory practices of the state. Soldiers more exposed to repression could die in larger numbers because authorities assign them to more dangerous posts, and they could be denied awards even if they deserve them. We conduct an array of empirical tests to assess the plausibility of this argument.

2. Soviet Repression and War Effort

Under Stalin's rule (1927-1953), the Soviet state officially convicted 3.8 million people for "counter-revolutionary" crimes, most in the 1930's.[4] The stated goal of this violence was to eliminate "anti-Soviet elements," but the regime had few means to identify who engaged in "counter-revolutionary" activities or held "anti-Soviet" views. Security officials targeted broadly-defined segments of the population, like residents of particular provinces or minorities, without discriminating between individuals within 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" (Gregory, 2009, 196). Eventually, people of all backgrounds, including party officials, the military, and security agencies, "everybody from the Politburo member down to the street cleaner," became potential victims of state violence (Ulam, 1973).

Central authorities provided little concrete guidance as to who should be repressed. The NKVD's Main Directorate for State Security issued numerical quotas of persons to be executed or sent to camps in each region and "everything else depended on the ingenuity of Security operations personnel" (Solzhenitsyn, 1973, 69-70). Local executives often engaged in "exceptional competition" to exceed their quotas and signal administrative zeal (Chukhin, 1999, 76). 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 (Jansen and Petrov, 2002, 86, 88).

[4] State Archive of Russian Federation (GARF), collection 9401, series 2, case 450, pp. 30-37.

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" (Conquest, 2008, 434).

After the German invasion on June 22, 1941, the Soviet Union mobilized all military-age males – over 30 million civilians throughout the war – to support its 4.5 million-strong standing Red Army. These ordinary citizens were the backbone of the Soviet defense against the Germans. The war became an "acid test" for Stalinism: would the people risk their lives for a regime that only recently had terrorized them (Thurston, 2000)?

In some respects, Stalinism passed the test: millions of soldiers fought for the Soviet state, often to death. Early on, the Red Army stumbled spectacularly due to prewar officer purges, politicized decision-making, and chronic mismanagement (Glantz, 2005). Ultimately, the Soviet Union won the war, and it did so largely by keeping its troops fighting *despite* organizational malaise and human loss (Reese, 2011). Yet many soldiers voted against Stalinism with their feet. This was the "first war fought by Russia in which a large force of its citizens joined the other side" (Conquest, 2008, 456). Half of all personnel losses in the war's first year comprised soldiers missing in action or captured. Thousands were detained for desertion, sabotage, or treason. While there were many reasons to shirk, widespread distaste for how the state treated its citizens did not help (Edele, 2017).

Moscow took draconian measures to hold its troops in line. On August 16, 1941, the Headquarters of the Supreme High Command issued Order 270, stipulating that those "who surrender to the enemy shall be considered malicious deserters, whose families could be arrested" (Zolotarev, 1997, 58-60). Commanders were to prepare bi-weekly reports for the General Staff, listing captured soldiers and their families' addresses (Kachuk, 2013). Among the victims of this order was Stalin's own son, Yakov, whose wife was sent to a labor camp after his capture by the Germans. People's Commissar of Defense Order 227 (July 28, 1942) went further, requiring every front to organize "penal units" staffed by men accused of indiscipline, and send them to the most dangerous sectors to "atone for their crimes against the Motherland with their blood" (Statiev, 2010, 726). The order also mandated the creation of "blocking units" to detain or execute retreating personnel.

Given these punitive measures, soldiers who fought instead of fleeing might have done so because they were intrinsically committed to the cause or because they feared punishment. Soldiers who witnessed state repression prior to war may have been especially sensitive to these 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.

3. Data and Measures

3.1. Military records

We extracted military records from a database maintained by the Russian Ministry of Defense, People's Memory (*Pamyat' Naroda*), which contains 106 million declassified Red Army personnel records. This 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 the soldier's discharge.

Due to illegible handwriting, digitization inaccuracies, 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 military ranks, unit names and numbers, assigning tactical units to parent divisions, corps, and armies, and standardizing and validating geographic names. Since the records do not contain a field like "social security number" that uniquely identifies soldiers across different sets of documents, we used a probabilistic record linkage approach (Enamorado, Fifield and Imai, 2019) which we tailored to be operable with our data (see Appendix A2).

Because we measure soldiers' exposure to repression based on their birth locations (see below), we excluded soldiers whose birthplaces were missing or could not be geocoded to the municipality level or lower. In our main analysis, we only include soldiers born within Soviet Russia (RSFSR), because data on arrest records — from which we construct our main independent variable — are sparse and less reliable outside RSFR.[5] Our final dataset contains 26,542,786 records for 11,606,552 soldiers.

[5]The data on arrest records come from a Russia-based NGO, *Memorial*, and it naturally has better coverage for Russia compared to other former Soviet republics. In Appendix A7.10, we replicate our analysis with an expanded sample, which includes soldiers born in Soviet Ukraine in pre-1939 borders.

3.2. Measuring combat resolve

Before the war ended, 46% of soldiers were discharged from the Red Army. Discharge categories included: killed or wounded in action (KIA/WIA), missing in action (MIA), prisoner of war (POW), deserted, defected or committed treason (DDT), or punished for misconduct (PUN). 17% of soldiers received medals for valor. As noted earlier, we use KIA/WIA as a proxy for compliance with orders, in the sense that a soldier fought until they exited the sample by being physically incapacitated (cf. Ager, Bursztyn and Voth, 2022).[6] We treat MIA, POW, DDT, or PUN as indicators of low combat resolve (cf. Costa and Kahn, 2003), and medals as indicators of high resolve (cf. Barber IV and Miller, 2019).

There are important caveats to consider when inferring soldiers' resolve from official wartime records. While it is quite plausible that soldiers who defected, deserted, committed treason (DDT) or were punished for misconduct (PUN) displayed low resolve by conventional standards, it is less clear whether becoming a POW is also evidence of shirking. Indeed, many soldiers surrendered en masse, sometimes 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 capture. Stalin's Order 270, which equated surrender with treason, stipulated that "every soldier is obligated ... to demand that their superiors fight to the end if part of their unit is surrounded."

Another caveat relates to MIA as a measure of shirking. Although this approach is consistent with past literature (Costa and Kahn, 2003), it may appear problematic because some of the unaccounted soldiers surely died fighting. Our interpretation is motivated by the fact that to avoid being held personally responsible, Soviet officers were notoriously reluctant to report unaccounted soldiers as DDTs or POWs. Instead, common wartime practice was to report them as MIA, as a Russian Defense Ministry official acknowledged:

> By official reports, out of our five million-plus missing in action, just 100 thousand 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.[7]

The fact that some missing soldiers might have been KIA leads to a measurement error that should attenuate our estimates. We do not expect this error to be very large, since the

[6] We classify deaths in captivity as POWs, and executions as PUN.

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

quoted numbers suggest that $\Pr(\text{POW}|\text{MIA}) = 4.5\text{M}/5\text{M} = 0.9$.[8]

Finally, one should be careful when using military decorations as indicators of high resolve. Soldiers could be decorated (or denied decoration) for reasons unrelated to performance, like perceived political loyalty or ethnic discrimination. We therefore focus on a subset of awards — Medals "For Courage", "For Battle Merit", Order of Glory, and Hero of the Soviet Union — that specifically recognized individual performance in situations involving a risk to life, and which required detailed descriptions of individual acts (see Appendix A3 for examples).[9] Our estimations also control for observable factors like ethnicity and soldier's region of origin, which commanders could have used to infer one's perceived political loyalty when deciding whether to recommend them for an award.

Ultimately, our measures of combat resolve are valid only if they correlate in a predictable fashion with military units' aggregate performance. To establish this, we constructed an aggregate panel dataset with monthly observations for each major Soviet combat unit. Using monthly orders-of-battle from secondary sources (Fes'kov, Kalashnikov and Golikov, 2003), we matched combat units to soldiers by unit assignment and discharge date.[10] The result was an unbalanced panel dataset of 56,225 unit-month observations, tracking 5,756 active combat units over 48 months from June 1941 to May 1945.

Using official descriptions of 225 major battles from the People's Memory database, we classified each battle as resulting in a territorial gain, loss, or no change for Soviet forces (see Appendix A4 for details), and added these battles to the panel dataset. Because these were large, army-level operations, linking battles to units required establishing the "parent" army for each corps, division, regiment, and battalion in our monthly panel, and filtering soldiers' records to include only those corresponding to the unit at the time of the battle.[11] For each unit-month, we calculated monthly proportions of soldiers who were KIA/WIA, MIA, POW, DDT, PUN, or received one of the four valor decorations.

Using this panel, we estimate a regression where the dependent variable is a dummy equal to one if the battle resulted in territorial gain and no loss, and the covariates are unit-

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

[9] We exclude career service awards and hybrid medals like the Order of the Patriotic War, which was awarded on both an individual basis and collectively to units, towns, factories and categories of veterans.

[10] These combat units correspond to the "operational-tactical" tier in Soviet military organization (Grechko, 1976), and include corps (6%), divisions (19%), brigades (19%), regiments (42%), battalions (9%), and other large formations under direct army command (5%). 7,531,315 soldiers (65% of probabilistically linked records) had information on both unit assignment and discharge dates.

[11] We were able to match 27,368 of the unit-months (49%) to at least one of the 225 battles.

month proportions of soldiers' battlefield outcomes. We include fixed effects for units, years, and months, to remove confounding due to unit-level and temporal factors. The results, in Table 1, show that the aggregate success of army units correlated positively with higher casualty rates and higher rates of medals, and negatively with all measures of shirking — MIA, POW, DDT, and PUN. Since any reasonable measure of combat resolve should be correlated with higher operational effectiveness, this suggests that our proxy measures correctly capture this latent quantity. In section 7, we exploit this correlation structure further to build a scalar index of combat resolve.

Predictor	Coefficient (S.E.)	Implied correlation with combat resolve
KIA/WIA	0.1 (0.1)*	Positive: KIA $\rightarrow$ high resolve
MIA	-0.3 (0.1)**	
POW	-0.6 (0.2)*	
DDT	-1.9 (0.3)**	Negative: MIA/POW/DDT/PUN $\rightarrow$ low resolve
PUN	-1.1 (0.5)*	
Medals	0.2 (0.1)**	Positive: Medal $\rightarrow$ high resolve

OLS coefficients from linear probability regression: dependent variable is one if Red Army gained territory, zero otherwise. Unit of analysis is military unit by month (N = 27,368). Predictors measured as proportions per unit-month. Model includes fixed effects for units, years, months. Standard errors clustered by unit, battle. Observations weighted by number of soldiers. Significance levels: $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\dagger}p < 0.1$.

Table 1: Observable battlefield outcomes and army unit performance.

3.3. Data on repression

Our data on repression come from the Victims of Political Terror archive, maintained by the Russian human rights organization Memorial. It draws on declassified case files from Russian federal and regional archives, the Commission for the Rehabilitation of Victims of Political Repression, regional NGOs, and "Memory Books." The 2,747,582 Memorial records do not include victims of famines, deportations, and counterinsurgency operations, which were largely concentrated in the national republics outside Russia. Due to the under-representation of former Soviet republics other than Russia in Memorial, we limited our geographic scope to the RSFSR in its 1937 borders.[12] Using the same approach as with

[12] See Appendix A7.10 for a pooled analysis of Soviet Russia and Ukraine.

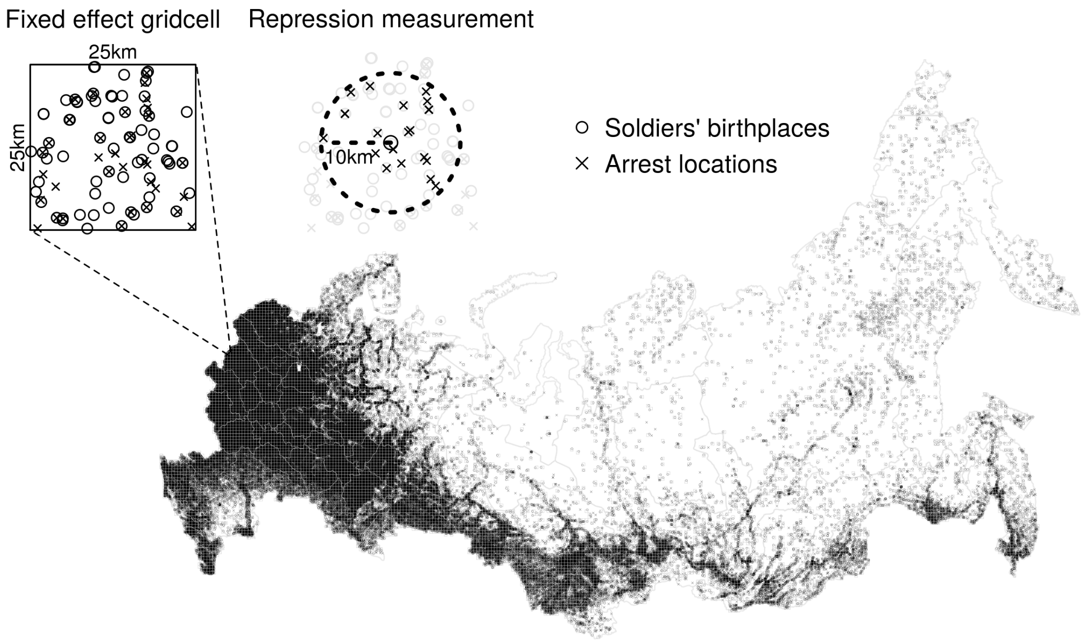


Figure 2: Geographic distribution of soldiers' birthplaces and exposure to repression.

military data, we found geographic coordinates for 2.15 million pre-WWII arrests (78%), using victims' residential addresses (where available) or birthplaces.

As Figure 2 illustrates, we measure exposure to repression by counting the number of arrests in the vicinity of a soldier's birth location:

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

where the logarithmic transformation is used to reduce skewness.[13] This geographic measure of repression rests on the idea that people are more aware of arrests 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 (*voenkomaty*) where soldiers were drafted. 18% of soldiers were born within 1 km of their draft

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

location, 23% within 10 km, 53% within 100 km. 29% of soldiers were drafted by the *voenkomat* closest to where they were born. If most soldiers remained in their areas of birth until mobilization, they were likely also there in 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. 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 pre-war censuses do not provide information on population counts below the district level.[14] In the analyses below, we adjust for several proxies of local population density (distance to the administrative center, road junctions, and farmland). 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. Most conservatively, we replicate our results at an aggregate, district level, directly controlling for local population size and urbanization from the 1926 census.

3.4. Additional data

We collected additional data on historical political economy, logistics and ethnicity. To measure local state capacity, we use the distance in 1935 from each birthplace to the nearest district administrative center, where local NKVD branches were based (TsIK, 1935). To distinguish between urban and rural areas, we calculated hectares of cropland within 10km of each birthplace, using geo-referenced maps of economic activity from the 1937 *Large Soviet Atlas of the World* (Gorkin et al., 1937, 155). To account for the targeting of peasants during collectivization, we calculated the number of state farms within 10km of each birthplace (Gorkin et al., 1937, 161). We also georeferenced information on historical road and railway junctions (Afonina, 1996), to help capture local economic development.

With few exceptions, the military records do not include information on soldiers' ethnicity. Because national minority status could confound the relationship between pre-war repression and wartime behavior, we address this issue by building a nationality classifier

[14] District-level geographic precision allows us to estimate population counts for small areas (e.g. grid cells, see footnote 17), but not point estimates for specific birth locations.

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 predict (with 97% out-of-sample accuracy) whether a surname's represents Russian nationality. We then assigned to each military personnel record a dummy variable equal to one if the surname's predicted nationality is Russian.[15]

4. Empirical Approaches

4.1. Ordinary Least Squares with Grid Cell Fixed Effects

Our first empirical strategy is motivated by the observation that Stalin's terror was locally arbitrary in its selection of targets, net of group-level factors. We can treat exposure to repression as plausibly exogenous conditional on the observables that the regime used in selecting victims. One such observable was ethnicity: authorities often viewed minorities as politically disloyal and subjected them to greater coercion. Another was socioeconomic: 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: western borderland regions, the Far East, and areas with a history of peasant uprisings faced higher arrest quotas (Getty and Naumov, 1999).

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, and distances to the nearest administrative district center and nearest road junction. 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

[15]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 A5).

balanced comparisons, we partition 1937 Soviet Russia's administrative regions into a regular 25×25km grid (shown on Figure 2), and include a fixed effect for the grid cell k in which soldier i was born. Because 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.[16]

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 25×25 km 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). Although arrest density varies between cells and regions, the *local* (within-cell) spatial distribution of repression appears quite arbitrary.

4.2. Geographic discontinuities

Our second empirical strategy utilizes the fact that regional state security officials had discretion when implementing central orders. A town located in a region run by a zealous state security chief could face more repression than a nearby town from a different region with less ambitious or cruel security officials. These discontinuities across regional borders can serve as a plausibly exogenous source of variation in repression levels.

The idiosyncratic qualities of local security personnel cannot be measured directly, but we can infer them indirectly by identifying administrative regions (*oblasti*) where repression was lower or higher than expected, conditional on observables. We first estimate how much the level of repression in each region deviated from what would be expected given basic observables like local population and urbanization. We do so using regression

$$\text{Repression}_j = \alpha + \beta_1 \cdot \ln\Big(\text{Population}_{k[j]}\Big) + \beta_2 \cdot \text{Urbanization}_{k[j]} + \epsilon_{kj}, \quad (4)$$

where j indexes birth locations and k indexes grid cells.[17] From the above regression,

[16] The median (maximum) distance between two birth locations in a grid cell is 8.9 km (16.9 km).

[17] Data on population and urbanization come from the 1926 Soviet Census, which reports them at the level of district (*rayon*). To disaggregate these 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.

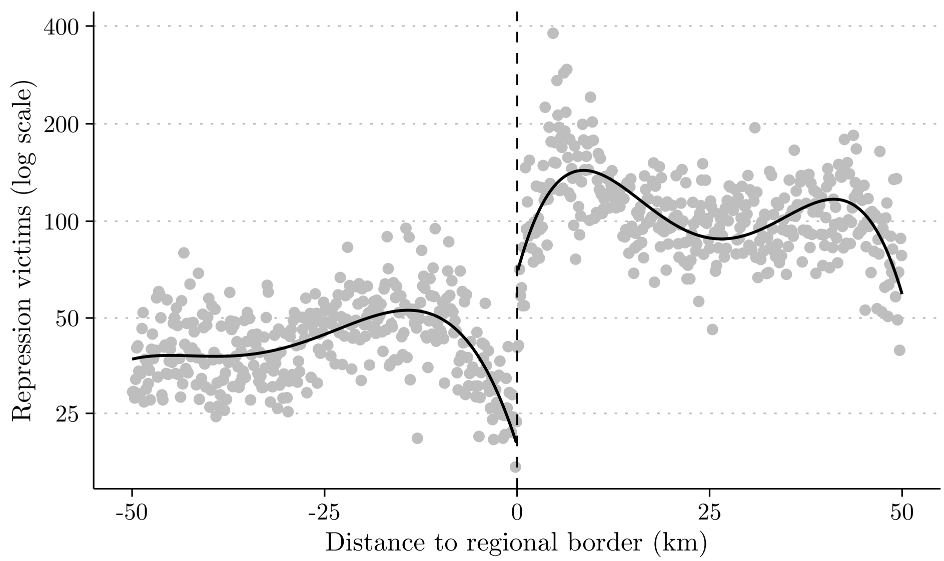


Figure 3: Discontinuity of repression at regional borders (local means and polynomial fit).

we calculate the average residual $\bar{\epsilon}_r$ for each region r. We then select pairs of adjacent regions (r, r') where the absolute difference between average residuals $\bar{\epsilon}_r$ and $\bar{\epsilon}'_r$ is at least one standard deviation; that is, we select adjacent regions with highly contrasting levels of repression that cannot be explained by their basic background characteristics.

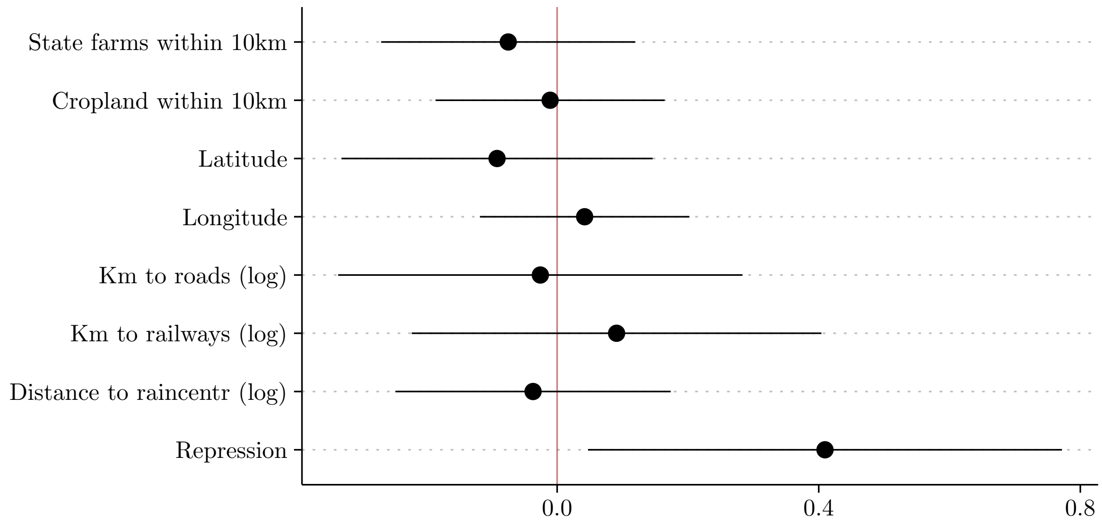
Let d_{jr} denote the distance from birth location j in the region r to the border of the nearest region. Define the forcing variable

$$\delta_{jr} = \begin{cases} d_{jr} & \text{if } \bar{\epsilon}_r > \bar{\epsilon}'_r, \\ -d_{jr} & \text{otherwise.} \end{cases}$$

To see how this forcing variable works, suppose that $\delta_{jr} = -2$. This means that birthplace j is inside a (relatively) 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 j in r' instead of r, the level of repression in j would have been higher, in expectation. This is a plausible counterfactual: just prior to the Great Terror, Soviet regions underwent a series of territorial reforms, which subdivided large regions into smaller, more "manageable" units (Shiryaev, 2011). To preclude comparisons of wildly different locations, we restrict this analysis to birthplaces within ±50 km of regional borders.[18]

Figure 3 plots the average levels of repression as a function of the forcing variable δ_{jr}.

[18] Appendix A6 reports a map of locations that are included in the regression discontinuity analysis.



RD coefficients with 95% confidence intervals adjusted for clustering by grid cell. All outcomes are normalized for comparability.

Figure 4: Balance tests for discontinuity at the border.

We see a discontinuous jump in repression levels as we move from less repressive to more repressive regions. The bias-corrected local-polynomial estimate of the discontinuity effect (Calonico, Cattaneo and Titiunik, 2015) is 0.0.85 (S.E. clustered by grid cells is 0.33) on the logarithmic scale. On the natural scale, estimated at the sample average, the effect is about 47 victims, as can be seen in Figure 3.

An important concern with this design is that other things might change discontinuously across borders. To assess the magnitude of this problem, we conduct a series of balance tests reported in Figure 4. It shows the estimated discontinuity effects of eight variables, normalized to have a standard deviation of one for comparability. Only repression shows a discontinuous jump, suggesting that border discontinuities are not capturing differences in any observables other than repression. However, this does not rule out the possibility that the borders affected other relevant variables that we cannot measure.

With the latter caveat in mind, we exploit the border effects in a fuzzy regression discontinuity design (FRDD) using two-stage least squares:

$$\begin{aligned} \text{Repression}_{j[i]} &= \alpha \cdot \mathbb{1}\{\delta_{jr[i]} > 0\} + g_1\left(\delta_{jr[i]}\right) + \boldsymbol{\beta}'\boldsymbol{X}_{ij} + s(\text{lon}_j, \text{lat}_j) + \epsilon_{1i}, \qquad (5) \\ y_i &= \gamma \cdot \widehat{\text{Repression}}_{j[i]} + g_2\left(\delta_{jr[i]}\right) + \boldsymbol{\beta}'\boldsymbol{X}_{ij} + s(\text{lon}_j, \text{lat}_j) + \epsilon_{2i}, \end{aligned}$$

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 for repression. Both stages include covariates and spatial splines but exclude grid cell fixed effects — because cells are nested within regions, by construction, the instrument cannot vary within cells.

4.3. Clustering and Weights

The outcome variables in our study are measured at 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 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).

5. Main Results

Table 2 reports estimated coefficients on repression ($\hat{\gamma}$) from fourteen regression models. For each of the seven battlefield outcomes, we report estimates from OLS with grid cell fixed effects, and FRDD with the border instrument. Estimates in the first row suggest that soldiers from places with more repression died or were wounded on the battlefield at higher rates. The conclusion is consistent across both designs. The 0.5 OLS coefficient implies that increasing repression in the soldier's birthplace within the same 25×25km grid cell from zero to 32 people (first quartile in sample) meant a $[\ln(32+1) - \ln(0+1)] \times 0.5 \approx 3.5 \times 0.5 = 1.8$ percentage point higher chance of death or injury, after adjusting for soldier-specific and location-specific covariates. Based on the FRDD estimate, the respective change is approximately $3.5 \times 1.8 = 6.3$ percentage points. Given that the mean KIA/WIA rate in the sample is about 21 percent, these magnitudes are substantial.

The next batch of outcomes represents soldiers' proclivity to shirk. The first outcome in this batch is an index *Flee*, indicating whether a soldier was reported as either missing, surrendered, deserted, defected, committed treason, or punished for misconduct. The coefficients on repression are negative for both OLS and FRDD but significant only at 90% confidence level after clustering standard errors by soldiers' birth location and grid cell.

	Coefficient on Repression		
Outcome	OLS	FRDD	Mean outcome
KIA/WIA	0.5 (0.1)**	1.8 (0.6)**	21.4
Flee (index)	-0.1 (0.1)$^{\prime}$	-0.5 (0.3)$^{\prime}$	25.6
MIA	-0.1 (0.1)*	-0.7 (0.2)**	20.1
POW	-0.03 (0.04)	0.3 (0.2)	5.7
DDT	0.01 (0.002)**	-0.002 (0.005)	0.2
PUN	-0.01 (0.005)	-0.04 (0.02)$^{\prime}$	0.8
Medals	-0.2 (0.1)**	-0.9 (0.3)**	17.9
Birthplaces	180,895	38,521	
Gridcells	12,176	2,094	
Soldiers	11,351,164	2,828,431	

Outcomes are on a percentage scale (0 to 100): killed or wounded in action (KIA/WIA), missing in action (MIA), prisoner of war (POW), defected, deserted, or committed treason (DDT), punished for battlefield misconduct (PUN), any of the last four (Flee), and receiving at least one valor decoration (Medal). Standard errors in parentheses, clustered by birth location and grid cell. All regressions include grid cell fixed effects, and individual and birth location-level covariates. Observations weighted by record clustering probability. FRDD excludes locations in non-matched regions and $>$ 50km from regional borders. FRDD first-stage $\mathcal{F}$ = 13.6. Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 2: Coefficients on repression for each battlefield outcome.

Substantively, the increase of repression from zero to the first quartile is associated with a reduced chance of shirking (flight) by 0.4 (3.5×-0.1) to 1.8 (3.5×-0.5) percentage points, depending on the specification. Coefficients for MIA, the most frequent flight indicator, are also consistently negative and significant at the 95% confidence level.

Estimates for the other individual indicators of shirking offer more mixed results, from negative for PUN to negative, null and inconsistently positive for POW and DDT. Some of this variability may reflect the idiosyncrasies of reporting. Most POW's, as we noted, were officially misreported as MIA's and cases of DDT were rare, raising the possibility that these outcomes emerged under qualitatively different circumstances. Notably, positive coefficients on POW and DDT do not survive robustness tests, but coefficients for MIA and PUN remain negative and significant in most specifications (Appendix A7.4).

The final row in Table 2 reports the estimated effects of repression on battlefield initiative, measured through a soldier's receiving at least one decoration for valor (*Medals*).

The estimated coefficients are negative and significant at the 99% level of confidence. In substantive terms, increasing repression from zero to the first quartile is associated with 0.7 (3.5×-0.2) or 3.1 (3.5×-0.9) percentage point lower probability of a medal. These magnitudes are substantial, given that 18 percent of soldiers received such awards.

6. Caveats and Robustness Checks

Some of the FRDD estimates are larger than OLS estimates, which may be concerning. This could be the artifact of FRDD analyses using a different sample than OLS: locations within 50km of matched borders. However, OLS coefficients are nearly identical if we restrict the OLS sample in the same way as FRDD (Appendix A7.7). A second explanation is that OLS estimates may be attenuated due to errors in the measurement of repression. As the instrument should alleviate attenuation bias, the larger FRDD estimates make sense.

A more fundamental consideration is that the estimates diverge because they represent different quantities. Under the design assumptions, FRDD estimates represent the local average treatment effect of repression induced by proximity to regional borders. Repression induced by exogenous factors, like border discontinuities, may have appeared more arbitrary and, as such, had a stronger overall impression on those who were exposed to it.

Finally, and most importantly, we cannot directly rule out the possibility that differences in coefficient size are due to violations of the exclusion restriction. The same idiosyncratic factors that led to higher arrest rates across regional borders could also have led to differential battlefield outcomes through channels other than repression. More zealous local administrators could have been more efficient in drafting soldiers and transporting them to the front in the early stages of the war, when death rates were high. They may also have had a better capacity to keep records of these soldiers. The evidence for these alternative pathways is weak, however. In additional analyses (Appendix A6), we find no discontinuous border effect on the timing at which soldiers started or terminated their service. We do find that soldiers in higher-repression border regions were slightly less likely to have missing discharge records, but there is little evidence that such missingness is consequential for our findings (see below). While partly reassuring, these results do not rule out the possibility of other violations of the exclusion restriction.

We conducted a battery of robustness tests (Appendix A7). One 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 (Imai et al., 2009), 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 41,274 clusters (22% of total), or 20,637 matched pairs (Appendix A7.1). We ran our analyses on the subset of soldiers who were born in these matched clusters. To further account for local population size, we averaged individual outcomes at the level of birthplace and ran the same regressions on these aggregated data, using full and matched samples. We did the same at the level of districts ($N = 361$), directly controlling for local population size and urbanization from the 1926 census. Table 3 reports estimates from these reanalyses of alternative samples and units of analysis, which are consistent in sign and significance with those in Table 2.

	KIA/WIA	Flee	Medals
Units of analysis	Soldiers (matched clusters), $N =$ 4,489,873		
Coef. for Repression	0.5 (0.1)**	-0.2 (0.1)$^{\prime}$	-0.4 (0.1)**
Units of analysis	Soldiers' birthplaces, $N =$ 180,895		
Coef. for Repression	0.6 (0.1)**	-0.2 (0.1)**	-0.2 (0.04)**
Units of analysis	Soldiers' birthplaces (matched clusters), $N =$ 41,272		
Coef. for Repression	0.6 (0.1)**	-0.1 (0.1)	-0.3 (0.1)**
Units of analysis	Soldiers' districts, $N = 336$		
Coef. for Repression	1.3 (0.4)**	-0.6 (0.4)	-0.6 (0.2)*

Outcomes on percentage scale. Clustered standard errors in parentheses. First set of models includes grid cell fixed effects, individual and birth location-level covariates. Second and third inlcude grid cell fixed effects, birth location-level covariates, and birth location-level averages of individual-level covariates. Fourth includes regional fixed effects and district averages of individual and birth location-level covariates. Observations weighted by record clustering probability (1) or number of soldiers (2-4). Significance levels (two-tailed): $^{\dagger}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 3: Alternative Samples and Units of Analysis.

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 war's end. Our conclusions remain robust when we exclude individuals whose discharge reasons are not observed (Appendix A7.2). Our results also hold when we consider a more selective set of medals (Appendix A7.3), when we consider only the subset of soldiers who did not flee, and when we separate out soldiers who received medals posthumously (Appendix A7.5).

Our analysis treats battlefield outcomes as independent across individual soldiers, although in reality soldiers do not make decisions in isolation. Following the econometric approach of Carrell, Sacerdote and West (2013) for the study of peer effects, we account for the interdependence of soldier-level outcomes within army units (Appendix A7.8). The estimates remain consistent with our main results. These additional results indicate that repression may have impacted soldiers' behavior not only directly, but also indirectly, through the behavior of their peers.

As an additional robustness check on our identification strategy, we exploit exogenous variation in repression due to the logistical costs of accessing and transporting arrestees to prison colonies by railroad. The estimates based on this alternative instrument align with those we report here (Appendix A7.9).

Finally, because limiting our analysis to soldiers born inside Russia risks missing a large part of the story, we expanded the sample to include soldiers born in Ukraine — the USSR's second-most populous republic. Our conclusions remain unchanged after running our models on this expanded sample (Appendix A7.10).

7. Distributional Test Using a Combat Resolve Index

So far we have tested our theoretical predictions using individual-level outcomes as indirect measures of combat resolve, which was in part justified by the specific correlation structure between the individual-level outcomes and the battlefield success (Table 1). We now exploit these correlations further by constructing a scalar index of combat resolve for each unit-month, to conduct a more direct, distributional test of our predictions.

Building on the results in Table 1, we first estimate a semiparametric regression

$$\text{Territorial gain}_{it} = \sum_k f_k(y_{itk}) + \epsilon_{it},$$

where the outcome variable is equal to one if the battle in which unit i took part at time t resulted in territorial gain and y_{itk} is the proportion of soldiers in unit i and year-month t with the outcome $k \in \{\text{KIA/WIA, MIA, POW, DDT, PUN, Medal}\}$. f_k is a smooth function for input k, approximated by cubic regression splines.

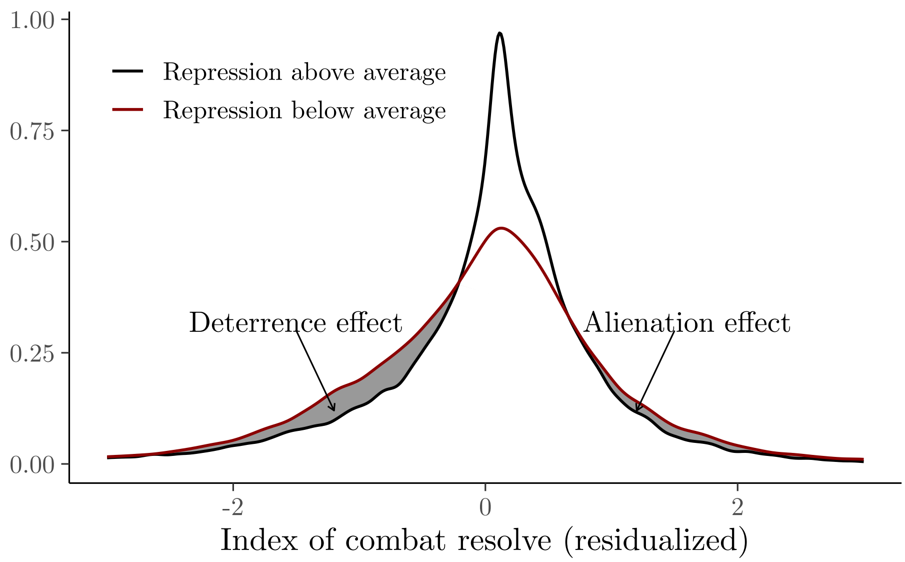


Figure 5: Density estimates of CRI by repression levels.

The combat resolve index (CRI) for unit i in year-month t is the predicted probability that a unit participated in a successful battle, conditional on the proportion of its service members having experienced each of the six battlefield outcomes:

$$\text{CRI}_{it} = \sum_k \hat{f}_k(y_{itk}),$$

where $\hat{f}_k$ is the estimated function. By Proposition 1, CRI should have a higher mean *and* lower variance when a unit comprises soldiers with higher average exposure to repression.

We residualize CRI and repression measures by regressing each of them on fixed effects for units, years, and months, and the covariates used in soldier-level regressions (averaged over unit-months). Figure 5 shows the kernel density estimates of residualized CRI for units with repression levels above and below the residualized sample average. The mean CRI in units with "above average repression" is 0.06 standard deviations higher than in units with "below average repression." The difference is statistically significant at the 95% level after accounting for the clustering of standard errors by unit and battle (S.E. = 0.02). Consistent with Proposition 1, *average* combat resolve increases in repression.

Figure 5 also shows that, in units where more soldiers were exposed to repression, the distribution of CRI is compressed from both sides toward the center. This is consistent with the prediction that repression decreases the variance of combat resolve. The empirical

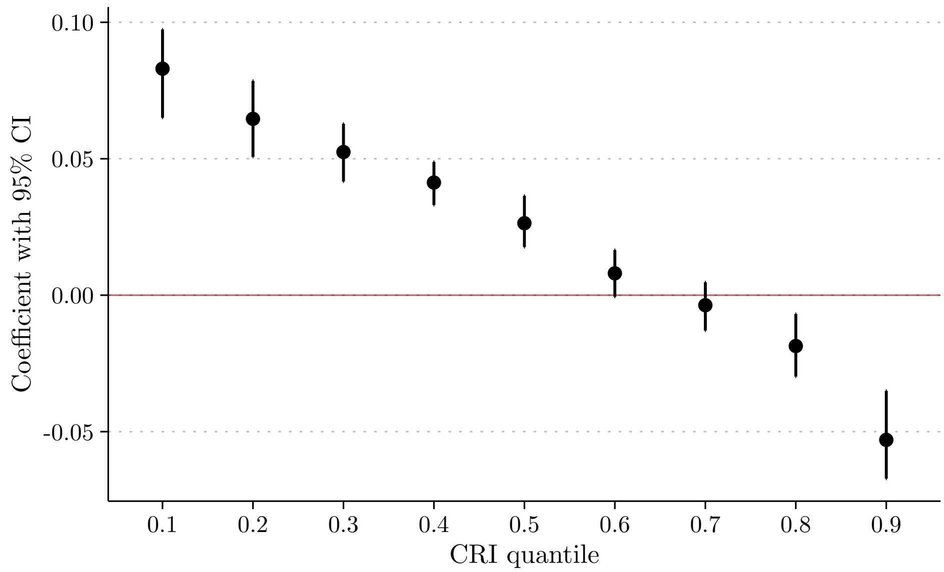


Figure 6: Quantile regression coefficients on repression.

patterns in Figure 5 broadly resemble the theoretical prediction shown in Figure 1. In units with more exposure to prewar repression, we observe a deterrent effect of low-level CRI being pushed up and an alienation effect where high-level CRI is pushed down.

We test the prediction about variance reduction more formally using conditional quantile regression. The dependent variable in this regression is the residualized CRI and the independent variable is residualized repression. Figure 6 shows estimated quantile regression coefficients for each decile of residualized CRI. The estimated coefficients are positive for the lower quantiles and negative for upper quantiles, consistent with the variance reduction hypothesis. Units with high average exposure to repression had higher CRI at the low end of the distribution and lower CRI at the high end of the distribution, compared to units with low average exposure to repression.

8. Interpretation

The above empirical patterns align closely with the logic of "perfunctory compliance." The fact that soldiers exposed to repression were more likely to die, less likely to flee, and less likely to receive awards for valor indicates that repression simultaneously increased extrinsic motivations to fight while sapping intrinsic ones. This evidence contradicts the view that repression uniformly incentivizes over- or under-performance (Edele, 2017).

Rather, repression spurs higher performance by soldiers who would otherwise show low resolve and lower performance by soldiers who would otherwise show high resolve.

The explanation that prewar terror conditioned soldiers into conformity resonates with historical accounts. As Merridale (2006, 45-46) writes, soldiers who 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 . . . than to remain alone, condemned to isolation and the fear of death." Glantz (2005, 446) notes that repression incentivized overly cautious decision-making and undermined independent thinking at all levels of the Red Army; those "who survived the terror were paralyzed [and] afraid to display initiative." As Overy (1998, 32) observes, "the result [of Stalin's terror] was the triumph of military illiteracy over military science, of political conformity over military initiative."

An important alternative interpretation of these results is that they reflect the differential treatment of soldiers by the state, rather than differential behavior by soldiers. The positive relationship between repression and combat deaths may exist because authorities used soldiers from heavily-repressed areas as "cannon fodder," with more dangerous assignments, worse equipment and leadership. The negative relationship with flight may suggest that authorities more closely monitored these soldiers, or sought to minimize their opportunities to cross the front line. The negative relationship between repression and medals may reflect unit commanders' hesitancy to recommend, and higher authorities' refusal to approve, decorations for soldiers from "problematic" parts of the country.

Some of these possibilities are more facially plausible than others. Because our treatment variable captures geographic exposure to repression, discrimination during a soldier's assignment to units would have required not only information about the soldier's personal history, but also about how many of the soldiers' neighbors the secret police had arrested. Military commissariats did not have this information — it was collected and closely guarded by a different agency — nor would they have had the time to process it at the height of war. While valor decorations required proper vetting of soldiers' backgrounds, enlistment was a fast-paced and haphazard process (Glantz, 2005, 470), which made this type of selective assignment difficult to implement consistently or at scale.

The same applies to possible discrimination of soldiers after they are assigned to units. As we document in Appendix A8, the Service Record Card Files available to unit commanders contained information on soldiers' families and political background, but no information on whether the soldiers in question were from "problematic" locations where many people had been arrested. Such information would be necessary for unit commanders to discriminate against individual soldiers on individual level (net of their other observable characteristics for which we control) by assigning them to more dangerous tasks or by monitoring them more closely to prevent shirking.

A discrimination interpretation of our findings for flight and medals is difficult to reconcile with other empirical patterns. To the extent that authorities sought to minimize exposed soldiers' opportunities for flight, the most direct means of doing so — assigning soldiers to rear duties — would run counter to the "cannon fodder" interpretation of our KIA/WIA results, since such assignments would also minimize soldiers' exposure to enemy fire. As regards medals, Table 2 shows that soldiers from high-repression locations were *not* more likely to be punished for real or presumed violations of military code, which is the opposite of what we should see if these soldiers were subject to more scrutiny.

We now consider these arguments more directly, and evaluate whether, instead of conformity, our empirical findings might reflect discrimination at the individual or unit level.

8.1. Rank Advancement

If soldiers from high-repression areas faced discrimination during the allocation of military awards, then it seems reasonable to expect them to have also suffered discrimination during promotions. Rank advancement decisions followed a structurally similar bureaucratic process to medals, but were more weakly tied to individual combat performance. Similar to medals, unit commanders would recommend individuals for promotion, with conferral authority residing with higher ministerial or party authorities (see Appendix A3). 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 fill higher-ranking billets, ethnic or religious quotas, disciplinary records, party membership, and other indicators of political loyalty. Opportunities for discrimination to enter the promotion process were more abundant than in the conferral of medals.

If discrimination drove our results on medals, we should see a similar negative effect for promotions. We re-estimate our baseline regressions with the outcome variable equal to one if a soldier advanced ranks at least once during the war, and zero otherwise. As the first column in Table 4 reports, we find no evidence that soldiers with higher exposure to repression were less likely to be promoted. Unless rank advancement was insulated from political considerations while the conferral of decorations was not (which seems implausible), discrimination does not appear to be a strong alternative to our explanation.

	Promotion	Infantry	Penal
Model		OLS	
Coef. for Repression	0.01 (0.03)	-0.1 (0.1)*	0.003 (0.004)
Mean Y	10.1	82.9	0.1
Birthplaces	158,154	116,204	116,204
Gridcells	10,944	9,580	9,580
Soldiers	6,951,642	2,342,735	2,342,735
Model		FRDD (First-stage $\mathcal{F}$ = 14.4)	
Coef. for Repression	0.02 (0.1)	-0.4 (0.3)	0.02 (0.01)$'$
Mean Y	10.1	82.2	0.1
Birthplaces	33,251	24,823	24,823
Gridcells	1,959	1,810	1,810
Soldiers	1,733,432	578,275	578,275

Outcome = at least one advancement in rank (Promotion), assignment to infantry branch (Infantry) or penal unit (Penal), measured on percentage scale (0 to 100). Standard errors in parentheses, clustered by birth location and grid cell. Models include grid cell fixed effects, individual and birth location-level covariates. Observations weighted by record linkage probability. Rank information unavailable for 39% of soldiers. Branch information unavailable for 79% of soldiers. FRDD 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 4: Repression, Promotions and Army Branch Assignment.

8.2. Assignment to Army Branches

Another mechanism through which discrimination could explain our results is that soldiers from repressed places died in larger numbers because they were selectively assigned to serve in more dangerous positions within the army. To assess this possibility, we check whether soldiers from highly repressed areas were more likely to be assigned to the infantry branch of the army — where direct exposure to enemy fire was higher than in other branches, like artillery and aviation — or to so-called "penal units," which were routinely ordered to charge through minefields and machine-gun fire.

Contrary to the "cannon fodder" hypothesis, the estimates in Table 4 (second column) suggest that conscripts from high-repression areas were *less* likely to serve in the infantry, although the result is significant only for OLS. There is mixed evidence of a positive correlation between prewar repression and assignment to penal units (third column). Even if we take the FRDD estimate at face value, its magnitude is small: a soldier from a location with 32 arrests (first quartile) was $3.5 \times 0.02 = 0.07$ percentage points more likely to serve in a penal unit than a soldier from a place with no arrests. Given that just 0.1% of soldiers were assigned to penal units during the war, this accounts for a small fraction of the estimated effect of repression on battlefield deaths and injuries. Indeed, our results on KIA/WIA are almost identical if we exclude soldiers assigned to penal units.[19]

8.3. Assignment to Combat Units

A third mechanism through which discrimination could explain our results for death rates operates at the level of combat units rather than army branches: soldiers from highly-repressed locations may have been selectively assigned to units with older equipment, less competent commanders, in more dangerous locations.

We can account for some of the variation by including fixed effects for the unit in which each soldier served and the time of their deployment in that unit. In the case of OLS, we augment the baseline regression to include fixed effects for combat units in which a soldier fought, and fixed effects for the month of the war, ranging from June 1941 to May 1945. The variable that indexes combat units has almost 12,000 unique values and identifies the smallest-echelon unit mentioned in each record.[20] Because 30% of soldiers had served in more than one unit, we disaggregated soldiers' records by unit assignment for this analysis.

Coefficient estimates remain similar to baseline specifications after we adjust for unit and time fixed effects (Table 5). Depending on the estimator, increasing repression from zero to the first quartile (32 arrests) increased one's chances of death or injury by 0.7 to 4.2 percentage points, decreased flight by 0.4 to 2.5 points, and chances of receiving a medal by 0.1 points (as before, percentage change $\approx 3.5\hat{\gamma}$). Although the medals result loses significance, the other estimates remain consistent with the logic of conformity, 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.

As regards the quality of equipment, documents from Soviet military archives reveal significant temporal variation in the supply of arms and ammunition across fronts — driven largely by industrial production, stockpiles, and the pressures of ongoing campaigns — but no evidence that authorities selectively withheld support from specific units

[19] The OLS coefficient in the restricted sample is 0.47 (S.E. = 0.08), close to the estimate in Table 2.

[20] 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		
Coef. for Repression	0.2 (0.04)**	-0.1 (0.03)**	-0.02 (0.03)
Mean Y	40.3	19.2	8.3
Birthplaces	134,351	134,351	134,351
Gridcells	9,808	9,808	9,808
Soldiers	5,470,129	5,470,129	5,470,129
Model	FRDD (First-stage $\mathcal{F}$ = 18.1)		
Coef. for Repression	1.2 (0.4)**	-0.7 (0.2)**	-0.03 (0.1)
Mean Y	51.9	20.9	7.6
Birthplaces	1,590	1,590	1,590
Gridcells	19,450	19,450	19,450
Soldiers	756,455	756,455	756,455

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 and date information. 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 5: Estimates Adjusting for Military Unit and Month.

on the basis of (average) pre-war repression levels (Appendix A8). The General Staff's monthly supply plans allocated resources across large formations like fronts and armies, but not operational-level units like divisions and regiments. If some units were nonetheless chronically under-supplied, our unit-level fixed effects should account for this variation. While we cannot rule out the possibility that discrimination existed within units at an individual level — net of age, ethnicity, and other observables — this would require that (a) unit commanders had information on prewar arrest levels near each soldier's birth location, and (b) commanders prioritized this information over subordinates' tactical needs when making decisions in battle, both of which are inconsistent with the qualitative evidence.

8.4. Wartime Learning

Additional evidence for the competing mechanisms underlying our results can be gained by examining how the relationship between pre-war repression and battlefield behavior changed over time. The incentive structure and the informational environment shifted substantially during the course of the war, due to increasingly harsh disciplinary measures within the Red Army and increasing awareness of German cruelty against Soviet civilians and captured soldiers. Depending on their exposure to pre-war repression, soldiers may have adjusted differently to this changing environment.

To examine temporal heterogeneity, we use our panel of combat units and estimate a regression where the coefficient on repression varies by month of the war:

$$y_{ut} = \text{Month}_t + \text{Unit}_u + \sum_t \gamma_t \cdot \text{Repression}_{ut} + \boldsymbol{\beta}'\mathbf{X}_{ut} + \epsilon_{ut}, \tag{6}$$

where y_{ut} is the percentage of a unit's soldiers that were KIA/WIA, fled (index), or received a medal at time $t \in \{\text{June 1941}, ..., \text{May 1945}\}$. Month_t and Unit_u are month and unit fixed effects. Repression_{ut} is the average exposure to repression in unit u and month t, and $\mathbf{X}_{ut}$ are the control variables from our baselines regressions, averaged by unit-month. Our quantities of interest are γ_t, the time-variant coefficients on repression for each of the three outcomes. As in the previous unit-level regressions, we weigh observations by the number of available records within each unit-month. We cluster standard errors by units.

Figure 7 shows the estimated coefficients with 95% confidence intervals for each of the three outcomes. With the understanding that one should be cautious in attributing temporal changes in coefficients to specific events, several patterns are worth noting.

The largest positive estimates for KIA/WIA and the largest negative estimates for Flee (late summer of 1942) both coincide with Stalin's issuance of Order No. 227 on July 1942, which outlawed "cowardice" and "panic." This change follows a similar negative turn in estimates for Flee following Stalin's Order No. 270 on August 1941, outlawing surrender. If such policies indeed drove these temporal shifts, this would indicate that soldiers exposed to pre-war repression were more sensitive to coercive incentives on the battlefield. These patterns could also reflect the higher responsiveness of repressed soldiers to reports about poor conditions in German POW camps, particularly in the second half of 1941.

Estimates for medals follow a different pattern, becoming more negative over time. This pattern reflects the fact that 90 percent of all valor decorations in the Red Army were for actions taken in the second half of the war, following the Battle of Kursk in July 1943. This battle was a key turning point in the war, marking Germany's final strategic offensive. As the Red Army began its long drive to Berlin, Soviet authorities sought to incentivize acts of bravery, by establishing new decorations (e.g. Order of Glory in

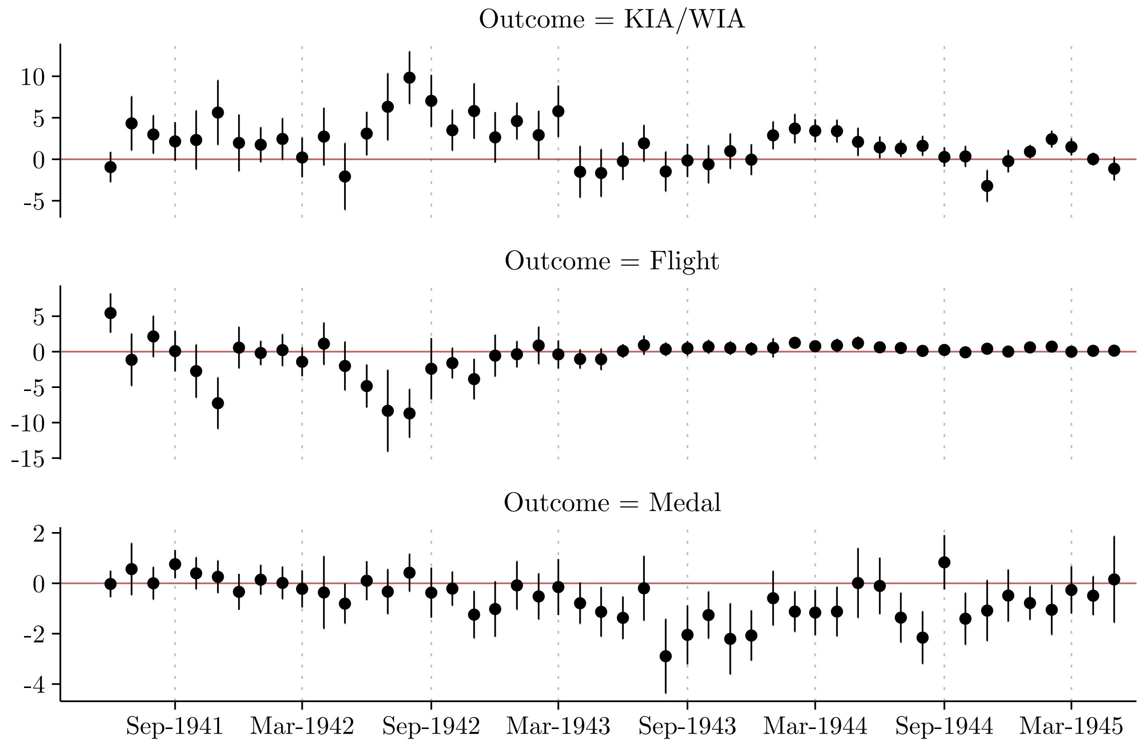


Figure 7: Time-variant coefficients on repression with 95% CI's.

November 1943) and amending eligibility criteria for others (e.g. "For Courage" in June 1943). The fact that effect estimates remain strongly negative throughout this period — at a time when battlefield exploits were more likely to be recognized — suggests that soldiers from repressed areas were not only more responsive to coercive incentives (deterrence effect), but they were also less responsive to positive inducements (alienation effect).

9. 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 shirk by fleeing the battlefields of the Second World War. They were also less likely to show personal initiative in battle, as far as we can infer from military decorations. While we cannot fully exclude the possibility that unobserved factors are driving these relationships, our analyses 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 zealots willing to go beyond their call of duty, but because they were more aware of what the state might do if they did not comply. Past repression may have compelled less-motivated soldiers to signal resolve, but it may also have decreased effort by highly-motivated types. The countervailing forces of deterrence and alienation helped resolve some principal-agent problems associated with fighting, but they did so by inculcating obedience at the expense of initiative, and raising the human costs of war.

References

Acemoglu, Daron, David H Autor, and David Lyle. 2004. "Women, war, and wages: The effect of female labor supply on the wage structure at midcentury." *Journal of political Economy*, 112(3): 497–551.

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.

Ager, Philipp, Leonardo Bursztyn, and Hans-Joachim Voth. 2022. "Killer incentives: Status competition and pilot performance during World War II." *Review of Economic Studies*, 89(5): 2257–2292.

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.

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, and David D. Laitin. 2008. "Religion, terrorism and public goods: Testing the club model." *Journal of Public Economics*, 92(10): 1942–1967.

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

Biddle, Stephen. 2004. *Military power: Explaining victory and defeat in modern battle.* Princeton University Press.

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.

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.

Cederman, Lars-Erik, Andreas Wimmer, and Brian Min. 2010. “Why do ethnic groups rebel? New data and analysis.” *World Politics*, 62(01): 87–119.

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.

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.

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*, 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.

Fes'kov, V.I., K.A. Kalashnikov, and I.F. Golikov. 2003. *Krasnaya Armiya v pobedakh i porazheniyakh 1941-1945 gg.[Red Army in victory and defeat 1941-1945].* Tomsk:Tomsk University Press.

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. 2005. *Colossus Reborn: The Red Army at War, 1941-1943.* Lawrence, Kansas:University of Kansas Press.

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

Grechko, Andrey Antonovich, ed. 1976. *Sovetskaya Voennaya Entsiklopediya [Soviet Military Encyclopedia].* Voenizdat.

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.

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).

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

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.

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.

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

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.

Scheidel, Walter. 2018. *The great leveler: Violence and the history of inequality from the stone age to the twenty-first century.* Princeton University Press.

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).

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, ed. 2015. *Velikaya Otechestvennaya voyna. Yubileynyi statisticheskiy sbornik [Grate Patriotic war. Anniversary statistical digest].* Moscow:Rosstat.

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*, 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, 169–191. Cambridge.

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

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.