The Economics of Civilian Victimization: Evidence from World War II Italy

Mattia Bertazzini Michela Giorcelli*

June 25, 2025

Abstract

This paper studies to what extent changes in economic incentives, and in particular a drop in soldiers' accountability, affect violence against civilians. It uses evidence from the Italian Campaign during World War II (1943-1945). Exploiting movements of the front lines as plausibly exogenous shocks to soldiers' accountability in a staggered difference-in-differences strategy, we find that: i) at front line activation, the probability of indiscriminate violence against civilians jumped 13 folds; ii) the effect was driven by a drop in accountability linked to a temporary break-down of the command chain; iii) crucially for policy, the effect is concentrated in proximity of less experienced units.

Keywords: civilian victimization, accountability, behavioral incentives, WWII, Italy

JEL Classification: D74, D90, N44

*Contact information: Mattia Bertazzini, University of Nottingham, CEPR, Sir Clive Granger Building, Nottingham, NG7 2QX. Email: mattia.bertazzini@nottingham.ac.uk; Michela Giorcelli, University of California, Los Angeles, NBER, and CEPR, 9262 Bunche Hall, 315 Portola Plaza, Los Angeles CA, 90095, USA. Email: mgiorcelli@econ.ucla.edu. Alessandro Brioschi, Dominik Loibner, and Daniel Perez provided excellent research assistance. We thank Sebastian Ottinger for his help in getting in touch with the BundesArchiv. We also acknowledge Jean-Paul Carvalho, Eric Chaney, Nicolas Duquette (discussant), James Fenske, Simon Gachter, Eleonora Guarnieri, Marco Molteni, Vincenzo Scrutinio, Martina Uccioli, Joachim Voth, and Noam Yuchtman, as well as seminar and conference participants at UC-Berkeley, UCLA, USC, Boston University, Oxford, Bocconi, Warwick, Trinity College Dublin, CERGE-EI, Zurich, St Andrews, Manchester, Birmingham, Trondheim, Kent, Nottingham, CCA, Bath, York, Bologna, and EHES 2022, Clio 2023, ASSA 2023, ASSA 2025 Conferences for comments and feedback.

1 Introduction

Civilian victimization — namely targeted violence against civilians during conflicts — is one of the most horrific consequences of war. Civilian killings represent its most extreme form. Previous research has documented that intrinsic preferences for violence are a key predictor for variation in its intensity, across conflicts, geographical areas, and perpetrator-victim combinations (Guarnieri, 2025; Guarnieri and Tur-Prats, 2023; Berman et al., 2017; Esteban et al., 2012).[1] Other scholars have highlighted the relationship between atrocities and preferences for following orders, as well as the role of conformity, when these incentivize killings (Milgram, 1974; Browning, 1993; Heldring, 2021). However, the same perpetrators can behave very differently within the same context, and while facing the same potential victims. The drivers of this erratic behavior remain a puzzle.

In this paper, we propose changes in behavioral economic incentives as a driver for civilian victimization. In particular, we study the effect of a drop in soldiers' accountability on the incidence of civilian killings. We use evidence from atrocities committed by German troops in the context of the Italian Campaign (1943-1945). The German Command allowed the use of selective violence against the Italian population, but pushed soldiers to exert a certain restraint towards civilians so to avoid indiscriminate one (Kesselring, 1954; Klinkhammer, 2006; Gentile, 2015).[2] However, his capacity to monitor and punish was affected by war-related events that affected soldiers' accountability.

We conceptualize our setting as a principal-agent problem, featuring information asymmetries, alongside a misalignment in incentives, between the principal (army command) and the agents (the soldiers) (Kalyvas, 2006). The principal is in favor of *selective* strategic violence, that targets specific civilians and is proportionate to context-specific war objectives, which is allowed. However, he forbids *indiscriminate* civilian victimization, that goes beyond the strategic goals of the armies and has a disproportionate, random component. In fact, this may alienate the local population, fuel insurgency and broadly jeopardize military occupation in the medium and long-term. The principal enforces restraint by punishing soldiers who do not follow his directives. Yet, he has limited monitoring capacity, due incomplete knowledge of the events on the ground, and disruptions to the command chain that affect his monitoring capacity.

We exploit the activation of a new front line as an exogenous shifter to the monitoring capacity of the

[1] This is in line with a large literature showing behavior is largely shaped by intrinsic preferences, often leading to "good" behavior due to intrinsic honesty. See, among others, Abeler et al. (2019) and Gächter and Schulz (2016)

[2] This general war conduct did not prevent the German Command, led by Field Marshal Kesselring, from explicitly (e.g. Fosse Ardeatine) or implicitly (e.g. Sant'Anna di Stazzema, Marzabotto) endorsing exemplar, large mass-killings, seen as necessary given the tactical situation and the counterinsurgency needs at specific points in the war. These episodes, however, were a tiny share of the total occurred during the Italian Campaign. In fact, "only" 18 out of 5,300 episodes involved more than 100 civilian victims, while 47 percent of them involved a single person.

principal. Between the Allied landing in Sicily in July 1943 and the surrender of Germany in May 1945, Allied and German armies confronted over several front lines gradually shifting from the South to the North of the country, following the Allied advance. The "activation" of a new front line, defined as the day when fighting started on a new defensive line, substantially decreased the German Command's monitoring capacity relative to the previous day, and resulted in a "fog of war" (Clausewitz (2003)[1832]). Tracing soldiers responsibilities and punishing them became more difficult as troops were shuffled around in response to Allied offensives, and the command chain was stretched in an attempt to contain the enemies (Ronchetti and Ferrara, 2014; Klinkhammer, 2006). This scenario arguably reduced soldiers' accountability (Collotti, 1963b; Kurowski, 2003; Gentile, 2012).

For identification, we rely on the fact that the exact location of the front lines, as well as the precise timing of their activation, were driven by strategic factors, orthogonal to the intention to kill civilians. We estimate a stacked difference-in-differences model that compares soldiers behavior in municipalities that fell into the combat zone upon the activation of a new front line but were outside it before (*treated* municipalities), with that of troops in municipalities that either remained outside or were already inside the combat zone (*comparison* municipalities). Treated and comparison municipalities were on statistically indistinguishable trends in civilian killings during the five days before front activation and were balanced in terms of proxies for fighting intensity - namely bombing and partisan activity - in the previous 5 to 20 days.

We use newly assembled panel data on civilian killings perpetrated by German soldiers during the Italian Campaign. We observe roughly 5,300 episodes that involved around 25,000 civilians. These represent close to the universe of killings against Italian civilians perpetrated by German soldiers, and correspond to roughly 20% of total civilians killed between 1943 and 1945 in Italy. For these episodes, we have detailed information on the number of people killed, their gender and age, whether they were civilians or partisans, as well as the identity of the perpetrators. Notably, the richness of our data allows us to distinguish episodes of indiscriminate and selective violence. We complement these data with information on front line position and activation, and use German military maps to geolocate the position of German units across the peninsula for the entire duration of the conflict. Finally, we collect information on other war related events, such as Allied bombings and resistance activity.

Consistent with incentives playing a central role viz-a-viz intrinsic preferences, we find that the activation of a new front line sharply increased indiscriminate violence. The likelihood of observing indiscriminate killings increased 13-fold, on average, upon front activation in treated municipalities relative to comparison ones. Such effect fades three days after the front line activation, which is consistent with the

monitoring capacity of the German Command being restored with a short lag. This pattern is driven by small-scale killings, likely unknown to the German Command as opposed to famous mass-killings. Conversely, the incidence of selective violence does not differentially change when a new front is activated. We interpret these results as the reduced-form effect of a drop in accountability.

Next, we investigate the first stage, namely the causal channel linking front line movements with the increase in violence. First, we rule out that the effect of front movements is driven by the selective sorting of troops. We show how front line activation did not systematically alter the identity of units at the front. Second, we proxy soldiers' insecurity with Allied bombing, distance from Allied troops and resistance activity, and show that none of these factors drive the estimated increase in indiscriminate violence. Third, we estimate a first stage in accountability. We document a significant drop in the monitoring capacity of the principal in the form of an sizeable increase in the likelihood for units at the front to become composite at activation. Composite units were patched-up divisions made from sub-units originally belonging to a different unit. These were typically difficult to control owing to their temporary nature and to a new division-level command chain. This result further corroborates that the front activation negatively shocked commanders' enforcement capacity, which in turn decreased soldiers' accountability.

We conclude by examining effect heterogeneity along several dimensions. First, we study whether underlying division characteristics affected soldiers' response to a drop in accountability. Importantly, we find that the more inexperienced and less trained units were more likely to use indiscriminate violence at front activation. We also document stronger violence against civilians in more urbanized and industrialized municipalities, where looting incentives were stronger, in line with the secondary literature (Klinkhammer, 2016). By contrast, duration of fronts and macro-war trends played a more limited role in shaping soldiers' attitude towards civilians.

The rest of this paper is organized as follows. Section 1.1 surveys the literature related to this study. Section 2 provides historical detail on our setting. Section 3 introduces the conceptual framework for our empirical investigation. Section 4 describes our original dataset. Section 5 describes the main estimating equation and the identification strategy. Sections 6 and 7 present the main results and the analysis of the mechanisms, respectively. Section 8 studies heterogeneity. Section 9 concludes.

1.1 Contribution to the Literature

Our paper relates to several strands of the literature.

First, it contributes to the literature studying the drivers of violence against civilians in armed conflicts.

Previous works have emphasized the role of preferences for violence as a key factor in explaining civilian victimization. In particular, Guarnieri (2023) and Guarnieri (2025) identify the cultural distance between victims, perpetrators, and armed groups as a determinant of violence both across and within conflicts. Other contributions have linked returns to civilian victimization (typically at the macro-level), mostly focusing on the strategic and economic incentives to kill of military or paramilitary units (Azam, 2006; Kalyvas, 2006; Blattman and Miguel, 2010; Esteban et al., 2015). Scholars have also emphasized the role of orders and incentive structures (Milgram, 1974; Browning, 1993; Acemoglu et al., 2020), and identified the economic relationships across ethnic groups (Grosfeld et al., 2020; Becker et al., 2022) and of state capacity (Heldring, 2021) as predictors of the severity of civilian victimization. In the same vein, studies on the drivers of genocides stress the importance of the underlying incentives structure at the macro-level in driving restraint and pursuit of violence by institutional actors (Midlarsky, 2005; Kathman and Wood, 2011). While analyzing the behavior of different combat groups within the same conflict, scholars have shown that the patterns of violence are correlated with soldiers' response to new information (Iyengar and Monten, 2022), the organizational characteristics of the fighting units (Humphreys and Weinstein, 2006; Weinstein, 2007), the role of political and ethnic cleavages (Balcells, 2010; Horowitz and Ye, 2013), and attempts to limit insurgencies and undercut guerilla fighters civilian basis (Valentino et al., 2004). Our work complements these findings by proposing economic incentives and , in particular, soldiers' accountability as a key, and often neglected, driver of civilian victimization in war settings. By focusing on small-scale killings, perpetrated without the consent of the army's leadership, we echoe studies on terrorism and civil wars that document how lack of leadership and discipline leads to indiscriminate violence (Leiby, 2009; Abrahms and Potter, 2015; Oppenheim and Weintraub, 2017). Our work is also in line with findings on the role of law in mitigating the consequences of perverse incentives in relation to victimization of civilians (Acemoglu et al., 2020).

Second, our paper relates to studies on behavioral responses to changes in economic incentives. In war contexts, public recognition and awards can increase risk-taking and gallantry actions (Ager et al., 2021, Carrozzi et al., 2024), while social capital and effective leadership are paramount for securing soldiers' loyalty and reducing desertion rates (Costa and Kahn, 2003; Dippel and Ferrara, 2025). Social pressure matters also in non-war settings, by affecting a range of behaviours, from charitable donations (DellaVigna et al., 2012), to campaign contributions (Perez-Truglia and Cruces, 2017), to voting decisions (DellaVigna et al., 2017). Compared to these studies, we focus on soldiers' *misbehavior*, documenting that armed groups may commit more horrific actions as their accountability drops and that less expert and trained units are more likely to respond to such behavioral incentives. To the best of our knowledge, this

is the first paper to test the role of accountability in a war setting, as most of the existing evidence relies on a theoretical or lab-experiment approach (see Abeler et al., 2019). Our work is in line with the literature showing the role of punishment in leading to "good" behavior in crime (Becker, 1968; Becker and Landes, 1974) and experimental economics (Fehr and Gächter, 2000, 2002). More broadly, our findings also relate to the long-standing debate in the social sciences that discusses the extent to which "good" behavior towards fellow humans stems from legally binding constraints created by a social contract (Hobbes, 1651; Rousseau, 1755; Olson, 2000), rather than from intrinsic preferences (Abeler et al., 2019; Gächter and Schulz, 2016).

Third, we contribute to the economic history and historical literature that analyzes the effects of Nazi occupation in Italy during World War II. Cannella et al. (2024) show that municipalities in the Nazi operational zones, which experienced more persecutions and violence against civilians during the war, developed an higher extent of government distrust in the war aftermath. Fontana et al. (2022) find that in areas where the German occupation was both longer and harsher parties related to the Resistance movement gained political votes until the late 1980s. In a paper closely related to ours, Costalli et al. (2020) suggest that civilian victimization in WWII Italy was related to vulnerability of German armies during retreat after the break of two major front lines, the Gothic and the Gustav Lines. By contrast, our paper focuses on the role of soldiers' accountability in determining their misbehavior, when Axis troops were actively fighting on front lines rather than retreating. In doing so, our results also shed new light on Nazi violence in Italy. While the historical literature has proven that German Command ordered or at least were aware of the most famous mass massacres (Collotti, 1963b; Klinkhammer, 2006; Gentile, 2015; Von Lingen, 2009), we argue that the individual responsibility and variation in accountability for soldiers mattered for small-scale episodes of indiscriminate violence.

2 Historical Background

2.1 World War II in Italy

World War II began in Europe on September 1, 1939, when Nazi Germany invaded Poland. Great Britain and France responded by declaring war on Germany two days later. Despite being formally in an alliance with Germany, Italy remained neutral until June 10, 1940, when it joined the conflict on the German side and declared war on France and Great Britain. During the first two years of war, Italy faced a series of military defeats, especially on the East and the African fronts, that depressed population morale and

weakened the Fascist regime. The situation escalated on July 10 1943 with a combined US and British landing near Siracusa in Sicily, code name Operation Husky. On July 25, 1943, after six days of intense Allied bombing on Rome, Italian dictator Benito Mussolini was forced to resign and imprisoned. The newly-formed Italian government, led by General Pietro Badoglio, officially declared that the war would continue alongside Germany, whose troops - not trusting Badoglio's intentions - took control of most peninsula and disarmed Italian soldiers. Badoglio indeed began secret negotiations with the Allies, which led to the signature of the Armistice of Cassibile on September 3. The Armistice, made public on September 8, opened a new active war front on the Italian territory between the German and Allied armies that lasted until May 2, 1945, known as Italian Campaign.

The Armistice also determined the breakout of a Civil War. On one side, the Italian Fascists remained pro-German and created the *Repubblica Sociale Italiana* (Italian Social Republic, RSI) in Saló, with Mussolini installed as leader after he was rescued from imprisonment by German paratroopers. On the opposite side, the anti-Fascist Partisans, under the control of the *Comitato di Liberazione Nazionale* (Committee of National Liberation, CLN), carried out guerrilla attacks against German troops and aided the Allied advance.

2.2 The Italian Campaign

The opening of a new front in Italy created substantial military problems for Nazi Germany, that was already under strong pressure on the Eastern front, and increasingly so on the Western front after the Allied landing in Normandy in June 1944. Since Hitler did not believe that defending the peninsula was essential to achieve victory in the war, troops and resources available for the Italian Campaign were limited (Kesselring, 1954). Therefore, the German Command, led in Italy Field-Marshal Albert Kesselring, devised a so-called "fighting withdrawal" strategy (Ronchetti, 2018). This aimed to slow down the Allied advance by fighting along pre-set defense lines, sequentially built based on a strict schedule. German troops were ordered to "contend in the bitterest way every inch of land" (Ronchetti and Ferrara, 2014) on a given defense line and to withdraw only when the next line was ready, destroying everything that could further delay the Allies. Such tactics not only caused a large number of fatalities to the enemies, but also bought time for German troops not at the front to build more powerful defense lines along the peninsula.

An important part of the "fighting withdrawal" strategy was choosing the location of the defensive line. German commanders typically decided to build such lines in easily defensible sites, exploiting the natural features of the soil, such as mountains, rivers, and caves (Short and Taylor, 2006). Front lines were

not uniformly fortified: for instance, more vulnerable parts, such as coastal plains and mountain passes featured more elaborate defenses, such as antitank ditches, concrete bunkers, tank turrets, and minefields. Elsewhere, defense lines were simpler: walkways and trenches of dry-stacked stone camouflaged in the soil and stone parapets. In protected areas with good visibility, positions for artillery were prepared. Moreover, in the ten kilometers preceding the lines, communication and electricity lines were dismantled (Ronchetti, 2018). These defensive lines would be activated as front lines when fighting started alongside them, as the Allied vanguards reached them. In the rest of the paper, we will refer to these situations as " front line activation". Generally, by the time of activation, German troops had fully fallen back to the new defensive position, with the withdrawal from the previous line usually starting between fifteen and five days before activation (Short and Taylor, 2006).

Despite the meticulous organization of the defense lines, German commanders often had to deviate from their original plans to adjust to the situation on the ground. For instance, it was not uncommon that the withdrawal was anticipated or delayed relative to what initially established based on the outcomes of the confrontation of German and Allied armies. Moreover, as ordered withdrawals were sometimes difficult under pressure, commanders were often forced to make a stand in positions different from the previously planned ones (Kesselring, 1954, Kurowski, 2003). Overall, the "fighting withdrawal" strategy proved successful. The Allies were only able to reach the Gothic Line in late 1944, more than one year after the landing in Sicily. This final defensive line was eventually broken in April 1945, only a few days before the official end of the war in Italy, on May 2, 1945.

2.3 German Military Occupation in Italy

With the signature of the Armistice of Cassibile, despite Nazi Germany officially recognizing the authority of the Italian Social Republic as *de jure* government, the country was placed under *de facto* military occupation and administration of the Wehrmacht, the German Armed Forces. Specifically, on September 12, 1943 Kesselring declared Italy as "war territory", therefore subject to the Third Reich war laws (Collotti, 1963b, p.95).

The Italian territory was divided into two areas, whose borders were constantly changing as the front line moved: the occupied territory, that included all the areas not directly exposed to the fighting, and the operation zone, that included the territories where military operations were taking place, typically the front and its rear, as well as coastal areas with a depth of roughly 30-40km (Gentile, 2015). The extension of the operation zone, defined by the German Command separately for each front line, covered an area

of approximately 250km behind the front (Gentile, 2015). It was further divided into a combat zone of around 40km behind the front, where German troops fought against the Allied forces and where offensives, counteroffensives, shelling and bombing were concentrated; and the rearguard, where the reserve divisions, the logistic services for troop supplies and the medical units were located (Collotti, 1963b, p.120; Gentile, 2015).

In the occupied territory, the military power was assigned to a Plenipotentiary General, who was responsible for administrative duties and for managing the war economy and the civilian manpower to ensure "the most unified exploitation of Italy for the war conduct" (Collotti, 1963b, p.120 and 224). The German uniformed police (Ordnungpolizei, OrPo) was in charge of dealing with partisan attacks and manpower strikes, and maintaining the discipline of German troops (Collotti, 1963b, Gentile, 2015). By contrast, in the operation zones, Kesselring had the supreme power on all military units and civilians, with division commanders in charge of warfare, counterinsurgency operations, regular policing and maintaining German troops' discipline (Gentile, 2012, p.100). Each commander was responsible for supervising the area around the division's headquarters (Collotti, 1963b, p.100).

3 Conceptual Framework

3.1 A Principal-Agent Problem in a Conventional War Setting

We think of our setting in the framework of a simple principal-agent model, where the military is an internal labor market (Rosen, 1992; Costa, 2014). The principal – the army commander – assigns tasks to the agents – the soldiers - that require a certain amount of effort E. Each soldier makes a decision on the amount of effort dedicated to following orders. For instance, effort can be applied to restrain interactions with civilians away from indiscriminate behavior, among many potential tasks. Soldiers receive a fixed wage W for implementing the commander's directives. As wages are typically fixed in military settings so that monetary incentives are lacking, the principal insures the required amount of effort is applied through monitoring and punishment. He pays a cost M to monitor effort.

The soldiers always receive wage W and exert a certain level of effort E, which costs them c(E). If they are caught not exerting enough effort, they incur in a disutility cost of punishment, normalized to -1.[3] The payoff for soldiers is thus $U_A = W - c(E) - q(E, M, I)$. q is the probability of being caught while not exerting the required amount of effort, which we refer to as "accountability". This is negative in effort,

[3]To simplify, we think of this as fixed and large, in line with the fact that capital punishment was often used to sanction failure to follow orders.

and positive in M. I is a productivity parameter that represents information available to the principal on the effort of the agent. The smaller I the smaller q for the same level of M and E. The principal's payoff is simply the soldiers' effort minus the monitoring cost: $U_P = E - M$ and can be thought of as proportional to the probability of winning the war. The first insight captured by this simple framework is the nature of the misalignment in incentives between the principal and the agents. While the principal maximizes utility through soldiers' effort by incurring M, soldiers minimize costly effort.

The equilibrium effort E^* is chosen by the soldier to maximize U_A. Intuitively, a reduction in the information available to the principal would result in the agent moving to a lower effort equilibrium, due to a drop in the probability of being caught. This intuition can be visualized in the example provided in Figure A.3 in the Appendix, which shows this for $c(E) = E^2$ and $q(E, M, I) = \frac{I*M}{E}$. With E on the y-axis and M on the x-axis, the equilibrium point of effort E^* is at the intersection of the agent and principal's best response (BR) curves. A drop in the monitoring productivity I due to reduced information leads to a downward rotation of BRA, resulting in a shift to lower effort equilibrium E'. This drop is due to a reduction in the probability of being caught, which results in a lower equilibrium level of effort.

In our setting, the army commander is in favor of a *selective* use of violence that targets specific civilians and is proportionate to precise war objectives (Kalyvas, 2006). Such restraint is costly for soldiers. Soldiers may instead prefer the application of indiscriminate violence against civilians, as this reduces E. A reduction in effort in following rules of engagement with civilians would thus result in more indiscriminate violence. One extension of this framework would allow for non-monetary benefits B that are negative in effort. We can think of these as increased security for the soldier if he defies orders (i.e. fighting less bravely). In the case of engagement with civilians, indiscriminate violence may boost security or facilitate looting. It is useful to note how, in this case, the lower equilibrium point E' could also be consistent with an increase in B.

3.2 Application of the Principal-Agent Problem to the Italian Campaign

The principal-agent framework presented above encapsulates the problem faced by the Germany Army during the Italian Campaign in relation to engagement with civilians. The German Command, led by Field-Marshal Albert Kesselring, was generally opposed to the use of indiscriminate violence against unarmed civilians, as he believed that an excessive civilian victimization would not only "create distrust in the German Army and help the enemies' propaganda" (Gentile, 2012, p.156) but also "have destabilizing effects on the internal structure of German troops" Gentile, 2012, p.150).[4] German troops had to be tough

[4]Kesselring's preferences are common in war contexts, as discussed by Kalyvas (2006).

with partisan fighters and saboteurs, but were required to exert an effort in moderation towards Italian civilians. Soldiers were monitored through the command chain and the military police, and were held accountable for the use of excessive violence. Based on the rules of engagement, those caught in flagrant "unjustified violence against the civil population" could be executed without trial (Gentile, 2012, p.155). Even suspected partisans and collaborators of the resistance had to be formally court-martialed and not executed without a legal, albeit quick, trial.[5] The Atlas of the Nazi and Fascist Massacres in Italy, our main source of data, indicates that these formal enforcement mechanisms were in place even in areas relatively close to the front-line. Suspected partisans were often court-martialed according to the rules:

> "[...] After the arrest the four men faced trial immediately at the Military tribunal that was already operational in Sulmona; their legal defense was carried out by lawyer Sigismondo Gravina. The trial ended with a death sentence for the four men due to the carrying of forbidden weapons. The sentence was immediately executed via shooting at 8 AM on 20 October 1943 by the fence of Sulmona's cemetery. According to local witnesses, shooting happened after the victims were forced to dig a ditch for their burial. Since the conclusion of the trial, the four men were assisted spiritually by Don Salvatore Ficorilli, priest at the local hospital, who stayed with them until their death."[6] *Source: http://www.straginazifasciste.it*, Sulmona (L'Aquila), October 20, 1943.

Although systematic data on the enforcement of Kesselring's directives are not available, anecdotal evidence indicates how German soldiers were actually monitored and punished for breaking the rules of engagement with innocent civilians:

> "[...] Until a German patrol intervened and, unable to disperse the mob, started to shoot and even to throw grenades against those poor desperate people. A child, Ivo Coccia, died immediately, while being transported to the hospital; a young woman, severely wounded, died a few hours later after having being transported to the hospital on a chair. Other than leaving an orphaned child, she was also pregnant. [...] Some witnesses reported how, a few days later, the Germans executed at Sulmona's cemetery a soldier that was considered responsible for the

[5] German authorities in Italy were relatively light-touched outside the operations zone (Gentile, 2015, pp.67-8).

[6] Authors' translation from: "[...] Dopo l'arresto i quattro furono immediatamente processati dal Tribunale militare penale tedesco che era già attivo in Sulmona; la loro difesa fu assunta dall'avv. Sigismondo Gravina. Il processo si concluse con la condanna a morte dei quattro per possesso di armi proibite. La sentenza fu immediatamente "ratificata" ed eseguita, tramite fucilazione, alle 8 del mattino del 20 Ottobre 1943 a ridosso del muro di cinta del Cimitero monumentale di Sulmona, dopo che, secondo alcune testimonianze, le vittime furono costrette anche a scavarsi la fossa comune. Fin dal momento della conclusione del processo i condannati furono assistiti spiritualmente da don Salvatore Ficorilli, Cappellano dell'Ospedale civile della SS.ma Annunziata di Sulmona, che resterà loro accanto sino alla morte."

mass killing."[7] *Source: http://www.straginazifasciste.it*, Sulmona (L'Aquila), October 11, 1943.

This monitoring mechanism guaranteed a certain moderation in the use of violence during the war (Collotti, 1963b). However, limits to the monitoring capacity of the principal were frequent and ubiquitous, in particular upon the activation of a new front. In fact, when a new front was activated, German divisions were redeployed and their headquarters moved. This strained the command chain, leading to a quintessential "fog of war" (Clausewitz, 2003). We think of these episodes as exogenous shocks to the information of the principal (I) in our framework, or in other words monitoring productivity. In such situations, even if the German Command's preferences towards violence did not change, soldiers could have increased indiscriminate violence due to a drop in the monitoring capacity of the principal. In other words, the activation of a new front line was a particularly severe shock to soldiers' accountability as the risk of being caught while exerting low effort (q) dropped, potentially leading to an increase in indiscriminate violence.[8] These shocks were temporary: when the front stabilized, accountability was quickly restored as the command chain returned fully operational. Oral testimonies confirm that this situation of chaos lasted a few days after the front activation and created a "no-man's land" within the new combat zones in which soldiers knew to have a limited accountability and were likely to commit severe crimes against the civilian population (Gentile, 2015, p.225; Klinkhammer, 2006).

4 Data

We assembled a new dataset, collected and digitized from several historical and administrative sources. In this section, we describe the data we collected and present key summary statistics. Additional details on the data collection and data sources are available in our Data Appendix.

4.1 Civilian Killings

We collected data on the Axis civilian killings from the *Atlante delle Stragi Naziste e Fasciste in Italia* (Atlas of the Nazi and Fascist Massacres in Italy).[9] The database reports "all the massacres and the individual

[7] Authors' translation from: "Fino a quando non intervenne una pattuglia tedesca che, vista vana ogni minaccia, cominciò a sparare e, addirittura, a lanciare alcune bombe a mano contro quella povera gente disperata. Un bambino, Ivo Coccia, morì quasi subito, vanamente trasportato all'ospedale; una giovane donna, ferita gravemente e trasportata in ospedale su di una sedia, morì qualche ora dopo. Da alcune testimonianze si apprese che la donna, oltre a lasciare orfana una bimba, era anche incinta. [...] Vi furono, infine, anche altri feriti. Altre testimonianze vorrebbero che, qualche giorno dopo, i Tedeschi fucilassero al Cimitero di Sulmona un loro commilitone ritenuto improvvido autore di questa strage."

[8] Soldiers may have also experienced an increase in the returns from violence (B), if activation shocked both I and B (for instance by making soldiers less secure). See Section 3.1

[9] We access this database, available at http://www.straginazifasciste.it, in November 2020.

murders of civilians and resistance fighters killed in Italy both by German soldiers and soldiers of the Italian Social Republic outside of the armed fights" between July 12, 1943 and May 2, 1945. It focuses only on targeted civilian killings and excludes collateral killings, which were committed unintentionally as a side effect of the fighting. For each killing episode, the Atlas reports its location and date, the number of victims, their gender, age and fighting status (whether civilians or partisans), the killing mode, and whether the perpetrators were Germans, Italian fascist troops, or both acting together.

The Atlas' data indicates that, between July 1943 and May 1945, 24,988 Italian civilians were killed by Axis soldiers. In the broader war context, this number is almost equal to all Italian civilians who died in the first three years of WWII and corresponds to 20.8% of civilian victims during the Italian Campaign (ISTAT, 1957). Since our theory focuses on the principal-agent interaction between German command and soldiers, we restrict our sample to killings perpetrated by German soldiers, either alone or in collaboration with fascist soldiers, that amount to 73% of all the episodes in the Atlas.[10] We collapse these data at municipality-day panel, obtaining a total of 3,892 episodes of civilian victimization.

While each episode counted an average 5.5 victims, the vast majority of them concerned either single killings (45% percent) or small collective killings between 2 and 10 victims (45 percent, Figure A.1, Panel A). In just 18 episodes (0.34 percent) more than 100 civilians were murdered. As clearly shown by these stylized facts, such large massacres, that were perpetrated with the more or less explicit consent of the German High Command, were the exception rather than the norm in the Italian context (see Von Lingen, 2009 for a summary of the literature on large massacres).

A natural question is whether the Atlas data capture the universe of targeted killings throughout the Campaign. While it is not possible to gauge the exact coverage of the Atlas, several elements suggest it is at least close to it. First, the data collection was performed by a team of German and Italian researchers, supported by partisan associations, who conducted a large-scale search across national and local sources using uniform criteria with the explicit intention of reporting all such episodes. While the stated objective of the project reassures against selective collection of episodes, the diversity of the team also limits concerns about intrinsic reporting bias, for instance with respect to the identity of the perpetrators. Second, we believe coeval reports on which this research effort was based to be comprehensive. On the one hand, the Italian police was still active under German occupation, which led to the compilation of several accurate reports, that were later complemented by post-war enquiries by the Italian state. On the other, relatives of the deceased victims had strong monetary incentives to report each crime both during and af-

[10]The remaining episodes were either perpetrated by Fascist units alone (22.7%), or the identity of the perpetrators could not be determined (3.9%). Summary statistics that include all the episodes are described in Appendix B.1. The inclusion of killings perpetrated by Fascist units does not change the interpretation of our findings (see Section A.2.1 and Table A.20).

ter the war, due to the granting of war pensions for the relatives of war victims. Notably, this pension was given for killings perpetrated by both Axis and Allied soldiers, thus reducing the incentives to attribute to German soldiers Allied killings.[11] The fact that roughly 50% of the episodes in the Atlas involved only one victim is consistent with a high degree of reporting and suggests that data are not likely to be truncated due to selective coverage by episode size.[12]

Indiscriminate and Selective Violence. We use these data to create several indicators that distinguish between episodes of indiscriminate violence against civilians from those of selective type. To capture this distinction, we start building on an established literature in political science, pioneered by Kalyvas (2006). This maintains that the difference between these two types of violence derives from the objective function of the perpetrators. Selective violence is proportionate to the strategic goals of the army and targets victims at the individual level, as the perpetrators attempt to "ascertain individual guilt" (Kalyvas, 2006, p.142). Conversely, indiscriminate violence goes beyond the war objectives, has a "disproportionate", "random" component and generally targets multiple victims, often in small groups.

Given the impossibility of measuring the intentions of the perpetrators, we proxy selective violence with single-killing episodes, while indiscriminate violence with collective ones (i.e. two victims or more). In Appendix B.2, we show that this approximation captures this distinction well, leveraging several examples from the Atlas. Based on Kalyvas (2006)'s definition, we document how, in our sample, episodes of selective and indiscriminate violence show a similar geographical distribution (Figure 1). The two types occurred with a similar frequency, as roughly 45% of the episodes involved a single killing and 56% a collective one (Figure A.2).

We then exploit additional information provided by the Atlas to refine these measures further. First, we use the detailed episodes' descriptions provided to identify killings that may have been related to counter-insurgency and guerilla operations (i.e. containing a word for "guerrilla fighter" (*partigiano*) in their description), and for which the use of violence may have been instrumental to German army's strategic goals. We then create an indicator for collective killings *not* related to partisans activity, that represents 22.2% of the episodes. Second, using information on the victims, we code an indicator for collective killings that involved the death of at least one vulnerable person (either a child, an elderly person or a woman), less likely to be a threat to German soldiers, an instance that occurred in 15.5% of the cases.

[11] Legge 10 agosto 1950, n. 648, art. 10, available at: $http://www.edizionieuropee.it/law/html/41/zn77_06_050.html$, accessed on 20/10/2022

[12] Even in case some killing episodes are not reported in the Atlas, this does not represent an issue for our analysis, as long as it is non-systematic and does not differentially affect treated and comparison municipalities. We provide a more complete discussion of this aspect in Section 5.3.

Finally, we read the detailed descriptions available for roughly 86% of the Atlas' episodes to code indicators based on whether they describe indiscriminate or selective violence. In the rest of the paper, we will refer to these two indicators as "hand-coded" measures.[13] Notably, also hand-coded indicators show a similar incidence of indiscriminate and selective violence, which account for 45 and 55% of all massacres respectively.[14]

Overall, restraint towards civilians is implicitly confirmed by our data. The probability of observing episodes of indiscriminate violence is low, with an incidence of 0.06% according to Kalyvas (2006)'s-based measure and 0.04% according to our hand-coded measure (Table 1). In other words, less than 1% of day-municipalities experienced an episode of indiscriminate violence through the campaign. Single killings occur with a probability of 0.05% for both measures. Killing episodes not related to partisan activity or that include violence against the vulnerable population are even less common, ranging between 0.02 and 0.03%.

4.2 Location of the Front lines and Combat Zone

We geo-localized all major front lines along which the German and the Allied troops fought from the invasion of Sicily in July 9, 1945 to the end of the war on May 2, 1945. To establish their exact locations, we rely on reconstructed maps collected from the secondary literature (Blumenson, 1993; Short and Taylor, 2006). We rely on the same body of work to establish with precision the exact timing of the activation of each front-line, defined as the day of the inception of fighting along a new front line. German and Allied troops confronted over 32 major defensive lines during the Italian Campaign. The first defense line in Western Sicily was active between July 11 and 15, 1943, right after the Allied landing. The last defense line, the Po Line, in North-East Italy was active from April 23 to May 2, 1945 when Nazi Germany surrendered (Figure 2). On average, each front line lasted 18 days. The two longest-lived lines were the Gustav Line, that, despite some adjustments, lasted 164 days from November 3, 1943 to May 14, 1944; and the Gothic Line and the related lines over the Apennine mountains that lasted 212 days, from August 30, 1944 to April 9, 1945.

We use information on these defensive lines, as well as the timing of their activation, to define, on the Axis side for each front, the position of the combat zone throughout the campaign, namely the area where bombing, shelling and Allied offensive would typically take place. As this area was not formally defined

[13] See Data Appendix B.2 for details on this procedure.

[14] We acknowledge that our proxies for indiscriminate and selective violence could be seen as extreme examples in the application of each type of violence, and by no means exclusive. Violence against civilians took many forms, including different type of physical, sexual, physiological and economic violence. While we are unable to observe and measure such episodes, we believe tracing violence with more granularity would lead to similar conclusions.

and changed depending on the front, we follow the secondary literature (Collotti, 1963b; Ronchetti and Ferrara, 2014; Klinkhammer, 2016) and use a 40 km distance cut-off from the front.[15] We use information on the activation date of each front line and respective combat zones as a shock to the principal's monitoring capacity and, thus, accountability.

4.3 Location and Characteristics of German Units

We retrieved information on the location and identity of German military units during the Italian Campaign from the maps prepared by the Wehrmacht High Command (OKM), stored at the German National Archives (Bundes Archiv). These were produced with a frequency spanning between 1 and 5 days, and thus reflected the situation on the ground quite precisely. We reconstruct the distribution and identity of German troops both before and after our episodes of front line activation as precisely as possible, based on the availability of military maps. More specifically, we are able to match 29 of our 32 episodes with a map from the day of activation or the day(s) immediately after. We also match 21 episodes with an additional map showing troops distribution immediately before activation.[16]

We use these maps to geo-localize the position of the headquarters of German army groups, armies, divisions, as well as detached regiments and battalions across municipalities. These data allow us to trace German troops movements across Italy, both at the front and in the rear, with precision. Moreover, these sources provide some key time-variant information about the state of military units. For instance, they report whether a division was composed of sub-units (typically regiments) that were not usually attached to that division. We refer to these as "composite" units. Such units were more likely to be created and deployed in periods of military crisis, and presented a management challenge, as they altered the usual command chain by creating temporary relationships between superior and subordinated commanders that did not usually work together. We also know whether a unit was being refurbished at any given time, or was temporarily detached from its main body.

We complement this information by collecting additional, time-invariant unit characteristics from a series of military compendia (*Orders of Battle*), such as the complete division name, the date and location of its creation, and previous theater of operations (if any).[17] Our sources show that 45 Axis divisions operated during the Italian Campaign for a total of around 450,000 soldiers. 57 percent of divisions were

[15] As 40km is an approximation (See Collotti, 1963b, p.120), in the empirical part of the paper we test sensitivity our results to alternative cut-offs (Tables A.11 and A.12).

[16] The Data Appendix B.4 provides additional information on the matching procedure.

[17] We collect information on the characteristics and history of German divisions through various compendia, but mainly from https://www.lexikon-der-wehrmacht.de/ and https://forum.axishistory.com/, last accessed on 31/5/2022. See Data Appendix B.4 for details.

infantry divisions, 34 percent Panzer (tank) or motorized infantry divisions, and the remaining ones paratroopers. Only 2 divisions were Italian, formed by soldiers loyal to the Republic of Salo'.

4.4 Italian Partisan Activity and Fighting Intensity

Other than capturing German soldiers' exposure to insurgency activity from the description of the killing episodes in the Atlas (see section 4.1), we also collect information on the location of partisan bands at the beginning of the conflict separately from Baldissara (2000).

Next, we collect data on Allied bombings to measure the intensity of the Allied attacks, that we use as proxy for German soldiers' fighting-related stress. Other than being a source of stress by itself, Allied bombing during the Italian Campaign was tactical: as explained in Bianchi and Giorcelli (2023), it had the goal of offering support to the Allied ground operations against the German troops and usually anticipated massive attacks. We retrieved such data from the Theater History of Operations Reports (T.H.O.R.; Lt Col Robertson et al., 2013), compiled by the US Air Force Research Institute, that reports location, date, type of target and tons of explosives of each Allied air strike during WWII.[18]

Finally, we use German military maps (see section 4.3) to compute the distance between a municipality and the closest Allied military unit, which allows us to trace exposure to Allied ground offensives.

4.5 1936 Population Census and Municipality Shapefile

We digitized the 1936 Population Census, that provides socio-economic characteristics of the 7,341 Italian municipalities. Crucially for our analysis, the Census reports maps with Italian administrative boundaries, that we used to create a novel shapefile at the municipality-level, our basic unit of analysis.

5 Empirical Strategy

Our empirical strategy relies on plausibly exogenous movements of the front line negatively shocking German soldiers' accountability in their engagement with civilians. This shock was due to a temporary drop in the monitoring capacity of German Command that experienced a reduction in the information available on the actions of the soldiers.

Estimating the effects of the front activation on civilian victimization presents two empirical challenges. First, the war front moved almost monotonically from South to North during the Italian Cam-

[18] We access this database,available at www.afri.au.af.mil/thor, in January 2017.

paign, as the Allies advanced. We only observe Axis crimes. By construction therefore, areas south of the front – those under the Allied control – cannot experience killings, thus leading to changes to the relevant comparison group and estimation sample alongside front movements. In this setting, a standard two-way fixed effect model could not take these mechanical changes into account, thus leading to potentially biased estimates. Second, the staggered nature of our treatment would include comparisons between early and late treated units, which may lead to invalid estimates due to OLS weights and heterogeneous effects (Goodman-Bacon, 2021).

To address these issues, we estimate a stacked difference-in-differences specification. This approach consists in constructing separate panels around each treatment, in our case front activation, that are vertically concatenated to form a stacked panel. The stacked panel allows us to estimate our effects only on the Axis side of the front and, by treating each activation episode as if in separate panels of equal weight, mitigates the problematic comparison between early and late treated units (Baker et al., 2022).

More specifically, for each war day, we construct t artificial panels across all Italian municipalities m, spanning between 5 days (k) before and after t, which we then stack together.[19] Across all episodes of front line activation, we compare municipalities that fell within the combat zone at time t but were outside it at time $t-1$ (treated municipalities) with municipalities in Axis territory, that either remained outside the combat zone at both times t and $t-1$, or that were inside it both at t and $t-1$ (comparison municipalities). Figure 3 provides a graphical intuition of the estimation we run for each episode of front activation. Formally, we estimate the following stacked panel difference-in-differences model:

$$\text{outcome}_{m,t,k} = \sum_{k=-5}^{5} \beta_k \text{Treated}_{m,t} \cdot \text{Day}_{t,k} + \sum_{k=-5}^{5} \alpha_k \text{Day}_{t,k} + \Gamma + \Theta + \epsilon_{m,t,k} \tag{1}$$

where $\text{outcome}_{m,t,k}$ is one of the key indicators of violence against civilians, observed in municipality m, day of front activation t (which we refer to as "experiment"), and day-relative-to-activation k. More specifically, we measure indiscriminate violence with: an indicator equal to one if a municipality experienced an episode of collective killings, where at least two people were killed by German soldiers; an indicator for collective killings unrelated to insurgency activity; an indicator for collective killings where at least a vulnerable (woman, child or elderly person) individual was killed; and a hand-coded indicator for episodes of indiscriminate killings based on the Atlas description. We measure selective violence with

[19] We restrict the baseline sample to those artificial panels in which the front line moved, which leaves us with 31 front-experiments. We further restrict our estimation to episodes for which we have data on civilian killings, which are only available between July 12, 1943 and May 2, 1945. This leaves us with 28 front-experiments. We limit the number of days per experiment to 5 days k before and after the front movement because a longer number of days would imply cutting at least a front movement, the last of which occurred 5 days before the end of the war. We limit the sample to these 28 experiments for computational purposes. The results do not change if we include all war-day artificial panels (700 in total).

an indicator for single killings, where a single person was killed by German soldiers; and a hand-coded indicator for episodes of selective killings based on the Atlas description.[20] $\text{Treated}_{m,t}$ is an indicator equal to one for municipalities that were not in the combat zone at time $t-1$ and fell into the combat zone at time t. We define the combat zone as the 40km behind the front, based on the secondary literature (Collotti, 1963a, p.120; Ronchetti and Ferrara, 2014; Klinkhammer, 2016), as explained in Section 4.[21]

$\text{Day}_{t,k}$ is a set of fixed effects for days k relative to front activation in day t. Γ is a full set of fixed effects that allows us to estimate our average treatment effect within municipality (γ_m), front-experiment (γ_t) and day relative to front-activation (γ_k). Importantly, γ_t absorbs treatment heterogeneity over time. Θ is a set of fixed effects designed to capture time trends. We account for municipality-experiment ($\theta_{m,t}$) and experiment-day-side of the front line ($\theta_{s,t,k}$) trends. Crucially in our setting, $\theta_{s,t,k}$ is an Allied-side, experiment trend, that allows us to estimate our results within the Axis-controlled side of the front line. This term makes the estimation equivalent to dropping all Allied-controlled municipalities in each experiment, an impossible procedure in a regular panel.

Our linear probability model estimates the reduced-form effect of accountability on soldiers behavior. If, indeed, movements of the front lines negatively shocked soldiers accountability we expect β_k to show a positive treatment effect on indiscriminate violence, but not on selective violence. β_k estimates the causal effect of front activation on indiscriminate violence based on three key assumptions, namely parallel trends, orthogonality of front movements with respect to intention to kill, and no differential reporting of episodes. In the rest of this section, we discuss the validity of these assumptions.

5.1 Parallel Trends

The crucial identification assumption for β_k estimates to be causal is that, in absence of the front activation, treated and comparison municipalities would have followed the same trend in indiscriminate and selective violence. To formally test this assumption, we study $k-5$ pre-trends across all event graphs in the paper and we report $k-2$ coefficient in all tables. Additionally, to directly test for the overall significance of the pre-trends across our six measures of indiscriminate and selective violence, we also regress them on a full set of indicators for each of the five days before front activation interacted with the *Treated* indicator. The coefficients estimated on the interaction terms are never statistically different from zero and

[20] These variables only include violence perpetrated by German soldiers, as they were directly exposed to the German High Command and its monitoring capacity. While German soldiers were responsible for more than 70% of the killings, our results are robust to the inclusion of killings by Fascist units (Table A.20). We exclude the Isle of Sardinia as no episodes are reported here throughout the Campaign.

[21] Since the extension of the combat zone is not precisely defined for each front line, in Section 6.1 we show that our results are robust to using 30 and 50 km as alternative cut-offs.

we always fail to reject the null hypothesis that all the interaction terms are jointly equal to zero (Figure A.4, Panels A-F). Overall, treated and comparison municipalities were on the same civilian killing trend in the five days before front activation.

5.2 Orthogonality of Front Movements to Intention to Kill

The parallel trends assumption does not require treatment and comparison units to be statistically equivalent in terms of observables and unobservables. Nevertheless, a threat to our identification may arise from correlated shocks, namely if treated and comparison municipalities differed in terms of time-varying characteristics that simultaneously affected the location of the front and violence against civilians. While mitigated by the staggered nature of our treatment, this possibility is not fully ruled out. For instance, if the Wehrmacht were to decide to systematically set up defensive lines in areas where they intended to kill more civilians, this may pose a threat to our identification.

This concern is partially attenuated by the historical evidence, which indicates that the timing of front lines activation, as well as their exact position, largely depended on geography and on the outcome of the fighting against the Allies in the previous front and was hard to predict even for German commanders (Kesselring, 1954; Kurowski, 2003). As discussed in Section 2.2, German defensive lines were built in easily defensible sites, that normally coincided with rivers and mountain chains and were sparsely populated (Short and Taylor, 2006). If anything, treated locations should have been negatively selected in terms of factors that might facilitate the explosion of violence against civilians. We corroborate this anecdotal evidence by comparing ever and never treated municipalities along their cross-sectional geographical and pre-war, socio-economic characteristics in Table A.2. On average, treated municipalities contained a larger number of watercourses, were more elevated, rugged and relatively close to the coastline (Panel A, columns 1-5). These more remote locations were, logically, sparsely populated, far from provincial capitals, less industrialized, more agricultural and without a particular concentration of partisan bands (Panel B, columns 1 to 5 respectively). If anything, remoteness should have made it harder for the German soldiers to find and kill civilians. It is worth pointing out that equation 1 includes municipality and municipality-experiment fixed effects that control for unobservable underlying differences between municipalities and for municipality-experiment differential trends.

While the contextual evidence is reassuring, time-varying factors - such as the intensity of partisan activity - may still simultaneously influence front line location and civilian killings. To deal with this potential issue more systematically, we compare treated and comparison municipalities in the 20 days before

front activation through our proxies for fighting intensity, namely exposure to Allied bombings and Resistance activity. The balancing test shows how, on average, treated municipalities did not look statistically different than the comparison ones in any of these measures (Table A.1), thus making correlated shocks unlikely. Overall, we find no evidence that treated and comparison municipalities were different before treatment based on observables.

5.3 Differential Reporting Between Treatment and Comparison Units

While non-systematic under-reporting of killing episodes is not a problem for our empirical analysis per se, selection bias may originate if this were differential between treated and comparison municipalities. For instance, if the records on killing episodes came from the German army, one may worry that better data was collected near the front due to larger personnel availability and strategic needs. However, given the nature of our data, we do not believe this is a likely scenario.

The data from the Atlas, in fact, relies either on coeval reports by the Italian police, which had no jurisdiction at the front as Italy was an occupied territory, or on inquires of Italian authorities after the end of the war, as explained in Section 4.1. These features of the data collection would imply more difficulties in recording killing episodes near the front, where Italian police was less likely to be present, especially during the chaotic days of its activation, and would result, if anything, in an under-reporting of violence in our treated municipalities.

Moreover, as explained in Section 4, the relatives of the victims had incentives to report war-related killings, since the Italian government paid war pensions to close relatives of civilians died during the war as a result of violent crimes. This feature made killing episodes very likely to be reported at some point (although possibly after the Germans had left), and not differentially so across treatment and comparison municipalities. Moreover, such pensions were paid for crimes committed both by the Allied and the German soldiers, therefore they did not create incentives to attribute to the Germans Allied violence.

6 The Effect of Front Activation on Civilian Victimization

Consistent with our hypothesis, the activation of a new front-line determined a sharp increase in indiscriminate violence against civilians. The likelihood of observing a collective killing episode perpetrated by German soldiers increased by 0.8 percentage points in treated municipalities relative to the comparison ones on the day of activation of a new front (Table 2, column 1). The effect is sizeable. Compared to a sample's mean probability of 0.06%, this corresponds to a 13-fold increase. This result is consistent

with the hypothesis that economic incentives can indeed influence soldiers' behavior with respect civilian victimization in such an extreme setting. We interpret this finding as consistent with a drop in soldiers' accountability leading to less effort in following the principal's directives and, in turn, to more indiscriminate violence. Reassuringly, results are not driven by large massacres, but rather by small collective killings involving 2 to 10 civilians (Table A.3).

The observed increase in collective killings could, at least in part, reflect a surge in violence associated to new threats faced by the soldiers in the combat zone. For instance, such episodes may have been a perhaps proportionate, albeit illegal, retaliation against civilians in response to partisan attacks, whose intensity could be correlated to the front activation itself. To rule out this possibility, we restrict the outcome variable to collective killing episodes *not* related to partisan resistance.[22] Collective killings unrelated to partisan activity increased by 0.3 percentage points in treated municipalities relative to the comparison ones at front activation (Table 2, column 2), corresponding to a 10-fold increase with respect to the sample mean.[23]. We then restrict collective killings to episodes that involved vulnerable population (at least one woman, child or elderly), which had a clear indiscriminate nature and were likely unrelated to the actual safety and tactical objectives of German troops. This type of collective killings also increased differentially with a similar magnitude (Table 2, column 3).

When plotting all β_k coefficients between 5 days before and 5 after treatment in an event graph, the difference in collective civilian killings between treated and comparison municipalities is small, and not statistically different from zero before activation. After treatment, it sharply increases, and persists for the following three days, only starting to fade away from day 4 (Figure 4, Panels A-C).

The estimated dynamic effects are consistent with the historical evidence describing how at the activation of new front line there was often a delay in the activation of policing services. This lag created two to three days of "no-man's land" where severe crimes against civilians were committed due soldiers' limited accountability (Klinkhammer, 2006; Gentile, 2012). A similar scenario was common also during the occupation of new territories on other fronts in the context of WWII (Aleksievič, 2018). For instance, while describing the Soviet campaign against Germany, Solzhenitsyn (1977, p.386) explains: "When we occupied a new territory, in the first three days we looted and killed. [...] However, after three days, one could be exposed to court-martial punishments for doing this".

So far, we have used a proxy for indiscriminate violence based on collective killings, following an

[22]Specifically, we analyze the descriptive text of each collective killing episode and we exclude those where the word "partigian*" (the way resistance fighters were called in Italy) features in the description.

[23]Table A.19 shows that our results are robust to excluding the front activation episodes in Summer 1944, during which Kesselring allowed more soldiers' violence against civilians when dealing with counterinsurgency activity (see Section 3.2)

established literature in political science (Kalyvas, 2006). However, one may worry this approach may not capture indiscriminate violence well in our setting. We therefore hand-coded episodes types based on the descriptions provided in the Atlas. The results confirm our main findings: upon front-activation, killings described as indiscriminate in the Atlas increased by 0.6 percentage points, a 14.5-fold increase relative to the sample's mean (Table 2, column 4).

Finally, our estimates may capture a general increase in violence of all types linked to renewed fighting, rather than a specific increase in indiscriminate one. To rule out this possibility, we study the response of selective violence to the front activation as a placebo exercise. Selective violence both measured by single killings, following Kalyvas (2006), and by our hand-coded indicator based on the Atlas description did not change differentially between treated and comparison municipalities upon front activation, indirectly confirming our interpretation (Table 2, column 5 and 6).

6.1 Robustness Checks

Our results are robust to a variety of robustness checks that we summarize here. Appendix A.2.1 provides a detailed description of all the robustness checks we perform.

First, we show that our results do not hinge on the definition of treated municipalities as those 40km behind the front line. Using 30 or 50km as alternative cut-offs produces estimates similar in magnitude to the baseline ones (Tables A.11 and A.12).

Second, our results are not driven by large-scale massacres, that may be carried out with knowledge from the high command. In fact, excluding the largest massacres, than involved more than 100 victims – that without doubt required some endorsement from the German Command – or even cases with more than 50 victims, does not affect our estimates (Table A.3, columns 2 and 3).[24] By contrast, our findings appear to be driven by small-scale episodes that involved between 2 and 10 victims (Table A.3, column 4). These numbers imply that in the absence of a principal enforcing accountability through a command chain, indiscriminate civilian killings would be 13 times higher, should agents be left to operate without constraints and based on their own preferences.

A third central challenge is that results may simply reflect a *higher* concentration of German soldiers in treated municipalities relative to the comparison ones due to proximity to the front itself. In other words, the observed differential increase in indiscriminate violence may be due to more soldiers killing with the same probability rather than the same number of soldiers killing more frequently. To rule out

[24]This pattern is consistent with the fact that these large-scale operations required significant planning and resources, and were typically not carried out in proximity of the front line, and surely not in the hectic days around the activation of a new front line (Gentile, 2015).

this possibility, we compare treated and comparison municipalities in Axis territory that were within 100, 50 and 25 km from German division headquarters and therefore exposed to a similar concentration of German troops. Despite a demanding drop of observations from more than 2 million to roughly 136,000 with the more stringent specification, these estimates are fully consistent with our baseline ones. Based on this finding, we conclude that the estimated effects are not driven by a stronger concentration of German soldiers (Tables A.4, A.5 and A.6).

Fourth, we test the sensitivity of our results to restricting the comparison group to municipalities increasingly closer to the front line. Limiting our sample to the operation zone only – which included municipalities directly exposed to the war and under the exclusive control of the Wehrmacht – or to municipalities located within 200 and 100 km from the front line on the Axis side does not change our results (Tables A.13, A.14 and A.15). Notably, restricting the comparison group to only municipalities that were within 40km from the front line both at time t and at time $t-1$ - effectively exploiting the pivoting of the front to compare new switchers with previous switchers so to hold proximity to the front at time t constant - leads to estimates virtually identical to the baseline ones (Table A.16).

Fifth, our results are robust to alternative estimators. First, heterogeneity-robust estimators for staggered treatments, and in particular De Chaisemartin and d'Haultfoeuille (2024) yields results consistent with and very similar to our baseline ones (Figure A.6).[25] Second, we can use a regular panel, rather than a stacked panel, to estimate a standard two-way fixed-effects model, as well as or non-linear models, such as logit, rare event logit and Zero Inflated Poisson estimator to account for the binary nature of our treatment (Table A.24).

Finally, we double-cluster the standard errors at larger geographical units, such as municipality-province, municipality-nearest division and municipality-experiment to confirm the robustness of our inference to spatial autocorrelation and heteroscedasticity within experiments (Tables A.21, A.22, and A.23).

7 Mechanisms and First-Stage in Accountability

Our main results indicate that the activation of a new front line led to more indiscriminate violence against civilians. We interpreted these results as a reduced form effect of decreased monitoring capacity of the principal and therefore of soldiers' accountability. In this section, we provide evidence to corroborate this interpretation. First, we rule out potential alternative mechanisms, namely selective sorting of troops and

[25] This is our favorite, as the only ones that accommodates non-absorbing treatments which is our case.

increased insecurity for the soldiers. Second, we study a first-stage in accountability, that confirms how the monitoring capacity of the principal reduced differentially at the front activation, consistent with a drop in accountability.

7.1 Sorting

A first potential mechanism behind our results relates to the sorting of troops. If upon the activation of a new defensive line, the German command systematically dispatched the more aggressive divisions at the front, as they were more capable to fight, this could also have affected violence against civilians. In this scenario, increased civilian victimization might be due to a higher concentration of 'bad people' at the front, rather than 'regular people' turning bad as behavioral incentives changed.

To rule out sorting as an alternative explanation for our results, we estimate equation 1 using time-invariant troops characteristics, observed before and after the front activation, as dependent variables. We were able to collect this additional information from the German military maps for 19 out of the 32 front movements in our sample.[26] We then use such troops' features to construct two proxies to capture heterogeneity across units.

Our first measure is an indicator for elite divisions (SS, paratroop, panzer and mountain divisions). These divisions – that correspond to 12 units in our sample – were the most combat-effective and motivated, but also the most aggressive towards civilians (Collotti, 1963b; Blumenson, 1993). Elite divisions were not more likely to be deployed in treated municipalities relative to the comparison ones at activation (Table 3, Panel A, column 1). The second measure simply uses the unit number included in each unit's official name. This is based on the fact that units with a lower number formed earlier on in the war and were generally better trained and with more combat experience (Gentile, 2015).[27] We do not find evidence of differential change in experience in treated municipalities relative to the comparison ones.[28] These results suggests lack of selective troops' sorting at front line activation. This degree of "randomness" in units deployment is consistent with the historical evidence, that explains that, due strategic constraints, all units were cyclically deployed in the combat zone (Kesselring, 1954; Collotti, 1963a).

As additional test, in Table A.8, we estimate equation 1 controlling for front-specific trends for the

[26] More specifically, we observe troops characteristics between one to three days before and on the day of front activation. Therefore, and differently from equation 1, we have only one, instead of daily, pre-front movement observation. We provide more details on how we match German military maps with front movements in Appendix B.4. Moreover, since the outcome variables vary at the division level, rather than at the municipality one, we double cluster our standard errors at the municipality-day-nearest-division level.

[27] For instance, the "26th Panzer Division"'s number would be 26.

[28] As a robustness check for comparability, we show that our baseline results remain unchanged if we estimate equation 1 using only the 19 front movements for which troops' characteristics are available (Table A.7).

nearest division at activation. This specification allows us to estimate the treatment effect within clusters of municipalities exposed to the same nearest division trend. The fact that these estimates are very close in magnitude to the baseline ones corroborates the evidence that troops' sorting is not driving our results. Finally, to test this idea more systematically, in Figure A.5, we report results from a linear probability model that predicts the probability for each of the 45 Axis divisions employed in the Italian Campaign to be located at the front. No unit was systematically more likely to be sent to the front, which further confirms the historical narrative. Taken together, these findings indicate that the estimated increase in violence is not driven by different soldiers being deployed at front activation, but rather by the same soldiers responding to behavioral incentives elicited by the war context.

7.2 Insecurity

Upon the activation of a new front line, troops may be exposed to increased stress due to war-related insecurity, which could in turn increase their gains from indiscriminate civilian victimization.[29] Such explanation may be consistent with the observed shift in the violence equilibrium that we document at front line activation. For instance, in a paper related to ours, Costalli et al. (2020) use evidence from the break of the Gustav and the Gothic lines and document an increase in atrocities against population due to growing uncertainty during the German retreat. In their empirical setting, soldiers stand to gain from misbehavior: by adopting a type of "shock-and-awe" strategy, they can better preserve supply lines and facilitate expropriation. While vulnerability of German soldiers clearly played a role on civilian killings throughout the war, it is worth noting that we focus on episodes of front activation, often happening on pre-set defensive lines, rather than on the collapses of previously defended positions. As such, German forces already deployed there should have not experienced a substantial drop in their perceived security relative to the day before front activation.

To test this empirically, we verify whether two proxies for security, namely Allied bombing and distance from the Allied troops, differentially changed at front activation. We show that the probability of observing Allied bombing in a day-municipality does not change differentially (Table 3, Panel A, column 3). Similarly, we document that the distance from the closest Allied unit reported in German military maps does not decrease when a new front line is activated (Table 3, Panel A, column 4). We rationalize these counter-intuitive results based on the fact that the front line was quite dangerous even before the beginning of the Allied ground offensives that define treatment timing. While small portions of it cer-

[29] This would correspond to a lower effort equilibrium which, as discussed in Section 3.1, may arise if soldiers benefits (B) from exerting low effort in following orders increase.

tainly became more dangerous at activation, the military pressure that the average soldier experienced in the combat zone in the very first days after activation was not substantially different from the days immediately before it. By contrast, insecurity at the front line became more and more severe close to the breaking of the line itself (Blumenson, 1993; Short and Taylor, 2006; Costalli et al., 2020), well after our 5-day post-activation window.

Other than the threats from regular warfare against the Allies, partisan activity could represent a distinct and separate source of war-related stress for German soldiers. As we do not observe partisan activity at a municipality-day level, we test for this possibility by implementing an exposure design that splits the sample based on proximity (<15km) to a partisan band's headquarters around September 1943, at the beginning of the occupation of Italy by the Nazis.[30] Concentration of the effect in areas more exposed to partisan presence would point in the direction of increased insecurity for the German soldiers as the main mechanisms. Instead, results from Table A.9 indicate no statistically different effects between the two subsamples. If anything, the estimates suggests that the effect might be stronger away from partisan bands. Taken together, these findings show that our estimated effects are unlikely driven by differential increases in returns to civilian victimization at front activation, linked to increased insecurity for the German soldiers.

7.3 Accountability

We have interpreted the increase in indiscriminate violence upon front activation as the result of the reduced monitoring capacity of the German Command leading to a drop in soldiers' accountability, which pushes agents to exert less effort in following rules of engagement with civilians. In Section 3.1, we have conceptualized this drop in accountability as a reduction in the probability q of being caught while not following orders, in turn due to a reduction in the productivity I of the monitoring cost M incurred by the principal. Our empirical setup exploits exogenous shocks in monitoring productivity, due to changes in the information available to the principal.

So far, however, we have only provided reduced-form evidence of the effect of the shock in information on the agents' behavior, without documenting a first stage in accountability. To provide direct evidence of this link, we estimate a first stage in accountability by leveraging variation in the quality of the command structure. We exploit time-varying characteristics of divisions, in particular the cohesiveness and strength of their command chain at any given point in time. To measure this, and proxy the division commander's *capacity* to effectively police his men and keep order across space and time, we

[30] We implement the exposure design by interacting the binary indicator with all regressors.

rely on information on whether army units were composite ("patchwork"), and we count sub-units. As described in Section 4.3, composite divisions were created in periods of tactical difficulty by attaching one or more sub-units (typically a regiment, at times a battalion) from multiple divisions to a different division command. At times, such units were made from scratch; however this is the exception rather than the norm in our setting. Usually, an existing division would be assigned one or more sub-units for a time, thus becoming composite. As opposed to unitary, cohesive divisions, commanding a composite unit made the task of maintaining order among the ranks more complicated for the division commander, as the new relationship between him and sub-commanders, and the temporary nature of the assignment created accountability problems, slow communication and limited cooperation within the internal command structure of each composite unit (Wray, 1987; Stein, 2010; Gressel, 2015).[31]

Consistent with our accountability hypothesis, treated municipality were differentially more likely to observe a change in the composition of their nearest division at front activation, from unitary to composite (Table 3, Panel B, column 1). The probability of being exposed to a composite unit increases by 13 percentage points, more than twice the sample's mean. We also replicate this exercise, using the count of the sub-units assigned to a composite one. The sample mean shows how most units were unitary. However, in treated locations, the average number of sub-units external to the division increases by roughly 0.2 sub-units, or 20% of the sample's mean at front activation (Table 3, Panel B, column 2). Notably, for both variables we do not observe anticipation effects in the distribution and condition of units before the front activation.

We account for potential sorting of composite divisions into treatment, by repeating the same analysis with the inclusion of experiment-nearest division fixed effects. This allows us to estimate our first-stage within groups of municipalities that shared the same nearest division in a given front-activation experiment. Results remain very similar in magnitude, in turn suggesting that the effect on indiscriminate violence operated through changes to the command structure of the nearest division, rather than a change in division (Table 3, Panel B, columns 3 and 4). This finding is fully consistent with the results on troops' sorting presented in Section 7.1. One additional alternative is that the creation of composite units may have been systematically applied to areas under heavier attack or for units that were suffering more losses, which were more frequent near the front. To account for this possibility, we show that our results remain virtually unchanged if we control for distance from the nearest Allied unit at front activation, a good proxy for exposure to fighting (Table A.10). This exercise strengthens our interpretation that composite units capture a break-down of the command chain that, however, was not systematically related to

[31] See Biddle (2004), for a discussion on the importance of cohesion of military units.

exposure to heavier fighting and, thus, potentially to higher gains from violence. Figure 4 also points in this direction, as the effect disappears within three days after front activation. This is consistent with the historical literature indicating that policing and command lines were restored with a relatively short lag with respect to front activation (Kurowski, 2003, Gentile, 2012).

Based on the evidence presented in this section, we confirm our hypothesis that front line activation led to more indiscriminate violence through reduced cost of misbehavior, namely a drop in soldiers' accountability. To the best of our knowledge this is the first paper to document the full causal chain leading from an information shock (front line activation), to a drop in accountability (the monitoring capacity of the principal), and in turn to a higher indiscriminate violence equilibrium. This is a novel explanation for the occurrence of violence against civilians, where the existing literature has focused on returns from indiscriminate violence (Esteban et al., 2012; Azam, 2006), and intrinsic preferences (Guarnieri, 2025; Abeler et al., 2019).

8 Heterogeneity Analysis

Upon front activation, German soldiers' behavior toward Italian civilians became, on average, significantly more violent. However, not all soldiers misbehaved every time accountability dropped. In fact, despite the estimated 13-fold increase, the overall probability of observing indiscriminate civilian killings remained low, moving from less than one in one-thousands before the front activation to one in one-hundred in treated municipalities. These features of the data beg the question of what underlying characteristics facilitated the adoption of particularly violent practices by German soldiers, when accountability dropped. In this section, we explore three different dimensions of heterogeneity. First, we examine how division characteristics influenced soldiers' behavior at front activation. Second, we consider the pre-war economic characteristics of the municipalities and their influence on soldiers' responses to changes in accountability. Finally, we analyze heterogeneity along macro war trends, to understand whether major events related to the general progression of the war had an influence on the behavior of soldiers at the micro-level.[32]

[32]We implement heterogeneity analysis by interacting all regressors from equation 1 with indicators for the various characteristics we study. This is equivalent to performing a manual sample split, with the advantage of allowing for a direct comparison of the magnitude of the coefficients, as well as the implementation of coefficients' equality tests.

8.1 Division Characteristics

Divisions' identity is often referred to as a major driver of civilian victimization in the context of WWII Italy (Gentile, 2015). As discussed in Section 7.1, we do not find evidence of systematic sorting into treatment. The fact that the identity of the division nearest a given municipality was not an outcome of the shifting of the front allows us to study treatment heterogeneity by division characteristics.

The first dimension that the historical literature mentions as a major determinant of violence is soldier affiliation with the Schutzstaffel (SS, Gentile, 2015). The SS were originally paramilitary units under the direct control of the Nazi party. Towards the end of WWII their number increased, and they began to represent a large share of active combat units. Although formally under the command of the Wehrmacht, these troops were characterized by ideological fanaticism and indoctrination and operated with more independently compared to other army units. These distinctive elements are often quoted as a major driver of civilian victimization in many war contexts (Ahmadov and Hughes, 2017). A few SS units were active in Italy throughout the campaign, most notably the I SS Division Adolph Hitler and the 16th SS Panzergrenedier Division. The latter was famously responsible for some of the most brutal collective killings on record, such as the massacre of Sant'Anna di Stazzema. We test whether the effect of fronts activation on civilian killings was stronger in municipalities whose nearest division belonged to the SS. Perhaps surprisingly, the results suggest that SS divisions did not apply more indiscriminate violence compared to other units when accountability dropped (Table 4, column 1, p-value of 0.9). It is worth noting that this finding does not imply that the SS units killed fewer civilians than other units. However, their killing inclination does not appear to have systematically changed upon front activation.

A second driver of violence against civilians that emerges in the literature relates to the quality and training of units (Oppenheim and Weintraub, 2017). Well-trained soldiers in homogenous units should be better able to cope with war-related stress factors and to operate professionally, even in the absence of a stable command line. We test this hypothesis by examining whether civilian victimization was stronger in municipalities whose closest units were not assembled and trained in Germany. In fact, German-drafted units received more complete training until relatively late in the war. They were also more complete in terms of personnel and generally ethnically and linguistically homogeneous. By contrast, divisions drafted outside of Germany were typically patched-up units made of heterogeneous sub-units from disbanded combat groups who had often suffered significant damage, or local volunteers. They normally received little or no training as a combat unit, were poorly equipped, and often lacked a well-drilled chain of command (Nafziger, 2000). However, also in this case we cannot reject the hypothesis of equality

between the two estimated coefficients for Germany-drafted and other units (Table 4, column 2, with a p-value of 0.41).

Another influential view on the determinants of German soldiers' misbehavior in the context of WWII argues that units deployed to the East against Soviet Union, where racial edges were felt more strongly and the German commands justified systematic violence against civilians more easily, committed more atrocities (Gentile, 2015). When deployed elsewhere, these units would have been more likely to victimize civilians because of experiences from their previous deployment. However, contrary to this hypothesis, we find that our results are largely driven by municipalities whose closest units had not been deployed on the Eastern Front (Table 4, column 3).

Finally, we test whether the combat experience of the units played a major role in determining how soldiers responded to front the accountability shock. The literature, in fact, does highlight combat experience as a major driver of soldiers behavior. More experienced soldiers (and units) should, *ceteris paribus*, be better able to cope with lack of a functioning command chain and with increased fighting stress (Mitchell, 2004; Oppenheim and Weintraub, 2017). We measure unit combat experience by splitting the sample by the sample's median of divisions' creation date (1942). The effects are fully concentrated near younger division, while close to more experienced units the estimated coefficient is small and insignificant (Table 4, column 4, with a p-value on testing for coefficient equality of 0.08).

Based on these results, we conclude that relatively inexperienced German divisions were more likely to apply indiscriminate violence when experiencing a reduction in accountability. This result highlights soldiers experience as a crucial factor in insuring they maintain an appropriate behavior towards civilians irrespective of the underlying level of accountability they face.

8.2 Economic Characteristics

Soldiers' may have responded to a drop in accountability upon front activation differently, based on local monetary incentives related to municipalities' economic characteristics (Grosfeld et al., 2020). In fact, during the Italian Campaign, the Wermacht exploited the economic resources of the Italian peninsula in a systematic manner to support the German war effort. These strategic needs entailed coercing labor and keeping industrial productive capacities at a maximum and, at times, also relocating industrial machinery and workers to the North when the front line moved too close to specific municipalities to continue production (Klinkhammer, 2016). We proceed along this line of inquiry and explore heterogeneity along pre-war municipality economic characteristics.

First, we test if atrocities against civilians upon front activation were more likely to happen in the more developed areas. We find that our results are largely driven by municipalities within 30 km of a provincial capital, where most of the industrial apparatus was generally located, and municipalities whose employment in the industrial sector was above the sample median, although - for the latter - the coefficients are not statistically distinguishable at conventional levels (Table 5, columns 1 and 2).

Next, we focus on heterogeneity along municipality size. The effects of front activation on civilian victimization are larger in more populated municipalities, above the median of total resident population in 1936 (Table 5, column 3), but not in highly densely populated areas, above the median of population density in 1936 (Table 5, column 4). This result suggests that, while the supply of potential victims in larger municipalities increases collective killings, having a more concentrated population on its own did not enhance the severity of civilian victimization.

All together, the heterogeneity analysis along pre-war economic characteristics shows how pre-existing economic conditions matter in creating the pre-conditions for the effect accountability effect to materialize. Stronger effects in populated and industrialized areas, are consistent with a systematic attempt on the part of the German troops to exert a firmer control of strategic locations and speaks to those contributions that link violence to looting (Blattman and Miguel, 2010).

8.3 Macro War Characteristics

Finally, we analyze heterogeneity of civilian victimization along front characteristics and macro war-trends. The characteristics of a specific front may have significantly influenced soldiers' behavior, for instance if they changed the perceived security of the troops deployed there. Moreover, previous works have highlighted how the behavior of the German army varied depending on the period of the war and the tactical characteristics of different front lines. For instance, there is evidence that the general deterioration of the morale of German troops, as the desperate situation of the Third Reich was becoming apparent towards the end of the war, led to more defections and a general increase in the incidence of violations of the military law (Klinkhammer, 2016). In the Italian setting, Costalli et al. (2020) point out how decreasing security for German troops towards the end of the Italian Campaign led to more frequent civilian victimization.

We first explore heterogeneity along front characteristics. We start this analysis by testing if civilian victimization was driven by the activation of short front, which lasted less than the median of the 32 fronts' duration of 18 days. At short-lasting front lines, German soldiers may have felt more insecure

than in more stable and likely better defended positions, and thus exercise less restraint when dealing with civilians. However, the estimated coefficients on collective killings in long fronts are twice as large than those of short fronts, even if such difference is not statistically significant (Table 6, column 1, with a p-value of 0.14).

As the average duration of a front line was the outcome of several factors, and not necessarily related with the actual security on the ground at front activation, as a second step, we look at fronts above and below the median of the standard deviation of terrain ruggedness, calculated across all treated municipalities separately for each front line. The rationale behind this measure is that a high average standard deviation in terrain ruggedness would capture those front lines that benefited from pronounced mountain ridges towering over relatively flat plains, a setting that would facilitate the defenders (Short and Taylor, 2006). Despite the coefficient being positive and significant in municipalities showing high ruggedness variability, the effect is larger in homogenous municipality, thus suggesting a limited role for soldiers' security - as measured by terrain features - in explaining the emergence of the effect (Table 6).

Next, we check whether different macro-conditions linked to the phase of the war had a mitigating or enhancing effect on civilian victimization. We test whether episodes of front activation before the landing of the Allies in Normandy on June 6 1944 show different patterns compared to later front lines. This episode, known as D-Day, had a negative impact on German hopes to obtain a victory in the war and might have consequently triggered a change in the behavior of the average German soldier. While the estimated effects are three times larger after the D-Day, both coefficients are positive and significant, suggesting that frustration may be an important prerequisite for indiscriminate violence, but not a crucial one (Table 6, column 3, difference is significant at 5%). In the same vein, we conclude by comparing front movements in the first and second half of the Italian Campaign, with the idea that soldiers frustration towards the end of the war may have further weakened the monitoring capacity of the German division commanders (Kesselring, 1954). Also in this case, even if we document larger effects in the second half of the war, the impact of front activation on collective killings remains positive and significant during the first half of the war as well (Table 6, column 4).

Taken together, these results suggest that, unlike the level of experience of troops and the underlying economic incentives for looting, front characteristics and macro-war trends had a limited importance in shaping patterns of civilian victimization throughout the Italian Campaign. While frustration towards the end of the war may have played a role in enhancing the severity of soldiers responses to accountability drops, the increase in indiscriminate violence is not strongly related to the underlying (macro) fighting conditions.

9 Conclusions

This paper studies the causal effect of a drop in soldiers' accountability on civilian victimization, using evidence from the Italian Campaign. To perform this analysis, we collected the universe of Nazi civilian killings between June 1943 and May 1945, that we linked with position of front lines and German divisions. We exploited plausibly exogenous variation in the movement of front lines as a negative information shock that diminished the capacity of the German command to police soldiers' misbehavior. We compared municipalities that fell into the combat zone on the day of activation of each front line with comparison municipalities that either remained far from the front line or stayed in the combat zone.

Our results showed that front activation increased indiscriminate violence in the form of civilian killings, against rules of engagement, by between 10 and 13 fold. By contrast, selective violence, accepted and used by German commanders to achieve strategic goals, did not differentially change upon front activation. We further documented a first-stage in accountability, by leveraging variation in the strength of the command chain. We showed how units assigned to treated municipalities were more likely to become composite at front activation, a state that likely prevented unit's commanders to exert effective control over soldiers. Finally, through heterogeneity analysis, we documented how effects were strongest near inexperienced units and where looting incentives were highest. We think of these results as showing the effect of one particularly salient exogenous shock to accountability. Challenges to the monitoring capacity of the principal, however, were ubiquitous and certainly were not limited to the front-activation events we study.

We interpret our evidence as consistent with economic incentives - and particularly the cost of misbehavior in terms of accountability - playing a paramount role viz-a-viz preferences in shaping patterns of violence against civilians. This is in line with the literature from crime economics (Becker and Landes, 1974; Glaeser et al., 1996), but novel in its application to a war setting and to an extreme form of violence. It is also consistent with the literature on the economics of mass killings and genocides, which has however focused on returns to victimization for the perpetrators (Esteban et al., 2015; Azam, 2006), preferences for violence towards specific opposing groups (Guarnieri, 2025) and for following orders (Browning, 1993). As civilian victimization has high social costs and could have a profound negative effect on economic development (Collier et al., 2003), our findings bear policy implications for both international organizations and army commands wanting to prevent episodes of indiscriminate violence in the future. We argue that, in addition to the importance of returns to victimization, also its costs is a key determinant of misbehavior and that specific unit characteristics, such inexperience, may exacerbate violent reactions. In

other words, while large-scale civilian victimization requires some knowledge or endorsement by army commanders, small-scale collective killings may depend on behavioral incentives that individual soldiers face in different fighting situations.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond, "Preferences for Truth-Telling," *Econometrica*, 2019, *87* (4), 1115–53.

Abrahms, Max and Philip BK Potter, "Explaining Terrorism: Leadership Deficits and Militant Group Tactics," *International Organization*, 2015, *69* (2), 311–42.

Acemoglu, Daron, Leopoldo Fergusson, James Robinson, Dario Romero, and Juan F Vargas, "The Perils of High-Powered Incentives: Evidence from Colombia's False Positives," *American Economic Journal: Economic Policy*, 2020, *12* (3), 1–43.

Ager, Philipp, Leonardo Bursztyn, Lukas Leucht, and Hans-Joachim Voth, "Killer Incentives: Rivalry, Performance and Risk-Taking among German Fighter Pilots, 1939-45," *The Review of Economic Studies*, 2021.

Ahmadov, Anar and James Hughes, *Ideology and Civilian Victimization in Civil War*, The London School of Economics and Political Science, 2017.

Aleksievič, Svetlana, *La Guerra Non Ha un Volto di Donna. L'Epopea delle Donne Sovietiche nella Seconda Guerra Mondiale*, Milano: Bompiani, 2018.

Azam, Jean-Paul, "On Thugs and Heroes: Why Warlords Victimize their Own Civilians," *Economics of Governance*, 2006, *7* (1), 53–73.

Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang, "How Much Should We Trust Staggered Difference-in-differences Estimates?," *Journal of Financial Economics*, 2022, *144* (2), 370–95.

Balcells, Laia, "Rivalry and Revenge: Violence against Civilians in Conventional Civil Wars," *International Studies Quarterly*, 2010, *54* (2), 291–313.

Baldissara, Luca, *Atlante Storico della Resistenza Italiana*, Milano: Bruno Mondadori, 2000.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *The Journal of Political Economy*, 1968, *76* (2), 169–217.

Becker, Gary S and William M Landes, *Essays in the Economics of Crime and Punishment*, New York: National Bureau of Economic Research and Columbia University Press, 1974.

Becker, Sascha O, Sharun Mukand, and Ivan Yotzov, "Persecution, Pogroms and Genocide: A Conceptual Framework and New Evidence," *Explorations in Economic History*, 2022, *86* (101471), 1–18.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig, "This Mine is Mine! How Minerals Fuel Conflicts in Africa," *The American Economic Review*, 2017, *107* (6), 1564–610.

Bianchi, Nicola and Michela Giorcelli, "Reconstruction Aid, Public Infrastructure, and Economic Development: The Case of the Marshall Plan in Italy," *Journal of Economic History*, 2023, *83* (2), 501–37.

Biddle, Stephen D., *Military Power: Explaining Victory and Defeat in Modern Battle*, Princeton, NJ: Princeton University Press, 2004.

Bishop, Chris, *The Military Atlas of World War II*, London: Igloo, 2005.

Blattman, Christopher and Edward Miguel, "Civil War," *Journal of Economic Literature*, 2010, *48* (1), 3–57.

Blumenson, Martin, *Salerno to Cassino*, Washington, DC: Center of Military History, United States Army, 1993.

Browning, Christopher, *Ordinary Men : Reserve Police Battalion 101 and the Final Solution in Poland*, New York: Harper Perennial, 1993.

Cannella, Mario, Alexey Makarin, and Ricardo Piqué, "The Political Legacy of Nazi Annexation," *Working Paper*, 2024.

Carrozzi, Felipe, Edward Pinchbeck, and Luca Repetto, "Scars of War: The Local Legacy of WWI Deaths on British Soldiers," *Working Paper*, 2024.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-Way Fixed Effects Estimations with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (5), 2964–96.

_ **and** _ , "Difference-in-Differences Estimators of Intertemporal Treatment Effects," *Review of Economics and Statistics*, 2024, pp. 1–45.

Clausewitz, Carl, *On War*, Penguin UK, 2003.

Collier, Paul, V.L Elliott, Havard Hegre, Anke Hoeffler, Marta Reynal-Querol, and Nicholas Sambanis, "Breaking the Conflict Trap: Civil War and Development Policy," *World Bank Policy Research Report*, 2003, *56793*.

Collotti, Enzo, "Documenti sull'Attività del Sicherheitsdienst nell'Italia Occupata," in "Il Movimento di liberazione in Italia" 1963, p. 77.

_ , *L'Amministrazione Tedesca dell'Italia Occupata, 1943-1945*, Milano: Lerici Editori, 1963.

Costa, Dora, "Leaders: Privilege, Sacrifice, Opportunity and Personnel Economics in the American Civil War," *The Journal of Law, Economics, and Organization*, 2014, *30* (3), 437–62.

_ **and Matthew Kahn**, "Cowards and Heroes: Group Loyalty in the American Civil War," *Quarterly Journal of Economics*, 2003, *118* (2).

Costalli, Stefano, Francesco Niccolò Moro, and Andrea Ruggeri, "The Logic of Vulnerability and Civilian Victimization: Shifting Front Lines in Italy (1943 1945)," *World Politics*, 2020, *72* (4), 679–718.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving," *The Quarterly Journal of Economics*, 2012, *127* (1), 1–56.

_ **, John List, Ulrike Malmendier, and Gautam Rao**, "Voting to Tell Others," *Review of Economic Studies*, 2017, *84* (1), 143–81.

Dippel, Christian and Andreas Ferrara, "What Makes Effective Leaders? Evidence from Civil War Captains and their Soldiers," *Working Paper*, 2025.

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and conflict: An empirical study," *American Economic Review*, 2012, *102* (4), 1310–1342.

_ **, Massimo Morelli, and Dominic Rohner**, "Strategic Mass Killings," *The Journal of Political Economy*, 2015, *123* (5), 1087–132.

Fehr, Ernst and Simon Gächter, "Cooperation and punishment in public goods experiments," *American Economic Review*, 2000, *90* (4), 980–994.

_ **and** _ , "Altruistic punishment in humans," *Nature*, 2002, *415* (6868), 137–140.

Fisher, Ernest, *Cassino to the Alps*, Washington DC: Center of Military History, United States Army, 1993.

Fontana, Nicola, Tommaso Nannicini, and Guido Tabellini, "Historical Roots of Political Extremism: The Effects of Nazi Occupation of Italy," *Journal of Comparative Economics*, 2022, *51* (3), 723–43.

Gächter, Simon and Jonathan F Schulz, "Intrinsic honesty and the prevalence of rule violations across societies," *Nature*, 2016, *531* (7595), 496–499.

Garland, Albert, Howard McGaw-Smyth, and Martin Blumenson, *Sicily and the Surrender of Italy*, Washington DC: Center of Military History, United States Army, 1993.

Gentile, Carlo, *Wehrmacht und Waffen-SS im Partisanenkrieg Italien 1943-1945*, Schoeningh Ferdinand GmbH, 2012.

_ , *I Crimini di Guerra Tedeschi in Italia*, Torino: Einaudi, 2015.

Glaeser, Edward L., Bruce Sacerdote, and José A. Scheinkman, "Crime and Social Interactions," *The Quarterly Journal of Economics*, 1996, *111* (2), 507–48.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–77.

Gressel, Gustav, *Russia's Quiet Military Revolution, and What It Means for Europe*, Policy Brief, 2015.

Grosfeld, Irena, Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya, "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire," *The Review of Economic Studies*, 2020, *87* (1), 289–342.

Guarnieri, Eleonora, "Cultural Distance and Ethnic Civil Conflict," 2023.

_ , "Cultural Distance and Ethnic Civil Conflict," *American Economic Review*, 2025, *forthcoming*.

_ **and Ana Tur-Prats**, "Cultural Distance and Conflict-Related Sexual Violence," *Quarterly Journal of Economics*, 2023, *138* (3), 1817–61.

Heldring, Leander, "The Origins of Violence in Rwanda," *The Review of Economic Studies*, 2021, *88* (2), 730–63.

Hobbes, Thomas, *Leviathan or The Matter, Forme and Power of a Commonwealth Ecclesiasticall and Civil*, London: Andrew Crook, 1651.

Horowitz, Shale and Min Ye, "Targeting Civilians in Ethno-Territorial Wars: Power- and Preference-Based Sources of Ethnic Cleansing and Mass Killing Strategies," *Studies in conflict and terrorism*, 2013, *36* (5), 372–93.

Humphreys, Macartan and Jeremy Weinstein, "Handling and Manhandling Civilians in Civil War," *American Political Science Review*, 2006, *100* (3).

ISTAT, *Morti e Dispersi per Cause Belliche negli Anni 1940-1945*, Istituto Centrale di Statistica, 1957.

Iyengar, Radha and Jonathan Monten, "Is There an Emboldenment Effect? Evidence from the Insurgency in Iraq," *Journal of Strategic Studies*, 2022, *45* (4).

Kalyvas, Stathis N., *The Logic of Violence in Civil War*, Cambridge: Cambridge University Press, 2006.

Kathman, Jacob D and Reed M Wood, "Managing Threat, Cost, and Incentive to Kill: The Short and Long-term Effects of Intervention in Mass Killings," *Journal of Conflict Resolution*, 2011, *55* (5), 735–60.

Kesselring, Albert, *Soldato Fino all'Ultimo Giorno*, Gorizia: LEG Edizioni, 1954.

King, Gary and Langche Zeng, "Logistic Regression in Rare Events Data," *Political Analysis*, 2001, (9).

Klinkhammer, Lutz, *Stragi Naziste in Italia*, Donzelli, 2006.

_ , *L'Occupazione Tedesca in Italia, 1943-1945*, Torino: Bollati Boringhieri, 2016.

Kurowski, Franz, *Battleground Italy, 1943-1945. The German Armed Forces in the Battle for the "Boot"*, Winnipeg: J.J. Fedorowicz Publishing, Inc., 2003.

Leiby, Michele L., "Wartime Sexual Violence in Guatemala and Peru," *International Studies Quarterly*, 2009, *53* (2), 445–68.

Lingen, Kerstin Von, *Kesselring's Last Battle: War Crimes Trials and Cold War Politics, 1945-1960*, United States of America: University Press of Kansas, 2009.

Lt Col Robertson, J.A., R. Burr, and B. Barth, "USAF THOR Database," 2013.

Midlarsky, Manus, *The Killing Trap: Genocide in the Twentieth Century*, Cambridge: Cambridge University Press, 2005.

Milgram, Stanley, *Obedience to Authority: An Experimental View*, London: Tavistock, 1974.

Mitchell, Neil J, *Agents of Atrocity: Leaders, Followers, and the Violation of Human Rights in Civil War*, New York ; Basingstoke: Palgrave Macmillan, 2004.

Nafziger, George, *The German order of battle : infantry in World War II*, London: Greenhill, 2000.

Olson, Mancur, *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*, New York: Basic Books, 2000.

Oppenheim, Ben and Michael Weintraub, "Doctrine and Violence: The Impact of Combatant Training on Civilian Killings," *Terrorism and Political Violence*, 2017, *29* (6), 1126–48.

Perez-Truglia, Ricardo and Guillermo Cruces, "Partisan Interactions: Evidence from a Field Experiment in the United States," *Journal of Political Economy*, 2017, *125* (4), 1208–43.

Ronchetti, Gabriele, *La Linea Gotica. I Luoghi dell'Ultimo Fronte di Guerra in Italia*, Fidenza: Mattioli 1885, 2018.

_ **and Maria Angela Ferrara**, *La Linea Gustav. I Luoghi delle Battaglie da Ortona a Cassino*, Fidenza: Mattioli 1885, 2014.

Rosen, Sherwin, "The Military as an Internal Labor Market: Some Allocation, Productivity, and Incentive Problems," *Social Science Quarterly*, 1992, *73*, 227–37.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Rousseau, Jean-Jacques, *Discours sur l'Origine et les Fondements de l'Ine'galite' parmi les Hommes*, Amsterdam: Marc-Michel Rey, 1755.

Short, Neil and Chris Taylor, *German Defences in Italy in World War II*, Osprey Publishing, 2006.

Solzhenitsyn, Aleksandr, *Prussian Nights: A Narrative Poem*, London: Collins Harvill Press, 1977.

Stein, Marcel, "A Flawed Genius: Field Marshal Walter Model. A Critical Biography," 2010.

Valentino, Benjamin, Paul Huth, and Dylan Balch-Lindsay, "'Draining the Sea': Mass Killing and Guerrilla Warfare," *International Organization*, 2004, *58* (2), 375–407.

War Crimes Commission, "The trial of Albert Kesselring," in "Law Reports of Trials of War Criminals. Volume VIII," London: War Crimes Commission, 1949.

Weinstein, Jeremy, *Inside Rebellion: The Politics of Insurgent Violence*, Cambridge: Cambridge University Press, 2007.

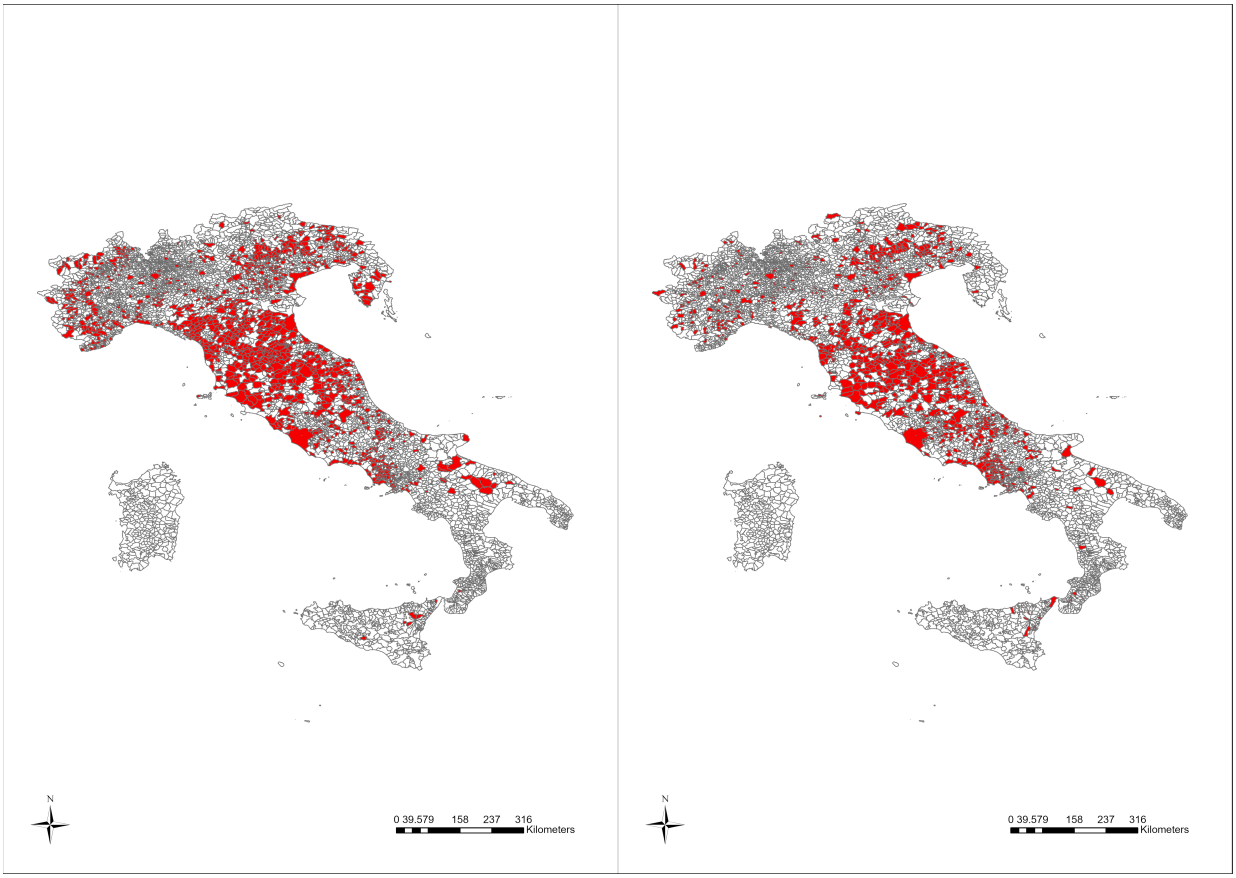
Wray, Timothy Allen, *Standing fast: German defensive doctrine on the Russian Front during World War II: prewar to March 1943*, Vol. 5, US Army Command and General Staff College, 1987.

Figures and Tables

Figure 1: GEOGRAPHICAL DISTRIBUTION OF INDISCRIMINATE AND SELECTIVE VIOLENCE DURING THE ITALIAN CAMPAIGN

(a) Indiscriminate Violence: Collective Killing

(b) Selective Violence: Single Killing



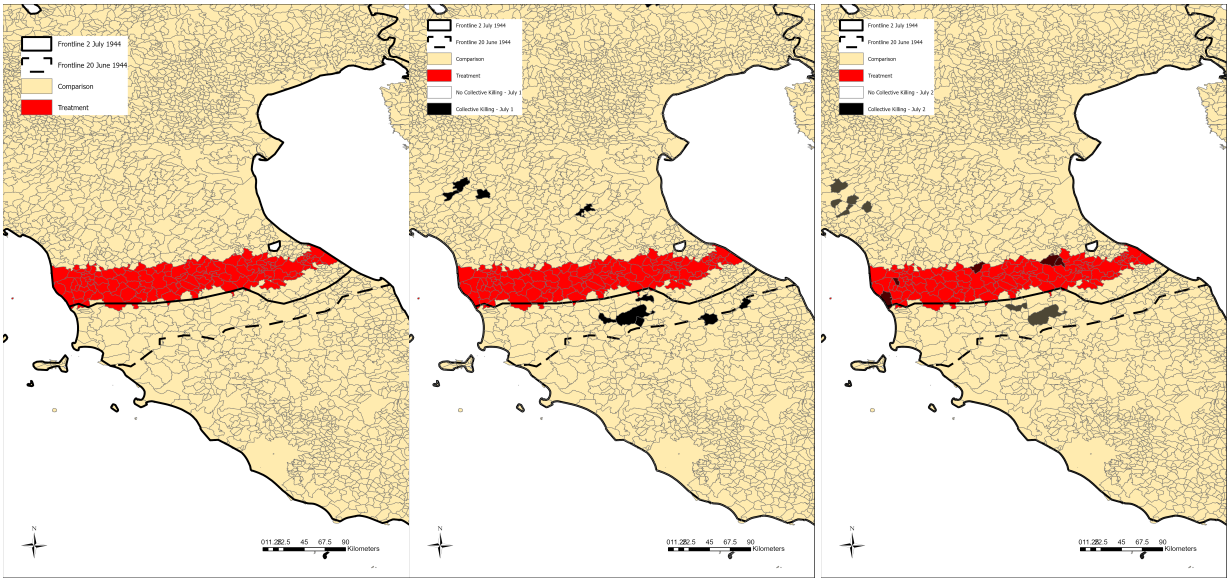
Notes. Panel (a) shows all Italian municipalities that experienced at least one episode of collective killing, where two or more civilians were killed by German soldiers, between July 15, 1943 and May 2, 1945. Panel (b) shows all Italian municipalities that experienced at least one episode of single killing, where only one civilian was killed by German soldiers, between July 15, 1943 and May 2, 1945. Municipalities are drawn using the 1936 borders. *Source*: http://www.straginazifasciste.it/

Figure 2: FRONT LINES 9 JULY 1943 TO 2 MAY 1945

Notes. The figure shows the 32 front lines along which Axis and Allied armies confronted during the Italian Campaign, between the Allied landing in Sicily on 9 July 1943 and Germany surrender on 2 May 1945.
Source: Blumenson (1993)

Figure 3: FRONT ACTIVATION AND INDISCRIMINATE VIOLENCE: AN EXAMPLE

(a) New front line, July 2nd 1944 (b) Collective Killing in k=-1 (c) Collective Killing in k=0

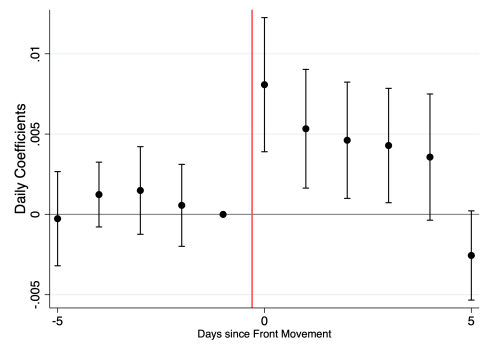


Notes. Panel (a) shows treated municipalies that were more than 40 km away from the June 20th 1944 front line and within 40 km of the 2nd of July 1944 one after the combat zone moves North. Panel (b) shows episodes of collective killings, where at least two civilians were killed by German soldiers, proxy for indiscriminate violence, that happened the day before the activation of the new fronline. Panel (c) shows episodes of collective killings that happened on the day of front activation. The units of observation are Italian municipalities with 1936 boundaries.
Source: Blumenson (1993) and http://www.straginazifasciste.it/

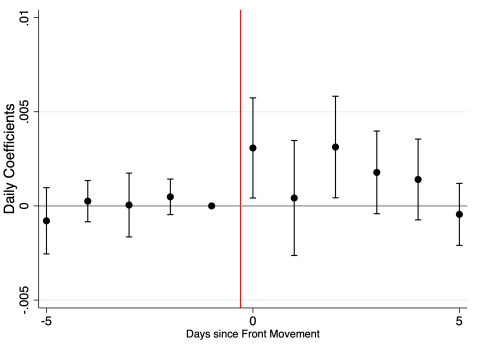
Figure 4: THE EFFECTS OF FRONT ACTIVATION ON CIVILIAN VICTIMIZATION

Indiscriminate Violence

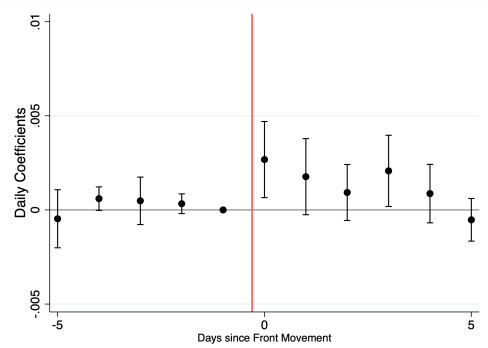
(a) Collective Killing



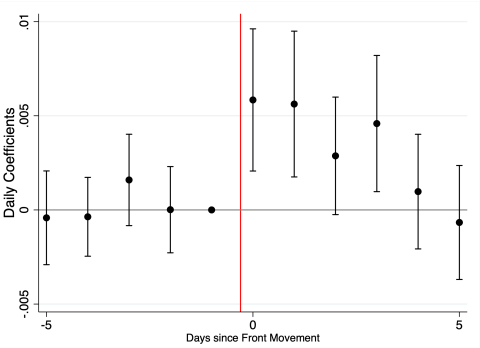
(b) No Partisans Killing



(c) Vulnerable Killing

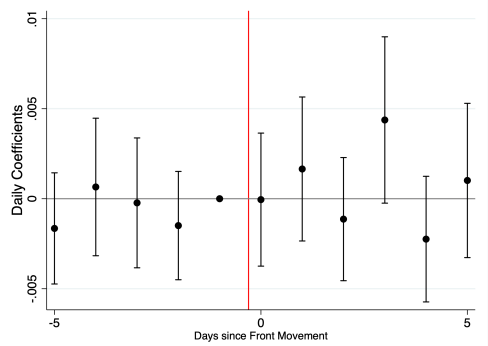


(d) Hand-Coded Indiscriminate Killing

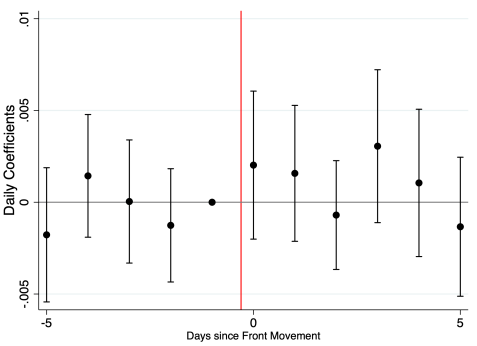


Selective Violence

(e) Single Killing



(f) Hand-Coded Selective Killing



Notes. Daily coefficients and 95% standard errors estimated from equation 1 from 5 days before to 5 days after front activation. Day $k = -1$ is the excluded category. Panels A-D measure indiscriminate violence; Panels E-F selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (Panel A); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (Panel B); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, Panel C); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (Panel D); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (Panel E); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (Panel F). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table 1: SUMMARY STATISTICS

	(1) Observations	(2) Mean	(3) St. Dev.	(4) Min	(5) Max
Collective yes/no	2,330,790	0.0006	0.0246	0	1
No Partisan yes/no	2,330,790	0.0003	0.0167	0	1
Vulnerable yes/no	2,330,790	0.0002	0.0135	0	1
Indiscriminate Hand-Coded yes/no	2,330,340	0.0004	0.0208	0	1
Single yes/no	2,330,790	0.0005	0.0220	0	1
Selective Hand-Coded yes/no	2,330,340	0.0005	0.0216	0	1
Treated * Day= 0	2,330,790	0.0012	0.0351	0	1
Elite Unit yes/no	2,330,790	0.0666	0.2493	0	1
Division number	2,320,218	31.9104	90.5363	0	715
Bombing yes/no	2,330,790	0.0011	0.0324	0	1
Dist. to Enemy (km)	2,330,790	115.9825	237.5615	0	1127
Composite yes/no	2,330,790	0.0135	0.1153	0	1
No. Sub-units	2,330,790	0.2908	0.5030	0	4

Notes. Summary statistics for all variables employed in the tables of the main paper. The sample is the same employed for the main regressions. It subsets to artificial panel containing a front movement, and excludes Sardinia, where no front movements and civilian killings took place.

Table 2: DOES A DROP IN ACCOUNTABILITY AFFECT SOLDIERS' (MIS)BEHAVIOR?

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0006	0.0005	0.0003	0.0000	-0.0015	-0.0013
	(0.0013)	(0.0008)	(0.0003)	(0.0012)	(0.0015)	(0.0016)
Treated * Day= 0	0.0081***	0.0031**	0.0027***	0.0058***	-0.0000	0.0020
	(0.0021)	(0.0014)	(0.0010)	(0.0019)	(0.0019)	(0.0021)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,790	2,330,790	2,330,790	2,330,340	2,330,790	2,330,340
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table 3: WHAT DRIVES THE INCREASE IN INDISCRIMINATE VIOLENCE?

Panel A:	Sorting		Insecurity	
Dependent variable:	Elite	Div. No.	Bombing	Dist. to Enemy
	(1)	(2)	(3)	(4)
Treated * Day= -2	0.0376	-4.3525	-0.0054	11.4130*
	(0.0377)	(7.7595)	(0.0033)	(6.1384)
Treated * Day= 0	-0.0033	10.4505	-0.0034	6.8603
	(0.0524)	(15.7523)	(0.0039)	(5.8357)
Mean dep. var.	0.2700	105.3842	0.0011	473.0849
Observations	402,591	399,432	402,591	402,591
Municipality clust.	7,063	7,063	7,063	7,063
Div. 1 x Front Clust.	398	392	398	398
Div. 1 x Front	NO	NO	NO	NO
Panel B:	Accountability			
Dependent variable:	Composite	No. sub-units	Composite	No. sub-units
	(1)	(2)	(3)	(4)
Treated * Day= -2	0.0072	0.0209	0.0084	0.0094
	(0.0242)	(0.0297)	(0.0109)	(0.0129)
Treated * Day= 0	0.1344**	0.2314**	0.1299**	0.2229*
	(0.0662)	(0.1142)	(0.0606)	(0.1153)
Mean dep. var.	0.0521	1.0773	0.0521	1.0773
Observations	402,591	402,591	402,588	402,588
Municipality clust.	7,063	7,063	7,063	7,063
Div. 1 x Front Clust.	398	398	395	395
Div. 1 x Front	NO	NO	YES	YES
Municipality FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Front FE	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES

Notes. Coefficients estimated from equation 1. In this case, $Day = 0$ corresponds to the first available military map on the day or immediately after front activation. $Day = -2$ corresponds to the second last military map before activation (often the previous front line). See text and Data Appendix for detail on the matching procedure. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t - 1$. In Panel A,*Elite* is an indicator for the presence of elite Nazi troops (SS, paratroops, panzer and mountain divisions, column 1), *Div No.* is the number of units included in each unit's official name (column 2), *Bombing* is an indicator for Allied bombing (column 3), and *Dist. to Enemy* is the distance from the Allied troops (column 4). In Panel B, *Composite* is an indicator for being assigned to a compositive division that features regiments from different units (columns 1 and 3); *No. Sub-Units* is the count of sub-units assigned to a composite one (columns 2 and 4). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/; Lt Col Robertson et al. (2013); Baldissara (2000); and military German maps collected from Bundes Archiv.

Table 4: HETEROGENEITY: DO MILITARY UNITS' CHARACTERISTICS MATTER?

Dependent variable:	Collective killing, yes/no			
	Indoctrination	Training	Fighting East	Combat Experience
	(1)	(2)	(3)	(4)
Treated x SS	0.0094			
	(0.0089)			
Treated x not SS	0.0083***			
	(0.0023)			
Treated x Germany		0.0185		
		(0.0127)		
Treated x not Germany		0.0080***		
		(0.0022)		
Treated x East			0.0031	
			(0.0024)	
Treated x West			0.0116***	
			(0.0032)	
Treated x new				0.0108***
				(0.0029)
Treated x old				0.0035
				(0.0030)
Mean dep. var.	0.0006	0.0006	0.0006	0.0006
Observations	2,327,945	2,330,614	2,330,614	2,330,614
Clusters	7,063	7,063	7,063	7,063
P-value equality test	0.90	0.41	0.03	0.08
Div. 1 x Front	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Front FE	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers. *SS* and *not SS* are indicators for municipalities whose closest German division was/was not a SS unit. *Germany* and *not Germany* are indicators for municipalities whose closest German division was/was not trained in Germany. *East* and *West* are indicators for municipalities whose closest German division was/was not previously deployed on the East Front. *Old* and *New* are indicators for municipalities whose closest German division was created before/after 1942, the sample's mean of creation date of divisions deployed during the Italian Campaign. Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table 5: HETEROGENEITY: DO LOCAL ECONOMIC CHARACTERISTICS MATTER?

Dependent variable:	Collective killing yes/no			
	(1)	(2)	(3)	(4)
Treated x <= 30 km capital	0.0095***			
	(0.0028)			
Treated x > 30 km capital	0.0025			
	(0.0041)			
Treated x industrial		0.0188***		
		(0.0050)		
Treated x rural		0.0032		
		(0.0022)		
Treated x large municipality			0.0126***	
			(0.0034)	
Treated x small municipality			0.0018	
			(0.0013)	
Treated x high density				0.0090***
				(0.0029)
Treated x low density				0.0093**
				(0.0036)
Mean dep. var.	0.0006	0.0006	0.0006	0.0006
Observations	2,330,141	2,330,273	2,330,306	2,330,284
Clusters	7,063	7,063	7,063	7,063
P-value equality test	0.16	0.00	0.00	0.94
Div. 1 x Front	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Front FE	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers. *<=30km* and *>30km* capital are indicators for municipalities within/beyond 30km of the provincial capital. *Industrial* and *Rural* are indicators for municipalities whose employment in the industrial sector was above/below the sample mean. *Large* and *small municipality* are indicators for municipalities above/below median of total resident population in 1936. *High* and *low density* are indicators for municipalities above/below the median of population density in 1936. Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/; and 1936 Population Census.

Table 6: HETEROGENEITY: DO MACRO WAR TRENDS MATTER?

Dependent variable:	Collective killing (YES/NO)			
	Front characteristics		War trends	
	(1)	(2)	(3)	(4)
Treated x long front	0.0136***			
	(0.0046)			
Treat t 0 x short front	0.0059**			
	(0.0024)			
Treated x rugged		0.0053**		
		(0.0023)		
Treated x not rugged		0.0149***		
		(0.0049)		
Treated x pre DD			0.0046**	
			(0.0021)	
Treated x post DD			0.0147***	
			(0.0046)	
Treated x early war				0.0052**
				(0.0023)
Treated x late war				0.0129***
				(0.0041)
Mean dep. var.	0.0006	0.0006	0.0006	0.0006
Observations	2,330,614	2,330,614	2,330,614	2,330,614
Clusters	7,063	7,063	7,063	7,063
P-value equality test	0.14	0.07	0.05	0.10
Div. 1 x Front	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Front FE	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers. *Long* and *short front* are indicators for fronts whose duration was above/below the median duration of the 32 front of 18 days. *Rugged* and *not rugged* are indicators for municipalities above/below the median of the standard deviation of terrain ruggedness, calculated across all treated municipalities separately for each front line. *Pre* and *Post D-Day* are indicators for front activations before/after the Allied landing in Normandy on June 6 1944. *Early* and *late war*are indicators for front activations in the first half and second half of the Italian Campaign. Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/; Blumenson (1993); and 1936 Population Census.

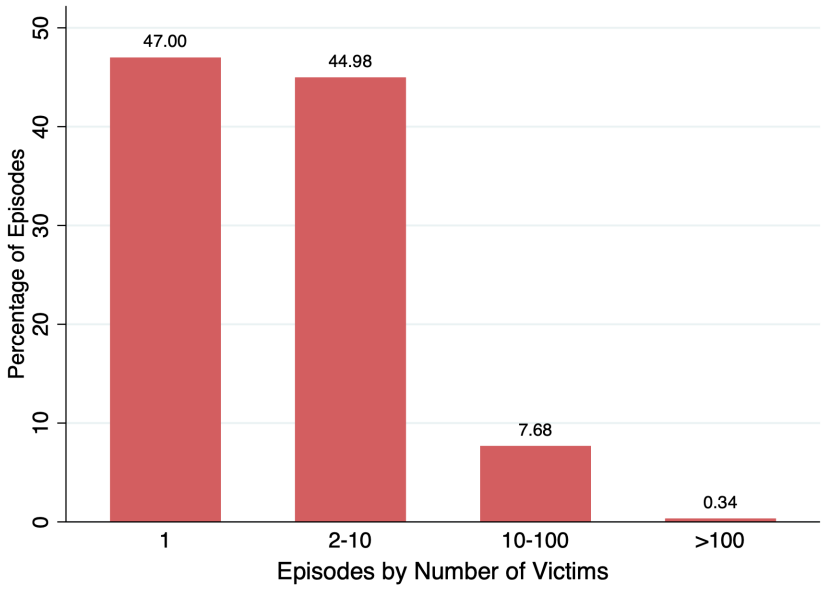
Online Results Appendix — Not for Publication

A Results Appendix

A.1 Additional Figures and Tables

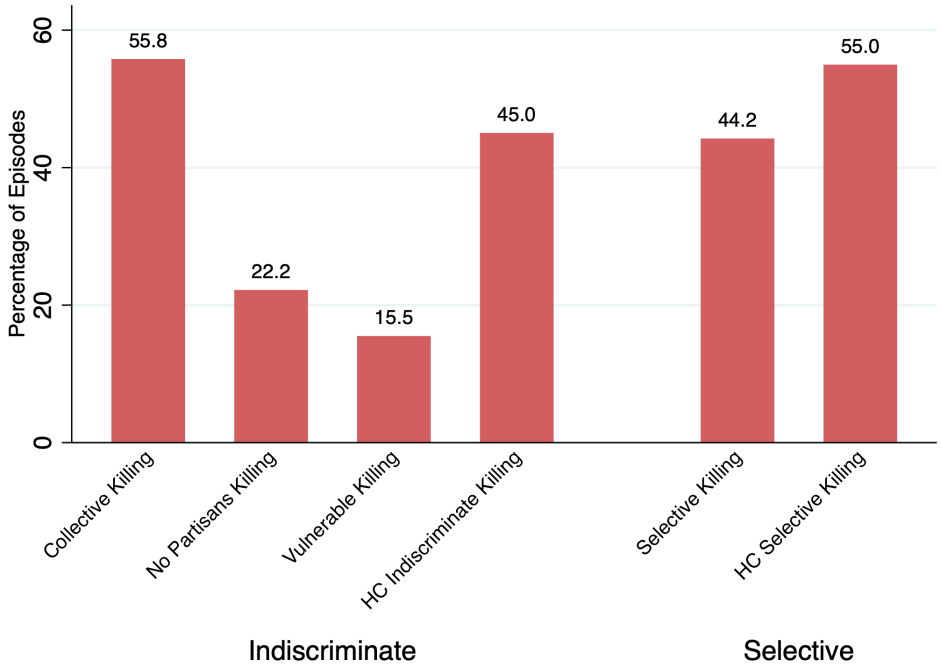
=

Figure A.1: CIVILIAN VICTIMIZATION BY EPISODE SIZE DURING THE ITALIAN CAMPAIGN



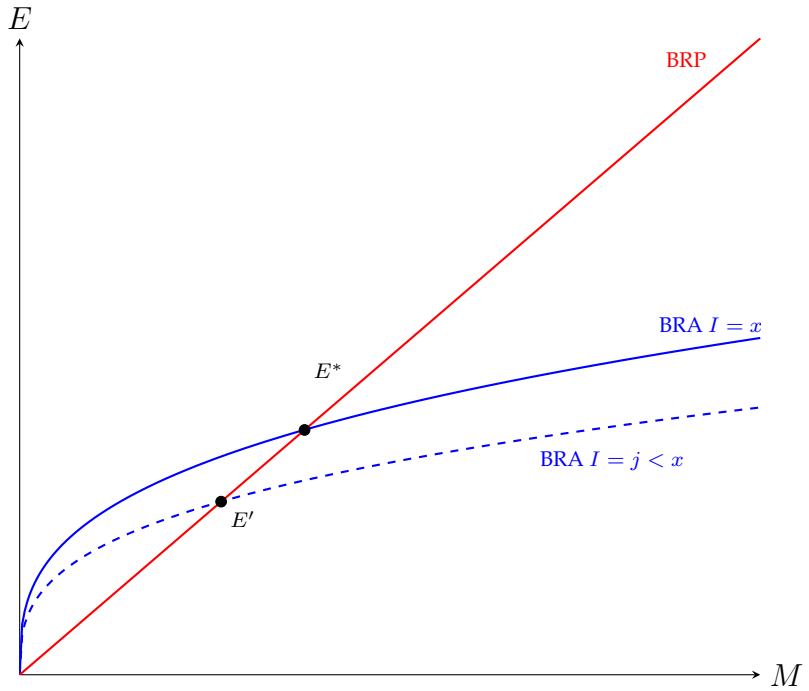
Notes. Percentage of episodes out a total of 5,291 with a single victim, between 2 and 10, 10 and 100 and more than 100 victims killed by German soldiers between July 15, 1943 and May 2, 1945. *Source*: http://www.straginazifasciste.it/

Figure A.2: CIVILIAN VICTIMIZATION BY VIOLENCE TYPE (INDISCRIMINATE VS SELECTIVE)



Notes. Percentage of episodes of indiscriminate and selective violence. *Collective Killing* is the share of episodes with at least two victims; *No Partisan Killing* is the share of episodes of collective killings not related to partisan activity; *Vulnerable Killing* is the share of episodes of collective killings that involved at least a vulnerable individual (woman, child, and elderly); *HC Indiscriminate Killing* is the share of hand-coded indiscriminate killings based on the Atlas description; *Single Killing* is the share of episodes with a single victim; *HC Selective Killing* is the share of hand-coded selective killings based on the Atlas description.
Source: http://www.straginazifasciste.it/

Figure A.3: EXAMPLE OF A DOWNWARD SHIFT IN EFFORT

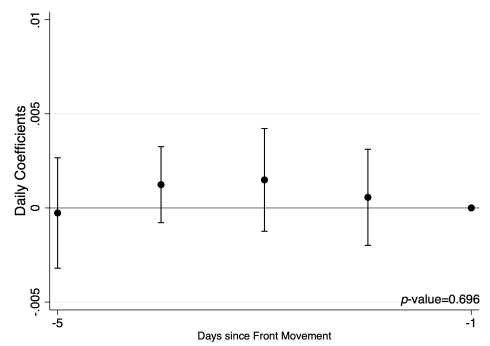


Notes. Best Response curves for Principal (P) and Agent (A) are represented. The graph provides an example of change in the equilibrium point of agents' (soldiers) effort (E) for a reduction in information (I) that is available to the principal. M is the monitoring cost paid by the Principal. See text for detail.

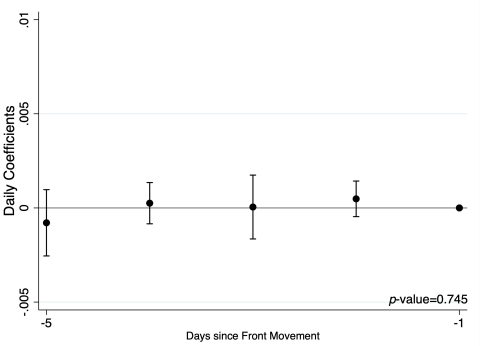
Figure A.4: TEST FOR PRE-TRENDS

Indiscriminate Violence

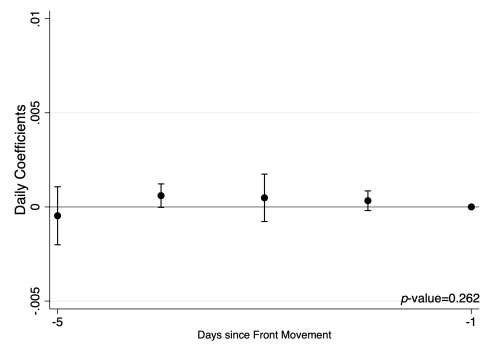
(a) Collective Killing



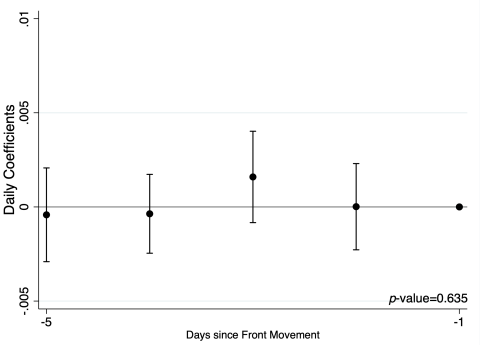
(b) No Partisan Killing



(c) Vulnerable Killing

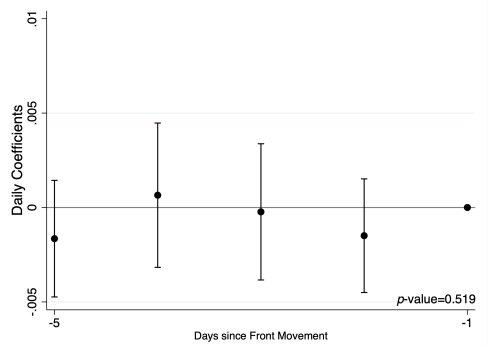


(d) Hand-Coded Indiscriminate Killing

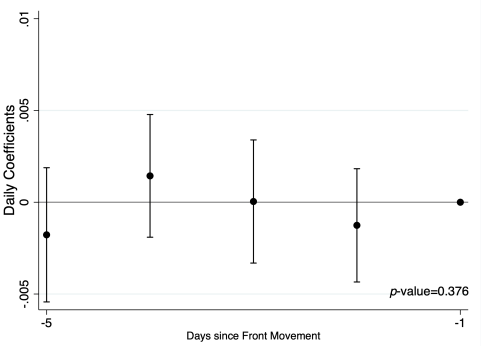


Selective Violence

(e) Single Killing

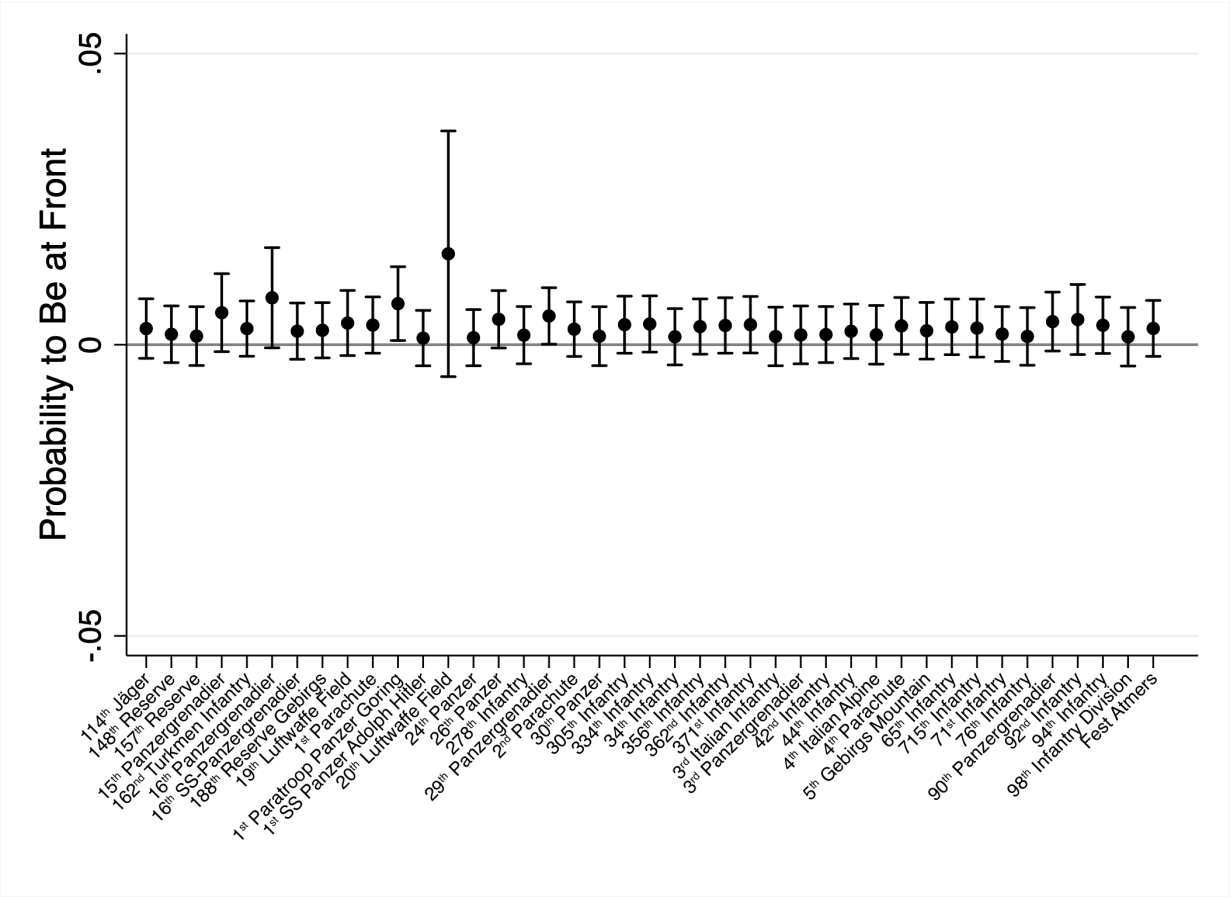


(f) Hand-Coded Selective Killing



Notes. Daily coefficients and 95% standard errors estimated from equation 1 in the 5 days before the front activation. Day $k = -1$ is the excluded category. Panels A-D measure indiscriminate violence; Panels E-F selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (Panel A); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (Panel B); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, Panel C); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (Panel D); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (Panel E); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (Panel F). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Figure A.5: PROBABILITY OF GERMAN DIVISIONS TO BE AT THE FRONT



Notes. Predicted probability for the 47 Divisions operating during the Italian Campaign to fight on the front line based on division fixed effects. The excluded division is 1st Paratroop Panzer Division Hermann Göring.

Table A.1: BALANCING TESTS BETWEEN TREATED AND COMPARISON MUNICIPALITIES

Day before Front Activation:	5 days before (1)	10 days before (2)	15 days before (3)	20 days before (4)
Panel A: Bombing yes/no				
Treated * Day= 0	0.0011	-0.0025	0.0012	0.0007
	(0.0012)	(0.0018)	(0.0015)	(0.0011)
Mean dep. var.	0.0011	0.0010	0.0010	0.0009
Observations	3,037,642	3,003,487	2,969,737	2,936,080
Clusters	6,939	6,939	6,939	6,939
Panel B: Partisan yes/no				
Treated * Day= 0	0.0024	0.0017	0.0003	-0.0000
	(0.0016)	(0.0012)	(0.0008)	(0.0007)
Mean dep. var.	0.0008	0.0007	0.0007	0.0007
Observations	3,037,460	3,003,317	2,969,574	2,935,923
Clusters	6,939	6,939	6,939	6,939
Municipality FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES

Notes. OLS regressions predicting fighting outcomes five days (Column 1), 10 days (Column 2), 15 days (Column 3), and 20 days (Column 4) before front activation. Each regression is estimated on a panel dataset covering Italian municipalities North of the front line (in Axis-controlled territory) in each day from June 8, 1943 to May 2, 1945. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. The dependent variables are binary indicators for Allied bombing (Panel A) and for Resistance activity (Panel B). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/; Lt Col Robertson et al. (2013); and Baldissara (2000).

Table A.2: CROSS-SECTIONAL COMPARISON BETWEEN EVER AND NEVER TREATED MUNICIPALITIES

	(1)	(2)	(3)	(4)	(5)
Panel A: Geographical Factors					
	No. Rivers	Elevation	Precipitations	Ruggedness	Dist. Coast
Treated * Day=0	1.1253***	27.3949	-46.2243***	-157.4497	43.8143***
	(0.0389)	(35.4309)	(3.9709)	(503.2580)	(3.0254)
Mean dep. var.	1.2203	458.2947	78.4683	6857.1418	67.7711
Observations	7,063	7,061	7,062	7,062	7,063
Panel B: Socio-economic Factors					
	Pop. Density	Dist. Prov. Cap.	Industry %	Agriculture %	No. Partisan Bands
Treated * Day=0	1.2989	-12.3546***	0.0024	-0.0121	0.0077
	(33.2219)	(2.0163)	(0.0101)	(0.0157)	(0.0177)
Mean dep. var.	187.5189	19.1849	0.1007	0.2928	0.0514
Observations	7,061	7,063	7,053	7,061	7,063

Notes. *Treated* is an indicator that equals one for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$ at least once during the Italian Campaign and zero for municipalities that were never within 40km of the combat zone at time t but were outside it at time $t-1$. Panel A uses geographical characteristics as outcome variables, such as the number of waterways in a municipality (column 1), the average municipality elevation in meters (column 2), the average municipality rainfall in centimeters measured in terms of average precipitations in the month of September between 1950 and 2000 (column 3), the average municipality ruggedness measured in per-thousant points (column 4), and municipality distance from the coast in kilometers (column 5). Panel B uses 1936 municipality socio-economic characteristics as outcome variables, such as population density (column 1), distance in kilometers to the nearest point of the boundary of a provincial capital (column 2), the share of the active population employed in the industrial sector (column 3), the share of the active population employed in the agricultural sector (column 4), and the number of partisan bands reported as active in the territory of a municipality at Armistice in September 1943 (column 5). All distance measures are calculated from municipalities' centroids. Robust standard errors are reported in parentheses. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/; and 1936 Population Census.

Table A.3: HETEROGENEITY BY EPISODE SIZE

Dependent variable:	Collective Killing			
	All	Between 2 and 100	Between 2 and 50	Between 2 and 10
	(1)	(2)	(3)	(4)
Treated * Day= -2	0.0006	0.0006	0.0004	0.0005
	(0.0013)	(0.0013)	(0.0013)	(0.0013)
Treated * Day= 0	0.0081***	0.0081***	0.0080***	0.0064***
	(0.0021)	(0.0021)	(0.0021)	(0.0020)
Mean dep. var.	0.0006	0.0006	0.0006	0.0005
Observations	2,330,790	2,330,790	2,330,790	2,330,790
Clusters	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Front FE	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers. Column 1 includes all collective killings; column 2 collective killings with between 2 and 100 victims; column 3 collective killings with between 2 and 50 victims; column 4 collective killings with between 2 and 10 victims. Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.4: ROBUSTNESS CHECK: ONLY MUNICIPALITIES WITHIN 100 KM OF GERMAN DIVISION

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0006	0.0004	0.0004	-0.0001	-0.0015	-0.0010
	(0.0014)	(0.0009)	(0.0003)	(0.0013)	(0.0017)	(0.0017)
Treated * Day= 0	0.0082***	0.0031**	0.0024**	0.0057***	-0.0007	0.0020
	(0.0022)	(0.0015)	(0.0011)	(0.0021)	(0.0020)	(0.0022)
Mean dep. var.	0.0008	0.0004	0.0002	0.0006	0.0007	0.0007
Observations	922,680	922,680	922,680	922,470	922,680	922,470
Clusters	6,646	6,646	6,646	6,646	6,646	6,646
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities within 100km of a German division . *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.*** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.5: ROBUSTNESS CHECK: ONLY MUNICIPALITIES WITHIN 50 KM OF GERMAN DIVISION

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0005	0.0006	0.0002	0.0011	-0.0009	-0.0018
	(0.0017)	(0.0011)	(0.0003)	(0.0013)	(0.0019)	(0.0020)
Treated * Day= 0	0.0079***	0.0030*	0.0021*	0.0076***	-0.0009	-0.0001
	(0.0026)	(0.0017)	(0.0012)	(0.0024)	(0.0022)	(0.0024)
Mean dep. var.	0.0013	0.0006	0.0003	0.0009	0.0012	0.0012
Observations	385,891	385,891	385,891	385,747	385,891	385,747
Clusters	5,504	5,504	5,504	5,504	5,504	5,504
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities within 50km of a German division . *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.6: ROBUSTNESS CHECK: ONLY MUNICIPALITIES WITHIN 25 KM OF GERMAN DIVISION

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0001	0.0004	0.0001	0.0031	-0.0016	-0.0044*
	(0.0023)	(0.0019)	(0.0002)	(0.0021)	(0.0025)	(0.0027)
Treated * Day= 0	0.0096***	0.0029	0.0018	0.0093***	0.0001	0.0008
	(0.0036)	(0.0025)	(0.0014)	(0.0033)	(0.0034)	(0.0038)
Mean dep. var.	0.0019	0.0011	0.0005	0.0015	0.0018	0.0017
Observations	141,141	141,141	141,141	141,066	141,141	141,066
Clusters	3,627	3,627	3,627	3,627	3,627	3,627
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities within 25km of a German division . *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.7: ROBUSTNESS CHECK: USING ONLY 19 FRONT MOVEMENTS FOR WHICH TROOPS' CHARACTERISTICS ARE AVAILABLE

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0001	-0.0001	-0.0001	0.0000	-0.0013	-0.0021
	(0.0019)	(0.0011)	(0.0001)	(0.0015)	(0.0024)	(0.0025)
Treated * Day= 0	0.0096***	0.0028	0.0037**	0.0068**	-0.0007	0.0022
	(0.0031)	(0.0019)	(0.0017)	(0.0027)	(0.0028)	(0.0032)
Mean dep. var.	0.0005	0.0003	0.0002	0.0004	0.0005	0.0004
Observations	1,476,167	1,476,167	1,476,167	1,475,895	1,476,167	1,475,895
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, limiting the sample to 19 experiments for which time-invariant troops' characteristics are available before and after the front movement. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/; and German military maps collected from Bundes Archiv.

Table A.8: TEST FOR TROOPS' SORTING BY CONTROLLING FOR NEAREST DIVISION FIXED EFFECTS

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0008	-0.0003	0.0011*	-0.0009	-0.0014	0.0000
	(0.0015)	(0.0011)	(0.0006)	(0.0016)	(0.0018)	(0.0017)
Treated * Day= 0	0.0087***	0.0029*	0.0035***	0.0047**	-0.0022	0.0017
	(0.0023)	(0.0016)	(0.0011)	(0.0020)	(0.0019)	(0.0021)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,614	2,330,614	2,330,614	2,330,164	2,330,614	2,330,164
Clusters 1	7,063	7,063	7,063	7,063	7,063	7,063
Front x Day x Div. 1	YES	YES	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, controlling for nearest division fixed effects. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.9: INSECURITY: DOES PARTISAN ACTIVITY MATTER?

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treat t 0 x <= 15 km partisans	0.0116***	0.0005	0.0048***	0.0034	-0.0037	0.0055
	(0.0042)	(0.0021)	(0.0018)	(0.0028)	(0.0033)	(0.0044)
Treat t 0 x > 15 km partisans	0.0054**	0.0045*	0.0021*	0.0045*	-0.0013	-0.0014
	(0.0024)	(0.0023)	(0.0012)	(0.0026)	(0.0023)	(0.0018)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,317	2,330,317	2,330,317	2,329,864	2,330,317	2,329,864
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
P-value equality test	0.20	0.21	0.20	0.78	0.55	0.15
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of fron activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. *<=15 and >15 km partisans* are indicators for municipalities within/beyond 15 km of a partisan headquarter at the signature of the Armistice (September 1943). The dependent variables are indicators for collective civilian killings, where two or more civilians were killed by Nazi soldiers (column 1), collective civilian killings excluding episodes related to Partisans activity (column 2), collective civilian killings that only include vulnerable population (women, children, and elderly, column 3), killings with an indiscriminate component from the Atlas' description (column 4), single civilian killing, where only one person was killed by Nazi soldiers (column 5), and killings with a selective component from the Atlas' description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.10: COMPOSITE UNITS AND INSECURITY

Dependent variable:	Accountability			
	Composite (1)	No. sub-units (2)	Composite (3)	No. sub-units (4)
Dist. Allied Unit (Km)	0.0004*	0.0010***	-0.0001	0.0001
	(0.0002)	(0.0004)	(0.0002)	(0.0002)
Treated * Day= -2	0.0027	0.0091	0.0089	0.0086
	(0.0242)	(0.0294)	(0.0114)	(0.0134)
Treated * Day= 0	0.1317**	0.2243*	0.1304**	0.2221*
	(0.0669)	(0.1162)	(0.0605)	(0.1153)
Mean dep. var.	0.0521	1.0773	0.0521	1.0773
Observations	402591	402591	402588	402588
Municipality clust.	7063	7063	7063	7063
Div. 1 x Front Clust.	398	398	395	395
Div. 1 x Front	NO	NO	YES	YES
Municipality FE	YES	YES	YES	YES
Day FE	YES	YES	YES	YES
Front FE	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of fron activation, controlling for distance from the nearest Allied unit. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. The dependent variables are an indicator for being assigned to a compositive division that features regiments from different units (columns 1 and 3) and the count of sub-units assigned to a composite one (columns 2 and 4). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/; Lt Col Robertson et al. (2013); Baldissara (2000); and military German maps collected from Bundes Archiv.

A.2 Sensitivity Checks

A.2.1 Robustness Checks

In this section, we check the robustness of our results to different definitions of treated and comparison municipalities, concentration of German troops and clustering at different levels of aggregation.

Definition of Treatment. In equation 1 we define treated municipalities as those within 40km of the front line, the average extension of the combat zone based on the historical evidence. As the combat zone extension varied across different front lines, we test the robustness of our results to using 30 and 50km as alternative cut-offs for the definition of treatment. The estimates are very similar to our baseline ones, suggesting that the results do not depend on the selection of a particular cut-off to define the combat zone (Tables A.11 and A.12).

Alternative Comparison Groups. One may be concerned that using all municipalities under Axis control as comparison group may not be effective in recovering a valid counterfactual. We show robustness to several alternative comparison groups. First, we limit our sample to the operation zone, which only includes municipalities directly exposed to the war and under the exclusive control of the Wehrmacht, obtaining estimates close in magnitude to our baseline ones (Table A.13). Second, we restrict the sample to municipalities located within 200 and 100 km from the front line on the Axis side to offer an even more stringent comparison between treated and comparison units. While the number of observations drops from more than 2 million in the baseline sample to roughly 150,000, our results are substantially unchanged (Tables A.14 and A.15). Third, we restrict the comparison group to municipalities that were within 40km from the front line both at time t and $t-1$ only. Thus, we effectively leverage those episodes of pivoting of the front lines. If our findings were driven by front proximity and not by its activation, we should not observe any difference in civilian killings between treated municipalities and those in the combat zone both before and after activation. Reassuringly, the estimated coefficients are very similar to the baseline ones, confirming that the sudden movement of the front line - rather than proximity to it - changed the behaviour of German soldiers (Table A.16). Fourth, some municipalities may have been exposed to more than one front shift in either the 5 days prior or after $k = 0$, especially if they were located close to particularly unstable defensive lines. Our results remain virtually unchanged if we exclude them from the sample (Table A.17). Finally, we show that, thanks the inclusion of side-front-day specific trends in the main specification, we can exclude the Allied-controlled part of the country for each front-experiment without changing the point estimates (Table A.18).

Excluding Summer 1944. An important exception in German Command's approach towards civilian

victimization during the conflict is represented by Summer 1944. In an attempt to contain partisans activity, Kesselring issued a series of orders that relaxed the formal constraints placed on his troops while carrying out counterinsurgency operations. This change of approach led to some of the best-documented and famous mass killings episodes of Italian civilians, such as Sant'Anna di Stazzema and Marzabotto, which took place with the more or less explicit consent of Kesselring. Interestingly, these orders were quickly withdrawn and the previous code of conduct towards civilians was re-instated in full already in late September 1944, as a consequence of Mussolini's complaints and the realization that these actions were leading to increasing participation and support for the resistance movement (War Crimes Commission, 1949, p.11), which indirectly confirms the overall committent to limiting violence on civilians in accordance with international law.[33]

Although, controlling for front-experiment fixed effects γ_t allows us to compare treated and comparison municipalities exposed to the same set of rules, if the estimated effect were solely driven by the events of Summer 1944, which coincided with unprecedented levels of military pressure exerted on the German forces, one may question that our results are driven by lower soldiers' accountability. Thus, in Table A.19 we show that our findings are robust to excluding the the front activation episodes between June 17 1944 and September 24 1944, the window of relaxation of the rules in dealing with unarmed civilians.

Including Killings from Fascist Units. Our dataset contains roughly 1,200 episodes of violence perpetrated by Italian Fascist units. Including such episodes in the sample confirms an increase in indiscriminate, but not selective violence upon the front activation, but leads to smaller coefficients than our baseline ones (Table A.20). This confirms that our findings are driven by German troops, more likely to be affected by changes in accountability.

Inference. In the main specification, we account for time-series heteroskedasticity by clustering the standard errors at the municipality level. As we worry that spatial autocorrelation may threaten inference, we provide two alternative specifications. First, in Table A.21, we double cluster at the municipality-province level. Second, in Table A.22, we repeat this exercise at the municipality-nearest division level. In both cases, the double-clustered errors are either the same or even smaller than the municipality-clustered ones, which suggests, if anything, a negative spatial autocorrelation. Finally, we present standard errors double clustered at the municipality-experiment level in Table A.23.

[33]Despite having reversed these orders after 2 months, Kesselring was found guilty of war crimes, among other allegations also due to the mass-murder episodes from Summer 1944, and sentenced to death (immediately commuted to life-sentence) in 1947 by a British Military court in Venice. The other crimes he was sentenced for were the Fosse Ardeatine massacre which he authorized, and the execution of a British commando that had landed in Liguria, in accordance with Hitler's directives but against international law.

Table A.11: ROBUSTNESS CHECK: 30 KM CUT-OFF FOR DEFINITION OF TREATMENT

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated 30km*Day=-2	0.0010	0.0006	0.0003	0.0005	-0.0018	-0.0017
	(0.0013)	(0.0010)	(0.0002)	(0.0015)	(0.0017)	(0.0014)
Treated 30km*Day=0	0.0063***	0.0032**	0.0020**	0.0036*	-0.0009	0.0021
	(0.0021)	(0.0015)	(0.0010)	(0.0019)	(0.0020)	(0.0022)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,790	2,330,790	2,330,790	2,330,340	2,330,790	2,330,340
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation. *Treated 30km* is an indicator for municipalities that were within 30km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.12: ROBUSTNESS CHECK: 50 KM CUT-OFF FOR DEFINITION OF TREATMENT

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated 50km*Day=-2	0.0014	0.0013	0.0004	0.0004	-0.0005	-0.0002
	(0.0015)	(0.0009)	(0.0003)	(0.0013)	(0.0015)	(0.0016)
Treated 50km*Day=0	0.0068***	0.0027**	0.0025***	0.0044**	-0.0007	0.0016
	(0.0020)	(0.0012)	(0.0009)	(0.0018)	(0.0017)	(0.0019)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,790	2,330,790	2,330,790	2,330,340	2,330,790	2,330,340
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation. *Treated 50km* is an indicator for municipalities that were within 50km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.13: ROBUSTNESS CHECK: ONLY MUNICIPALITIES IN OPERATION ZONE

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0003	0.0003	-0.0001	-0.0005	-0.0019	-0.0015
	(0.0015)	(0.0009)	(0.0001)	(0.0014)	(0.0018)	(0.0018)
Treated * Day= 0	0.0083***	0.0023	0.0027**	0.0053**	-0.0011	0.0021
	(0.0024)	(0.0014)	(0.0012)	(0.0021)	(0.0022)	(0.0025)
Mean dep. var.	0.0016	0.0008	0.0004	0.0012	0.0014	0.0013
Observations	601,898	601,898	601,898	601,554	601,898	601,554
Clusters	5,904	5,904	5,904	5,904	5,904	5,904
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities in the operation zone, directly exposed to war and under the exclusive military control of the Wehrmacht. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.14: ROBUSTNESS CHECK: ONLY MUNICIPALITIES WITHIN 200 KM OF FRONT LINE - AXIS SIDE

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0003	0.0001	-0.0000	-0.0005	-0.0014	-0.0011
	(0.0014)	(0.0009)	(0.0002)	(0.0012)	(0.0017)	(0.0017)
Treated * Day= 0	0.0087***	0.0030**	0.0026**	0.0057***	-0.0010	0.0021
	(0.0023)	(0.0015)	(0.0011)	(0.0021)	(0.0021)	(0.0023)
Mean dep. var.	0.0017	0.0007	0.0004	0.0013	0.0014	0.0014
Observations	367,664	367,664	367,664	367,504	367,664	367,504
Clusters	6,040	6,040	6,040	6,040	6,040	6,040
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities within 200km of the front line. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.15: ROBUSTNESS CHECK: ONLY MUNICIPALITIES WITHIN 100 KM OF FRONT LINE - AXIS SIDE

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	-0.0002	-0.0002	-0.0002	-0.0009	-0.0017	-0.0013
	(0.0015)	(0.0010)	(0.0002)	(0.0014)	(0.0017)	(0.0018)
Treated * Day= 0	0.0075***	0.0030**	0.0024**	0.0058***	-0.0013	0.0004
	(0.0022)	(0.0015)	(0.0010)	(0.0021)	(0.0021)	(0.0023)
Mean dep. var.	0.0028	0.0014	0.0007	0.0021	0.0026	0.0025
Observations	158,653	158,653	158,653	158,534	158,653	158,534
Clusters	4,019	4,019	4,019	4,019	4,019	4,019
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities within 100km of the front line. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.16: ROBUSTNESS CHECK: ONLY MUNICIPALITIES WITHIN 40 KM OF FRONT LINE - AXIS SIDE

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0020	0.0003	0.0000	-0.0020	-0.0025	0.0014
	(0.0022)	(0.0019)	(0.0007)	(0.0020)	(0.0026)	(0.0029)
Treated * Day= 0	0.0082***	0.0038*	0.0017	0.0042	-0.0030	0.0001
	(0.0025)	(0.0020)	(0.0011)	(0.0028)	(0.0026)	(0.0023)
Mean dep. var.	0.0035	0.0020	0.0010	0.0029	0.0033	0.0030
Observations	66,495	66,495	66,495	66,430	66,495	66,430
Clusters	2,404	2,404	2,404	2,404	2,404	2,404
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities within 40km of the front line. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.17: ROBUSTNESS CHECKS: EXCLUDING MUNICIPALITIES WITH MULTIPLE TREATMENTS

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0006	0.0005	0.0003	0.0000	-0.0015	-0.0013
	(0.0013)	(0.0008)	(0.0003)	(0.0012)	(0.0015)	(0.0016)
Treated * Day= 0	0.0081***	0.0031**	0.0027***	0.0058***	0.0000	0.0021
	(0.0021)	(0.0014)	(0.0010)	(0.0019)	(0.0019)	(0.0021)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,321,495	2,321,495	2,321,495	2,321,051	2,321,495	2,321,051
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, excluding municipalities that were treated multiple times within a front-experiment due the front lines shifting multiple times within a short time-frame. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.18: ROBUSTNESS CHECK: ONLY MUNICIPALITIES ON AXIS-SIDE

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0006	0.0005	0.0003	0.0000	-0.0015	-0.0013
	(0.0013)	(0.0008)	(0.0003)	(0.0012)	(0.0015)	(0.0016)
Treated * Day= 0	0.0081***	0.0031**	0.0027***	0.0058***	-0.0000	0.0020
	(0.0021)	(0.0014)	(0.0010)	(0.0019)	(0.0019)	(0.0021)
Mean dep. var.	0.0006	0.0002	0.0002	0.0004	0.0005	0.0005
Observations	1,639,935	1,639,935	1,639,935	1,639,695	1,639,935	1,639,695
Clusters	6,939	6,939	6,939	6,939	6,939	6,939
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, restricting the sample to municipalities in the Axis-controlled territory. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** $p < 0.01$, ** $p < 0.5$, * $p < 0.1$.
Source: http://www.straginazifasciste.it/

Table A.19: ROBUSTNESS CHECK: EXCLUDING SUMMER 1944

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0013	0.0012	0.0004	-0.0006	-0.0020	0.0001
	(0.0011)	(0.0009)	(0.0003)	(0.0011)	(0.0014)	(0.0013)
Treated * Day= 0	0.0040**	0.0034**	0.0011*	0.0040**	-0.0000	-0.0003
	(0.0016)	(0.0014)	(0.0007)	(0.0019)	(0.0018)	(0.0014)
Mean dep. var.	0.0005	0.0003	0.0002	0.0003	0.0004	0.0004
Observations	1,873,806	1,873,806	1,873,806	1,873,547	1,873,806	1,873,547
Clusters	7,341	7,341	7,341	7,341	7,341	7,341
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, excluding front movements happened between June 17 1944 and September 24 1944, when Kesselring gave more freedom to his troops in dealing with civilians. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.20: ROBUSTNESS CHECK: INCLUDING KILLINGS BY FASCIST SOLDIERS

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0001	0.0009	0.0003	0.0000	0.0000	-0.0002
	(0.0015)	(0.0009)	(0.0003)	(0.0013)	(0.0017)	(0.0018)
Treated * Day= 0	0.0071***	0.0030**	0.0027**	0.0055***	0.0001	0.0015
	(0.0022)	(0.0014)	(0.0010)	(0.0020)	(0.0019)	(0.0021)
Mean dep. var.	0.0007	0.0003	0.0002	0.0005	0.0006	0.0006
Observations	2,330,790	2,330,790	2,330,790	2,326,786	2,330,790	2,326,786
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation, including in the sample killings perpetrated by Italian Fascist units. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** $p < 0.01$, ** $p < 0.5$, * $p < 0.1$.
Source: http://www.straginazifasciste.it/

Table A.21: ROBUSTNESS CHECK: CLUSTERING AT MUNICIPALITY-PROVINCE LEVEL

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0006	0.0005	0.0003*	0.0000	-0.0015	-0.0013
	(0.0013)	(0.0008)	(0.0002)	(0.0012)	(0.0015)	(0.0017)
Treated * Day= 0	0.0081***	0.0031**	0.0027**	0.0058*	-0.0000	0.0020
	(0.0022)	(0.0013)	(0.0012)	(0.0031)	(0.0019)	(0.0019)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,790	2,330,790	2,330,790	2,330,340	2,330,790	2,330,340
Clusters 1	7,063	7,063	7,063	7,063	7,063	7,063
Clusters 2	92	92	92	92	92	92
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality-province level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.22: ROBUSTNESS CHECK: CLUSTERING AT MUNICIPALITY-NEAREST DIVISION LEVEL

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0006	0.0005	0.0003	0.0000	-0.0015	-0.0013
	(0.0010)	(0.0008)	(0.0003)	(0.0010)	(0.0016)	(0.0015)
Treated * Day= 0	0.0081***	0.0031*	0.0027**	0.0058**	-0.0000	0.0020
	(0.0022)	(0.0016)	(0.0013)	(0.0027)	(0.0027)	(0.0018)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,790	2,330,790	2,330,790	2,330,340	2,330,790	2,330,340
Clusters 1	7,063	7,063	7,063	7,063	7,063	7,063
Clusters 2	43	43	43	43	43	43
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality-nearest division level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.23: ROBUSTNESS CHECK: CLUSTERING AT MUNICIPALITY-EXPERIMENT LEVEL

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day= -2	0.0006	0.0005	0.0003	0.0000	-0.0015*	-0.0013
	(0.0012)	(0.0008)	(0.0004)	(0.0008)	(0.0009)	(0.0016)
Treated * Day= 0	0.0081***	0.0031**	0.0027**	0.0058***	-0.0000	0.0020
	(0.0027)	(0.0015)	(0.0010)	(0.0017)	(0.0011)	(0.0017)
Mean dep. var.	0.0006	0.0003	0.0002	0.0004	0.0005	0.0005
Observations	2,330,790	2,330,790	2,330,790	2,330,340	2,330,790	2,330,340
Clusters 1	7,063	7,063	7,063	7,063	7,063	7,063
Clusters 2	30	30	30	30	30	30
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1 and reported for two days before and the day of front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality-experiment level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

A.2.2 Robustness to Alternative Estimators

As discussed in Section 5, estimating the effects of the front activation on civilian victimization presents two potential issues in our setting. First, as the we only observe civilian killings committed by Axis soldiers, the movement of the front line mechanically alters the relevant estimating sample and comparison group. This issue makes using a regular panel for estimation problematic, as it would not allow to properly account for changes to the relevant estimating sample that are specific to each episode of front line activation. Second, a simple Two-Way Fixed Effects model would not be able to deal with the staggered treatment adoption present in our sample, which may lead severe estimates bias, particularly in the presence of heterogeneous treatment effects (Roth et al., 2023).

For these reasons, in equation 1, we estimate a stacked difference-in-differences model. By building separate artificial panels that are then combined to form a concatenated stacked panel, this methodology allows to treat each episode of front line activation as a separate event. On the one hand, this allows to solve the issue of mechanical changes in the estimating sample, by allowing the inclusion of experiment-specific, side of the front trends that capture the mechanical decline of killings on the Allied side of the front after each front movement. On the other, by saturating the model with experiment-specific fixed effect, we can recover the Average Treatment Effect (ATT) as the (equally) weighted average of each treatment (Baker et al., 2022).

While we believe our proposed approach to be the most appropriate in our setting, we also document the robustness of our estimates to a more mainstream estimator that is robust to staggered treatment rollout and heterogeneous treatment effects. As our treatment is non-absorbing (i.e. front line activation is a 1-day event), we rely on the estimator developed by De Chaisemartin and d'Haultfoeuille (2024). This allows to model the treatment variable as non-absorbing (to switch on and off), while accounting for dynamic effects. Reassuringly, results from this estimator are similar to the ones from the staggered DiD design (Figure A.6).

While not a reliable estimator in our setting due to the challenges discussed above, it is important to show we can recover the treatment effect with a simpler Two-Way Fixed Effects estimator for robustness, as we have a non-absorbing treatment with limited (i.e. short-lived) dynamic effects. The TWFE confirms our main findings and produces slightly larger estimates (Table A.24, column 1).[34] This is due to the fact that, without being able to properly account for side-front specific trends (we are only able to introduce side of front as a control in Table A.24) as the front moves North, some municipalities South of the front

[34] For these less relevant estimators we only replicate column 1 of Table 2

in the control group will have no killings by construction. Therefore, the TWFE estimates should be interpreted as an upper bound of the impact of front activation on civilian victimization. Our main outcome variable is an indicator for episodes of indiscriminate violence. Given its binary nature, we also estimate a logit specification, whose results are in line with our baseline ones. More specifically, the logit estimates indicate a 73-percent increase in the probability of indiscriminate killing episodes (Table A.24, column 2).

In our data, episodes of civilian victimization are rare. The probability of observing a killing episode is less than one out of 1,000 people. We therefore estimate a rare event logistic regression, following King and Zeng (2001).[35]. The results are larger than the estimates from the logit model, which indicates that our findings are not driven by rare events (Table A.24, column 3). As our events are not just rare, but also zero-inflated due to the presence of several municipalities under Allied control, we estimate a Zero-Inflated Poisson regression. The estimates are robust to this additional check (Table A.24, column 4).

Finally, the fact the killings episodes are rare may raise the concern that our results are driven by outliers. In Table A.25, we show that our stacked panel results are robust to collapsing days within an experiment by larger 3 days time-units, which allows to estimate only one pre-trend and two post-treatment period, each containing a higher number of episodes.[36]

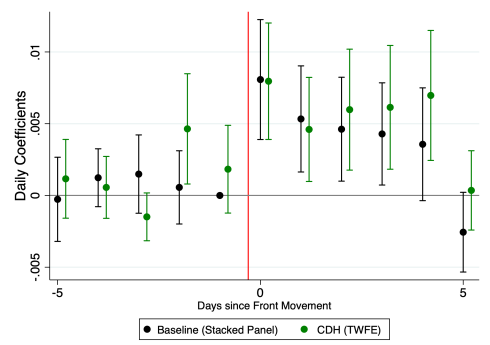
[35]More specifically, we use the Stata package `relogit` that estimates the same logit model as the `logit` command, but with an estimator that gives lower mean square error in the presence of rare events data for coefficients.

[36]This procedure collapses together k-5 and k-4; k-3, k-2 and k-1 (reference period); k0, k+1 and k+2; k+3, k+4 and k+5.

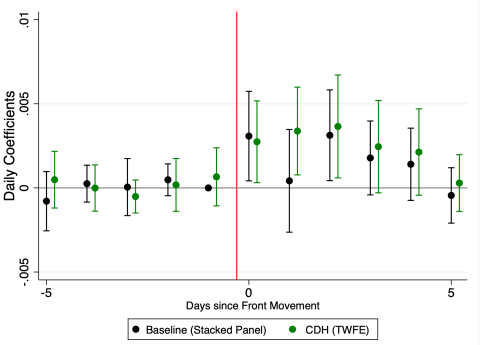
Figure A.6: TWFE PANEL ESTIMATION – DE CHAISEMARTIN AND D'HAULTFOEUILLE (2020)

Indiscriminate Violence

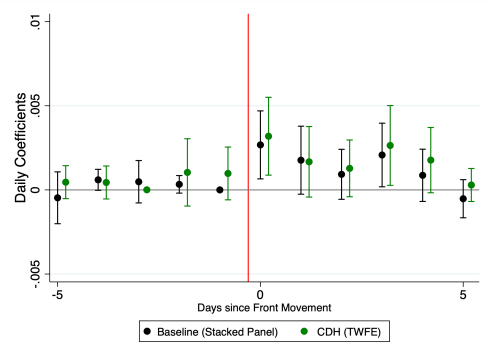
(a) Collective Killing



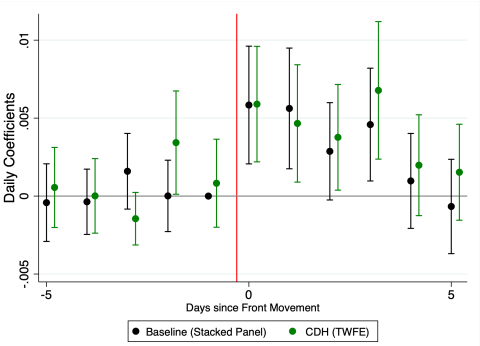
(b) No Partisan Killing



(c) Vulnerable Killing

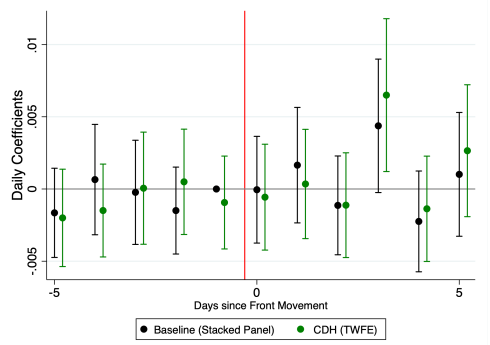


(d) Hand-Coded Indiscriminate Killing

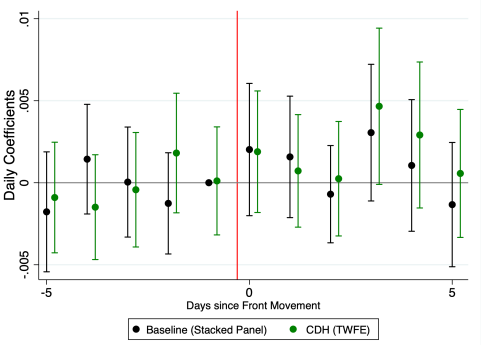


Selective Violence

(e) Single Killing



(f) Hand-Coded Selective Killing



Notes. Daily coefficients and 95% standard errors estimated from the stacked difference-in-differences equation 1 (black dots) and twoway FE panel estimates following De Chaisemartin and d'Haultfoeuille (2020) (CDH, green dots) from 5 days before to 5 days after front activation. The estimations of the twoway FE is implemented using the Stata commands `didmultiplegt`. Day $k = -1$ is the excluded category. Panels A-D measure indiscriminate violence; Panels E-F selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (Panel A); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (Panel B); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, Panel C); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (Panel D); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (Panel E); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (Panel F). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

Table A.24: ALTERNATIVE ESTIMATORS

Dependent variable:	Collective Killing			
Model:	TWFE	Logit	Rare Logit	ZIP
	(1)	(2)	(3)	(4)
Treated	0.0091***	2.1792***	2.9529***	2.9269***
	(0.0019)	(0.1983)	(0.1847)	(0.1828)
Axis side	0.0011***	1.5897***	0.7415***	
	(0.0001)	(0.1510)	(0.0740)	
Axis side				-16.2812***
				(0.1855)
Mean dep. var.	0.0005	0.0026	0.0005	0.0005
Observations	4,654,517	823,750	4,654,517	4,654,517
Clusters	7,063	1,250		7,063
Municipality FE	YES	YES	NO	NO
Day FE	YES	NO	NO	NO

Notes. Coefficients estimated from a Two-Way Fixed Effect (TWFE) model (column 1), a logit regression (column 2), a rare event logistic regression following King and Zeng (2001) (column 3), and a Zero-Inflated Poisson (ZIP) regression (column 4). *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. *Axis side* is an indicator for municipalities under the Axis control. *Collective Killing* is an indicator for episodes of collective killings where at least two civilians were killed by German soldiers. Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1. *Source*: http://www.straginazifasciste.it/

Table A.25: ROBUSTNESS CHECK: COLLAPSING DATA TO 4 DAYS EXPERIMENTS

Dependent variable:	Indiscriminate Violence				Selective Violence	
	Collective Killing (1)	No Partisan Killing (2)	Vulnerable Killing (3)	Hand-Coded Killing (4)	Single Killing (5)	Hand-Coded Killing (6)
Treated * Day=-5 to -4	-0.0028	-0.0016	-0.0007	-0.0034*	-0.0022	-0.0026
	(0.0022)	(0.0014)	(0.0011)	(0.0020)	(0.0025)	(0.0027)
Treated * Day=0 to 2	0.0138***	0.0086***	0.0047***	0.0125***	0.0025	0.0033
	(0.0034)	(0.0025)	(0.0017)	(0.0031)	(0.0029)	(0.0030)
Treated * Day=3 to 5	0.0025	0.0018	0.0017	0.0036	0.0048	0.0033
	(0.0031)	(0.0019)	(0.0015)	(0.0029)	(0.0032)	(0.0033)
Mean dep. var.	0.0016	0.0007	0.0005	0.0011	0.0013	0.0012
Observations	847,560	847,560	847,560	847,541	847,560	847,541
Clusters	7,063	7,063	7,063	7,063	7,063	7,063
Municipality FE	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES
Front FE	YES	YES	YES	YES	YES	YES
Front x Municipality FE	YES	YES	YES	YES	YES	YES
Front x Day FE	YES	YES	YES	YES	YES	YES
Front x Day x side FE	YES	YES	YES	YES	YES	YES

Notes. Daily coefficients estimated from equation 1, collapsing the data into 4 and 5 days before front activation, 3 to 1 days before front activation (excluded reference period) 0 to 2 days after the front activation and 3 to 5 days after front activation. *Treated* is an indicator for municipalities that were within 40km of the combat zone at time t but were outside it at time $t-1$. Columns 1-4 measure indiscriminate violence; columns 5-6 selective violence. *Collective Killing* is an indicator for episodes of collective killings, where at least two civilians were killed by German soldiers (column 1); *No Partisan Killing* is an indicator for episodes of collective killings by German soldiers not related to partisan activity (column 2); *Vulnerable Killing* is an indicator for episodes of collective killings by German soldiers that involved at least a vulnerable individual (woman, child, and elderly, column 3); *Hand-Coded Indiscriminate Killing* is an indicator for episodes of indiscriminate killings by German soldiers, based on the Atlas description (column 4); *Single Killing* is an indicator for episodes of single killings, where a single civilian was killed by German soldiers (column 5); *Hand-Coded Selective Killing* is an indicator for episodes of selective killings by German soldiers, based on the Atlas description (column 6). Standard errors are clustered at the municipality level. *** p < 0.01, ** p < 0.5, * p < 0.1.
Source: http://www.straginazifasciste.it/

B Data Appendix

This Data Appendix describes the data collection procedures behind our paper more in detail. It presents the sources, their coverage, and the way we selected them.

B.1 Civilian Killings

The first step of the data collection aimed at gathering data on civilian killings perpetrated by Axis soldiers during the Italian Campaign. We collected these data from the *Atlante delle Stragi Naziste e Fasciste in Italia* (Atlas of the Nazi and Fascist Massacres in Italy), available at http://www.straginazifasciste.it, and accessed in November 2020. The Atlas "lists and analyzes all the massacres and the individual murders of civilians and resistance fighters killed during the Italian Campaign both by German soldiers and soldiers of the Italian Social Republic outside of the armed fights between July 12, 1943 and May 2, 1945".

For most killing episodes, the Atlas reports its location and date, the number of victims, their gender, age and fighting status (whether civilians or partisans), their identity if known, the killing mode, and whether the perpetrators were Nazi, Fascists or both acting together. For a few episodes, whether the killings were associated to other forms of violence (for instance, robberies, looting, tortures and rapes) is also stated. Finally, 24 episodes for which the killing date is unknown or for which there are disagreements among the sources are excluded from the online database and therefore from our analysis.

The Atlas was commissioned by the German Foreign Ministry in collaboration with the Italian government through the creation of a "German-Italian Fund for the future" and was promoted in conjunction with the INSMLI (National Institute for the History of the Italian Liberation Movement" and the ANPI (National Association of Italian Partisans).

The historical inquiry was conducted locally by more than ninety researchers who tapped in the results of previous studies of the same kind made in Apulia, Campania, Tuscany, Emilia Romagna and Piedmont and used three main national common sources: 1) the data base of violent crimes perpetrated against civilians during the German occupation of Italy, established by the Joint Historical Italian-German Commission, based on the reports of the police reports stored in the Archives of the Historical Office of Army General Staff and the Historical Archives of the Carabinieri of Rome. 2) the General Repository of war crime reports collected from 1945 by the Army Prosecutor's office in Rome (dismantled illegally in 1960) that was found by the Parliamentary Commission of Inquiry while investigating on the reasons for the concealment of some files about Nazi-fascist crimes (14th Parliamentary term). 3) the rulings and files of the judiciary proceedings debated at the Military courts during the last trial season (from 1994 until

now).

According to the Atlas, between July 15, 1943 (the first day in which our data report a victim) and May 2, 1945 (when WWII ended on the Italian front), 24,988 Italian civilians were killed by Axis soldiers. This figure represents the 17.1% of the 145,996 Italian civilians that died in Italy due to war-related causes between 1940 and 1945 (ISTAT, 1957)[37]; and roughly 20% of the victims after the signature of the Italian Armistice.

As our unit of analysis is at municipality-day level, we collapse the data accordingly, obtaining 5,291 episodes of violence against Italian civilians over 700 days of war. Figure B.1, Panels A and B show the distribution of killing episodes across Italy. Those were concentrated in the central and northern part of the country, where fighting between German and Allied troops was more intense and the occupation prolonged. By contrast, Sicily and Southern Italy registered only a few episodes.[38]

On average, each episode registered 4.7 victims, 3.8 male and 0.6 female. While the great majority of episodes involved either a single (47% of the cases) or between 2 and 10 causalties (45%), three of them were particularly bloody (Figure B.1, Panel B). The most violent massacre happened in Marzabotto near Bologna on September 29, 1944, when 1,012 civilians were killed. The other two episodes were Sant'Anna di Stazzema near Lucca on August 8, 1944 and the "Fosse Ardeatine" massacre near Rome on March 24, 1944, when respectively 391 and 335 civilians were murdered. Sadly, 1,032 (4.2%) of all the victims were children.

In terms of the perpetrators, German soldiers, either alone or in collaboration of Fascist units, were responsible for the vast majority of killings (73.4%). Fascist soldiers alone committed 22.7% of murders, while for the a few remaining episodes the identity of the perpetrators could not be determined (3.9%). As explained in Section 4.1, most of our analysis focuses on atrocities committed by German units.

While in principle the Atlas reports if killings were committed along with other crimes, only 276 episodes (5.2%) do so. Almost all other forms of violence were robberies and looting, while only 5 instances describe rapes. This feature of the data prevents us from studying violence against women as a complementary indicator of civilian victimization. Finally, it is worth noting that mere 4 descriptions contain references to race-related crimes, confirming how violence against civilians during the Italian Campaign, and unlikely most civil war contexts, was not driven by racial motivations.

Does the Atlas cover (close to) the universe of *documented* Axis civilian killings during the Italian Campaign? According to the authors, the careful data reconstruction and the collaboration with both German

[37] These include collateral victims from war targeted killings.
[38] In the island Sardinia there was no fighting after the invasion of Sicily and so no killing episodes.

and partisan associations made possible to "paint a complete picture of the violence perpetrated against civilians by the German army and its allied Fascists in Italy between 1943 and 1945." The collaboration across Italian and German researchers and partisan associations likely limited incentives to mis-report the episodes. Moreover, the fact that all the episodes have been collected with the same methodology and from the same data sources ensures the application of systematic criteria in compiling the Atlas. Based on this, we believe that the vast majority of available sources have been searched for and included in the Atlas.

A different set of concerns relates to whether all killings episodes were recorded or reported during the war or immediately after that. While it is of course impossible to know for sure if this is the case, the significant presence of small-scale episodes involving only 1 civilian, reassures with respect to potential truncation of the sample. Moreover, in terms of selection into sample, we only worry in terms of it being differential between treatment and comparison units. In this respect, it is worth noting that the Atlas' sources were mostly based on either coeval reports by the Italian police, or on reports from civilians and inquires by Italian authorities after the end of the war.

Regarding coeval information, the fact that the reports were collected by the Italian police and not by German authorities reduces the probability of more reporting closer to German units due to higher availability of personnel, an instance likely to happen in the proximity of the front. If anything, Italian police, which was subordinated to the German authority as Italy was an occupied territory (Collotti, 1963b), was less likely to be present in the combat zone. Moreover, Italian police has less incentives to hide violent episodes than the German authorities.

Regarding reports submitted after the end of the war, relatives of the victims had strong incentives to report as they were eligible to receive a war pension.[39] Importantly, these pensions were granted regardless of the identity of the perpetrators, reducing the incentives to attribute to the Germans crimes committed by the Allies, and also to report differentially between treatment and comparison units. As discussed, by itself, the fact that about 50% of our episodes involved only a single victim is consistent with strong incentives to report and makes data truncation due to episode size unlikely.

All together, these considerations suggest that the Atlas is a comprehensive source of Axis killings during the Italian Campaign, and that differential reporting is unlikely.

[39]More specifically, the "Legge 10 Agosto 1950, n. 648, art. 10" states, among other eligibility criteria for living individuals, that war pensions are conferred to the families of Italian citizens died for war-related fatalities perpetrated by national or foreign Armies, allied or enemies. See: $http://www.edizionieuropee.it/law/html/41/zn77_06_050.html$

B.2 Indiscriminate and Selective Violence

Disentangling episodes of indiscriminate violence from those of selective violence is crucial in our paper. As explained by Kalyvas (2006), the two types of violence differ in terms of the motivation and mechanisms through which civilian victimization happens. More specifically, selective violence reflects the willingness of the perpetrators to use violence against civilians in a manner proportionate to the context-specific war objective and to ascertain individual responsibilities. By contrast, indiscriminate violence goes beyond the actual strategic needs of the occupying army, is thus disproportionate, and has a unjustified, "random" component. Empirically, as the intentions of the perpetrators cannot be observed, the difference between selective and indiscriminate violence has been mapped onto the distinction between single and collective killings (Kalyvas, 2006, p. 142).

Several examples from the Atlas are fully consistent with this distinction. Selective single killings were the norm rather than the exception. This fact indirectly supports the idea that German soldiers often exerted restraint without automatically resorting to indiscriminate violence, in line with the engagement rules set out by Kesselring, and even in areas exposed to partisan operations, as shown by this example:

> "[...] The partisan group "Di Vincenzo" operated in the area. A German patrol engaged in anti-partisan operations, bumped into five young men, who fled under fire. In the village of Fugnette the five men were captured by a second patrol. One of these, Franco Gambacurta, suspected to be a resistance sniper, was separated from the other three and executed at Capolaforca. The other men were set free."[40] *Source: http://www.straginazifasciste.it*, L'Aquila, September 25, 1943.

Conversely, small-scale collective killings often reflected an application of indiscriminate violence, in breach of the formal rules of engagement:

> "On 18 October 1943, in the village of Caira, near Monte Cairo, a German patrol tried to steal a cow from a local farmer, Felice Nardone, who resisted the violence shooting with a rifle against the soldiers. The soldiers, having escaped, came back with reinforcements. Unable to find Nardone, they stop two other locals Mariani and Nardoni to help them find the aggressor. Unable to find him, they kill Mariani with a pistol shot, while Nardoni was strapped to a

[40] Authors' translation from: "Nella zona operava la banda partigiana "Di Vincenzo". Una pattuglia di soldati tedeschi, in un'operazione di rastrellamento, si imbatté nei pressi di Filetto in cinque giovani, che, fatti oggetto di colpi d'arma da fuoco, si dettero alla fuga. In località Fugnette i cinque vennero catturati da una seconda pattuglia. Uno di questi, Franco Gambacurta, sospettato di essere un franco tiratore, fu separato dagli altri e passato per le armi a Capolaforca. Gli altri giovani furono liberati".

truck and dragged for a few kilometers. He died a few days after for the wounds."[41] *Source: http://www.straginazifasciste.it*, Cassino (Frosinone), October, 18 1943.

Therefore, our first measure of indiscriminate violence follows Kalyvas (2006) and is proxied by an indicator for episodes with at least two victims; by contrast, selective violence is proxied by an indicator for episodes with a single victim. Their occurrence in our data is quite similar. Out of 5,291 episodes of violence from the Atlas, those indicative of indiscriminate violence are 53% (2,804, Figure A.1), while those of selective violence 47% (2,487). Focusing on killings perpetrated by German soldiers only, the total episodes are 3,885: 56% (2,167) of indiscriminate violence and 44% (1,718) of selective violence.

The richness of our data allows us to refine these measures in several ways. Out of 5,291 episodes from the Atlas, 5,118 (96%) report a description. We exploit such detailed narrative to search for terms that refer to resistance fighters and activity. Namely we search for the word "partigian", which captures all combinations of the words indicating resistance fighters in Italy during WWII. This methodology suggests that roughly half (2,709 events) of killings happened in a context where some partisan activity was in place. We then construct an indicator for collective killings *not* related to partisan activity, whose nature is more likely to be indiscriminate than related to the actual safety of German troops. Moreover, we use victim characteristics to code an indicator for episodes that involved vulnerable population (at least one woman, child or elderly), unlikely to be a threat for the German army or its strategic goals. This represents 12% of all episodes and 16% of the killings by German soldiers.

In addition to that, we use the description of each episode to hand-code episodes of indiscriminate and selective violence. Even if the 96% of the episodes report a description, in a few cases, it is too short, too generic or uninformative to recover the intention of the perpetrators. Thus, we are were only able to hand-code indiscriminate or selective violence for 4,586 events (86.7%).

More specifically, we hand-coded as indicative of indiscriminate violence those episodes that clearly indicated: targeting of innocent people; lack of any formal trial procedure; impulsive and unjustified killings; involvement of vulnerable people not reported to have openly engaged in hostile activities towards the Germans; involvement of war prisoners that were protected under martial law; targeting of an innocent relative of a guilty individual; reporting of sexual violence alongside killing; collateral damage of direct targeting of civilian houses (i.e. locked and burned in a house).

By contrast, we considered selective violence those cases in accordance with the rules of engagement:

[41] Authors' translation from: "Il 18 ottobre 1943, nella frazione di Caira, ai piedi di Monte Cairo, una pattuglia tedesca cerca di rubare una mucca al contadino della zone Felice Nardone, che resiste alla violenza sparando con un fucile da caccia contro i soldati. Questi, scampati ai colpi, si allontanano per tornare poi con i rinforzi. Non trovando Nardone, fermano due abitanti del posto, Mariani e Nardoni, per trovare l'aggressore. Dinanzi all'inutilità della ricerca, Mariani viene ucciso con un colpo di pistole mentre Nardoni, legato ad un camion e trascinato per qualche km, muore due giorni dopo in seguito alle ferite riportate."

killings that involved people that refused compulsory labor services for the Germans; execution of civilians that had attacked German soldiers; killings of civilians suspected to be part of a partisan organisation, deserters, or that openly opposed requisitions; execution for having committed a crime (whether observed or reported); killings of men refusing to report for conscription; killings of people that refused to obey evacuation orders or that refused to obey a German command. Out of 4,321 killings, roughly 45% are coded as indicative of indiscriminate violence and 55% of selective violence.

B.3 Front Lines Locations, Activation and Movements

We geo-localize the position of Axis defensive line between July 9, 1943 and May 2, 1945. We collected this information from the secondary literature (Blumenson, 1993; Short and Taylor, 2006; Fisher, 1993; Bishop, 2005; Garland et al., 1993). We then geocoded and digitized each cartographic material in GIS, creating a series of shapefiles. The military historians' reconstruction of these lines is facilitated by their tendency to follow natural features such as rivers and mountain ridges and by the significant survival of remains of the fortifications. Therefore, we are confident to be able to locate all the major front lines that were active at any point of the campaign across the Italian peninsula. However, we acknowledge that smaller adjustments to the front may not be captured by our data. The same sources indicate the date of activation of each front line, defined as the the day of the inception of fighting along that line. This ensures full consistency between position and activation of front lines. After geo-referencing each map in GIS and creating a set of shapefiles, we compute the Euclidean distance between each municipality centroid to the closest vertex of each front line.

We cross-check the accuracy of front line data with the position of the fronts from military maps compiled by the Wehrmacht during the Italian Campaign (collected from the BundesArchiv in Berlin, accessed in January 2022). These maps, whose an example is shown in Figure B.2, were compiled by the Supreme Military Command of the Wehrmacht (OKM).[42] In principle, one advantage of using these maps is that they have a consistent level of resolution and are compiled with an average frequency of 3 days by the Wehrmatch throughout the Italian Campaign. However, they do not contain information on the front line activation. Moreover, it is not always clear whether the reported location of the front line was its *intended* position or the actual one. In other words, while we may improve on frequency, the actual location contains potential imprecisions. For this reason, we decided to rely on the maps compiled by the secondary literature. Repeating our analysis using the German maps, however, leads to very similar results com-

[42]More specifically, such maps belong to the series "Sud" and "Italien un Balkan" and were produced by the General Staff of the German Army (OKH) between 1943 and 1945.

pared to our baseline ones, which indicates that their higher accuracy would have a minimal impact on our findings (results available on request).

B.4 German Units

We use the German military maps described in Section B.3 to recover the locations of German and Allied units over the entire duration of the Italian Campaign. We select maps that portray the deployment of troops at each episode of front movement (see, for an example, Figure B.2). We complement these data with information on the deployment of troops in the first available date before the activation of a new front line, to trace changes in the distribution and composition of Axis units when fronts are activated. This information is usually available between -3 and +3 days before and after front activation. We successfully match pre-post maps with 19 of our front line activation episodes. A key advantage of these maps is that they provide an overview of all units deployed in Italy: those actively fighting at the front, those in reserve as well as those deployed elsewhere with occupation duties. These pieces of information are not available in zoomed-in versions of military maps that exist only for the frontline area. We georeference each map and create a troops shapefile per map. From these maps we record, for each unit, name, location, the number of sub-units from a different division that were attached to it at any given point (what we use to define "composite units"), whether a sub-unit was isolated from its main division, and whether it was incomplete (either in formation or under refurbishment).

One potential concern with the data extracted from these maps is their level of precision in locating units' headquarters as well as troops positions. The majority of the units reported on the maps are divisions, made up of between 10 to 20,000 men. In most cases their headquarters' location would be clearly marked as a flag, in others the unit's number would be printed on the map. We take the approach of recording a unit's location by placing a dot in the middle of the name in the absence of a flag. We assume that unit's troops would be located near the flag or the name of that particular unit. As we mostly use these data for nearest unit analysis, we think of this as a reasonable approximation. However, we also cross-validate the precision of the locations for a handful of units at the front by comparing the locations of maps form the "Sud" and "Italien un Balkan" series with those from a more precise one called "Italien", which were likely produced by the German command in Italy and only covers the front's sector. The match across series for proximate dates is remarkable with minimal spatial error. One good example of this is the comparison between map IB RH 2-KART/9472 (22 October 1943) from the "Italien un Balkan" series and map I RH 2 KART/9641 (27 October 1943) from the "Italien" series. The location of units at

the front in the larger "Italien un Balkan" series' map very closely matches the one shown in the more granular "Italien" series' map.

As some maps attempt to show the location of German sub-units in a more granular manner, we adopt the criterion of only recording the location of German units that are named. These are typically detached regiments and battalions. For those few examples where the presence of small military units is reported without a name, we implicitly assume these were small sub-units of the nearest largest named unit and we do not record them separately. We complement information on each unit that we gathered from these maps with additional data on the background characteristics of these units. At the division level, we gather information on the place and date of formation of each division, its last deployment before the Italian campaign, whether it was restructured or re-formed before (or during) the Italian campaign with the date, as well as the period of deployment in Italy. These data are collected from online repositories, mostly from the Lexicon Der Wehrmacht, which provides complete information and references for each unit.[43] Data of this latter type are time-invariant, as they do not change over our study period.

B.5 Partisan Activity

We measure partisan activity in two ways. First, using the description in the Atlas, we create an indicator for episodes of violence that contain the word "partigian" in their description (used in the balancing tests of Table A.1). While this variable can be used as a proxy for partisan presence, it is an imprecise one: at least a killing should take place for partisan presence to be reported.

Therefore, we complement our information on partisan activity by geo-localizing the position of the Italian partisan bands in September 1943 *Historical Atlas of Italian Resistance* (Baldissara, 2000). This Atlas gathers documentation stored at more than 60 local institutes for the history of the Resistence between 1943 and 1945 (Istituti Territoriali per la Storia della Resistenza), under the supervision of the Italian Ministry of Culture. Overall, 184 partisan bands were active during the Italian Campaign. Most of them were concentrated in the central part of Italy and in the North-Eastern and Western part of the country, in the last phase of war. We use this information to proxy war-related stress of German soldiers, by implementing an exposure design that splits our sample based on proximity of partisan bands' (see Section 7.2 and Table A.9).

[43] https://www.lexikon-der-wehrmacht.de/, last accessed January 2022.

B.6 Allied Bombing during the Italian Campaign

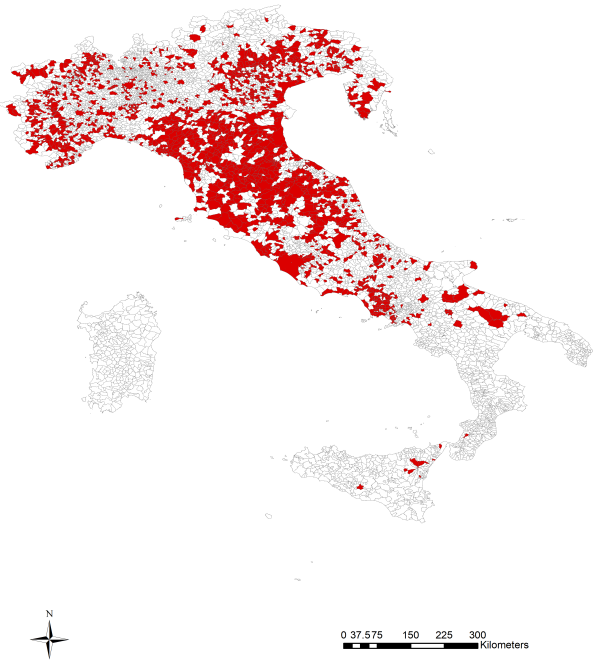
We retrieved data on Allied bombing in Italy from the Theater History of Operations Reports (T.H.O.R.; Lt Col Robertson et al., 2013), compiled by the US Air Force Research Institute, that we collected from www.afri.au.af.mil/thor, in January 2017. The database reports municipality, date, type of target and tons of explosives for each Allied air strikes during WWII. Between June 11, 1940 and the end of conflict, Italy was targeted by 22,325 air attacks that hit 877 different municipalities, for a total of 402,045 tons of explosives dropped. 21 percent of the attacks and 18 percent of the explosives were concentrated in the Italian Campaign (Figure B.3). In this phase of the war, the attacks hit the areas around the fronts more heavily. As explained by Bianchi and Giorcelli (2023), in this phase of the war, the Allies used bombing in a tactical manner, to facilitate ground operation of their troops or block Germans' retreat. By contrast, in the first three years of the war, bombing was used to weaker the enemies' production capacity and depress the population's morale, targeting the most industrialized and populated areas of the country. We construct indicators for Allied bombing at the front line activation to proxy soldiers' insecurity during battles (see Section 7.2).

B.7 1936 Population Census

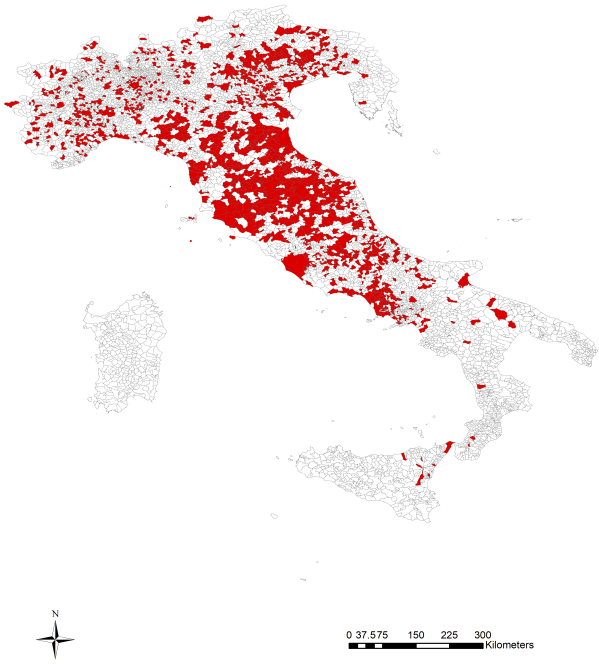
We digitized the 1936 Population Census, that provides data for the 7,341 Italian municipalities existing at that time on population, total employment, employment by sector (agriculture, industry, and services), inactive population, municipality borders and surface. The resident population amounted to 43,059,372 people. Out of 18,368,193 employed individuals, 48 percent worked in agriculture, 29 percent in industry and 23 percent in services. We also use the administrative documentation attached to the census to create a novel shapefile of Italian municipalities with 1936 boundaries. We then aggregate information on killings, front-lines and their movements, partisan activity and bombing at the municipality level, that is our unit of analysis.

Figure B.1: CIVILIAN VICTIMIZATION DURING THE ITALIAN CAMPAIGN

(a) Collective Killing



(b) Single Killing

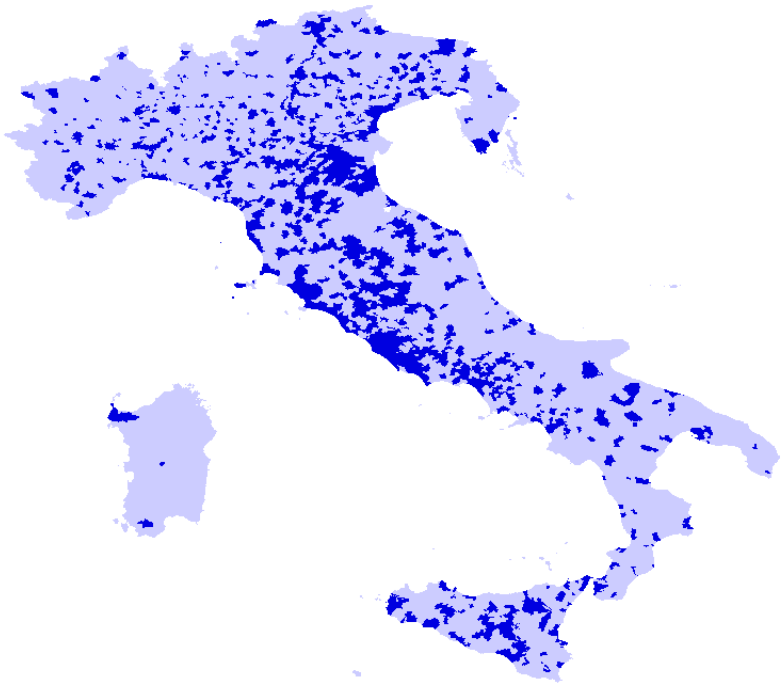


Notes. Panel (a) shows all Italian municipalities that experienced at least one episode of collective killing where two ore more civilians were killed by Axis soldiers, between July 15, 1943 and May 2, 1945. Panel (b) shows all Italian municipalities that experienced at least one episode of single killing where only one civilian was killed by Axis soldiers, between July 15, 1943 and May 2, 1945. Municipalities are drawn using the 1936 borders. *Source*: http://www.straginazifasciste.it/

Figure B.2: EXAMPLE OF MAP SHOWING GERMAN DIVISIONS POSITION – SEPTEMBER 9, 1943

Notes. Map prepared by the Wehrmacht High Command (OKM) showing the front line location and the position of German divisions on September 9, 1943.
Source: German National Archives (*Bundes Archiv*)

Figure B.3: DISTRIBUTION OF ALLIED BOMBING DURING THE ITALIAN CAMPAIGN



Notes. Dark blue municipalities experienced at least one Allied strike attack between between September 8, 1943 and May 2, 1945. *Source*: USAF Theater History of Operations Reports (THOR) Database, available at www.afri.au.af.mil/thor.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond, "Preferences for Truth-Telling," *Econometrica*, 2019, *87* (4), 1115–53.

Abrahms, Max and Philip BK Potter, "Explaining Terrorism: Leadership Deficits and Militant Group Tactics," *International Organization*, 2015, *69* (2), 311–42.

Acemoglu, Daron, Leopoldo Fergusson, James Robinson, Dario Romero, and Juan F Vargas, "The Perils of High-Powered Incentives: Evidence from Colombia's False Positives," *American Economic Journal: Economic Policy*, 2020, *12* (3), 1–43.

Ager, Philipp, Leonardo Bursztyn, Lukas Leucht, and Hans-Joachim Voth, "Killer Incentives: Rivalry, Performance and Risk-Taking among German Fighter Pilots, 1939-45," *The Review of Economic Studies*, 2021.

Ahmadov, Anar and James Hughes, *Ideology and Civilian Victimization in Civil War*, The London School of Economics and Political Science, 2017.

Aleksievič, Svetlana, *La Guerra Non Ha un Volto di Donna. L'Epopea delle Donne Sovietiche nella Seconda Guerra Mondiale*, Milano: Bompiani, 2018.

Azam, Jean-Paul, "On Thugs and Heroes: Why Warlords Victimize their Own Civilians," *Economics of Governance*, 2006, *7* (1), 53–73.

Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang, "How Much Should We Trust Staggered Difference-in-differences Estimates?," *Journal of Financial Economics*, 2022, *144* (2), 370–95.

Balcells, Laia, "Rivalry and Revenge: Violence against Civilians in Conventional Civil Wars," *International Studies Quarterly*, 2010, *54* (2), 291–313.

Baldissara, Luca, *Atlante Storico della Resistenza Italiana*, Milano: Bruno Mondadori, 2000.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *The Journal of Political Economy*, 1968, *76* (2), 169–217.

Becker, Gary S and William M Landes, *Essays in the Economics of Crime and Punishment*, New York: National Bureau of Economic Research and Columbia University Press, 1974.

Becker, Sascha O, Sharun Mukand, and Ivan Yotzov, "Persecution, Pogroms and Genocide: A Conceptual Framework and New Evidence," *Explorations in Economic History*, 2022, *86* (101471), 1–18.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig, "This Mine is Mine! How Minerals Fuel Conflicts in Africa," *The American Economic Review*, 2017, *107* (6), 1564–610.

Bianchi, Nicola and Michela Giorcelli, "Reconstruction Aid, Public Infrastructure, and Economic Development: The Case of the Marshall Plan in Italy," *Journal of Economic History*, 2023, *83* (2), 501–37.

Biddle, Stephen D., *Military Power: Explaining Victory and Defeat in Modern Battle*, Princeton, NJ: Princeton University Press, 2004.

Bishop, Chris, *The Military Atlas of World War II*, London: Igloo, 2005.

Blattman, Christopher and Edward Miguel, "Civil War," *Journal of Economic Literature*, 2010, *48* (1), 3–57.

Blumenson, Martin, *Salerno to Cassino*, Washington, DC: Center of Military History, United States Army, 1993.

Browning, Christopher, *Ordinary Men : Reserve Police Battalion 101 and the Final Solution in Poland*, New York: Harper Perennial, 1993.

Cannella, Mario, Alexey Makarin, and Ricardo Piqué, "The Political Legacy of Nazi Annexation," *Working Paper*, 2024.

Carrozzi, Felipe, Edward Pinchbeck, and Luca Repetto, "Scars of War: The Local Legacy of WWI Deaths on British Soldiers," *Working Paper*, 2024.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-Way Fixed Effects Estimations with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (5), 2964–96.

_ **and** _ , "Difference-in-Differences Estimators of Intertemporal Treatment Effects," *Review of Economics and Statistics*, 2024, pp. 1–45.

Clausewitz, Carl, *On War*, Penguin UK, 2003.

Collier, Paul, V.L Elliott, Havard Hegre, Anke Hoeffler, Marta Reynal-Querol, and Nicholas Sambanis, "Breaking the Conflict Trap: Civil War and Development Policy," *World Bank Policy Research Report*, 2003, *56793*.

Collotti, Enzo, "Documenti sull'Attività del Sicherheitsdienst nell'Italia Occupata," in "Il Movimento di liberazione in Italia" 1963, p. 77.

_ , *L'Amministrazione Tedesca dell'Italia Occupata, 1943-1945*, Milano: Lerici Editori, 1963.

Costa, Dora, "Leaders: Privilege, Sacrifice, Opportunity and Personnel Economics in the American Civil War," *The Journal of Law, Economics, and Organization*, 2014, *30* (3), 437–62.

_ **and Matthew Kahn**, "Cowards and Heroes: Group Loyalty in the American Civil War," *Quarterly Journal of Economics*, 2003, *118* (2).

Costalli, Stefano, Francesco Niccolò Moro, and Andrea Ruggeri, "The Logic of Vulnerability and Civilian Victimization: Shifting Front Lines in Italy (1943 1945)," *World Politics*, 2020, *72* (4), 679–718.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving," *The Quarterly Journal of Economics*, 2012, *127* (1), 1–56.

_ **, John List, Ulrike Malmendier, and Gautam Rao**, "Voting to Tell Others," *Review of Economic Studies*, 2017, *84* (1), 143–81.

Dippel, Christian and Andreas Ferrara, "What Makes Effective Leaders? Evidence from Civil War Captains and their Soldiers," *Working Paper*, 2025.

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and conflict: An empirical study," *American Economic Review*, 2012, *102* (4), 1310–1342.

_ **, Massimo Morelli, and Dominic Rohner**, "Strategic Mass Killings," *The Journal of Political Economy*, 2015, *123* (5), 1087–132.

Fehr, Ernst and Simon Gächter, "Cooperation and punishment in public goods experiments," *American Economic Review*, 2000, *90* (4), 980–994.

_ **and** _ , "Altruistic punishment in humans," *Nature*, 2002, *415* (6868), 137–140.

Fisher, Ernest, *Cassino to the Alps*, Washington DC: Center of Military History, United States Army, 1993.

Fontana, Nicola, Tommaso Nannicini, and Guido Tabellini, "Historical Roots of Political Extremism: The Effects of Nazi Occupation of Italy," *Journal of Comparative Economics*, 2022, *51* (3), 723–43.

Gächter, Simon and Jonathan F Schulz, "Intrinsic honesty and the prevalence of rule violations across societies," *Nature*, 2016, *531* (7595), 496–499.

Garland, Albert, Howard McGaw-Smyth, and Martin Blumenson, *Sicily and the Surrender of Italy*, Washington DC: Center of Military History, United States Army, 1993.

Gentile, Carlo, *Wehrmacht und Waffen-SS im Partisanenkrieg Italien 1943-1945*, Schoeningh Ferdinand GmbH, 2012.

_ , *I Crimini di Guerra Tedeschi in Italia*, Torino: Einaudi, 2015.

Glaeser, Edward L., Bruce Sacerdote, and José A. Scheinkman, "Crime and Social Interactions," *The Quarterly Journal of Economics*, 1996, *111* (2), 507–48.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–77.

Gressel, Gustav, *Russia's Quiet Military Revolution, and What It Means for Europe*, Policy Brief, 2015.

Grosfeld, Irena, Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya, "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire," *The Review of Economic Studies*, 2020, *87* (1), 289–342.

Guarnieri, Eleonora, "Cultural Distance and Ethnic Civil Conflict," 2023.

_ , "Cultural Distance and Ethnic Civil Conflict," *American Economic Review*, 2025, *forthcoming*.

_ **and Ana Tur-Prats**, "Cultural Distance and Conflict-Related Sexual Violence," *Quarterly Journal of Economics*, 2023, *138* (3), 1817–61.

Heldring, Leander, "The Origins of Violence in Rwanda," *The Review of Economic Studies*, 2021, *88* (2), 730–63.

Hobbes, Thomas, *Leviathan or The Matter, Forme and Power of a Commonwealth Ecclesiasticall and Civil*, London: Andrew Crook, 1651.

Horowitz, Shale and Min Ye, "Targeting Civilians in Ethno-Territorial Wars: Power- and Preference-Based Sources of Ethnic Cleansing and Mass Killing Strategies," *Studies in conflict and terrorism*, 2013, *36* (5), 372–93.

Humphreys, Macartan and Jeremy Weinstein, "Handling and Manhandling Civilians in Civil War," *American Political Science Review*, 2006, *100* (3).

ISTAT, *Morti e Dispersi per Cause Belliche negli Anni 1940-1945*, Istituto Centrale di Statistica, 1957.

Iyengar, Radha and Jonathan Monten, "Is There an Emboldenment Effect? Evidence from the Insurgency in Iraq," *Journal of Strategic Studies*, 2022, *45* (4).

Kalyvas, Stathis N., *The Logic of Violence in Civil War*, Cambridge: Cambridge University Press, 2006.

Kathman, Jacob D and Reed M Wood, "Managing Threat, Cost, and Incentive to Kill: The Short and Long-term Effects of Intervention in Mass Killings," *Journal of Conflict Resolution*, 2011, *55* (5), 735–60.

Kesselring, Albert, *Soldato Fino all'Ultimo Giorno*, Gorizia: LEG Edizioni, 1954.

King, Gary and Langche Zeng, "Logistic Regression in Rare Events Data," *Political Analysis*, 2001, (9).

Klinkhammer, Lutz, *Stragi Naziste in Italia*, Donzelli, 2006.

_ , *L'Occupazione Tedesca in Italia, 1943-1945*, Torino: Bollati Boringhieri, 2016.

Kurowski, Franz, *Battleground Italy, 1943-1945. The German Armed Forces in the Battle for the "Boot"*, Winnipeg: J.J. Fedorowicz Publishing, Inc., 2003.

Leiby, Michele L., "Wartime Sexual Violence in Guatemala and Peru," *International Studies Quarterly*, 2009, *53* (2), 445–68.

Lingen, Kerstin Von, *Kesselring's Last Battle: War Crimes Trials and Cold War Politics, 1945-1960*, United States of America: University Press of Kansas, 2009.

Lt Col Robertson, J.A., R. Burr, and B. Barth, "USAF THOR Database," 2013.

Midlarsky, Manus, *The Killing Trap: Genocide in the Twentieth Century*, Cambridge: Cambridge University Press, 2005.

Milgram, Stanley, *Obedience to Authority: An Experimental View*, London: Tavistock, 1974.

Mitchell, Neil J, *Agents of Atrocity: Leaders, Followers, and the Violation of Human Rights in Civil War*, New York ; Basingstoke: Palgrave Macmillan, 2004.

Nafziger, George, *The German order of battle : infantry in World War II*, London: Greenhill, 2000.

Olson, Mancur, *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*, New York: Basic Books, 2000.

Oppenheim, Ben and Michael Weintraub, "Doctrine and Violence: The Impact of Combatant Training on Civilian Killings," *Terrorism and Political Violence*, 2017, *29* (6), 1126–48.

Perez-Truglia, Ricardo and Guillermo Cruces, "Partisan Interactions: Evidence from a Field Experiment in the United States," *Journal of Political Economy*, 2017, *125* (4), 1208–43.

Ronchetti, Gabriele, *La Linea Gotica. I Luoghi dell'Ultimo Fronte di Guerra in Italia*, Fidenza: Mattioli 1885, 2018.

_ **and Maria Angela Ferrara**, *La Linea Gustav. I Luoghi delle Battaglie da Ortona a Cassino*, Fidenza: Mattioli 1885, 2014.

Rosen, Sherwin, "The Military as an Internal Labor Market: Some Allocation, Productivity, and Incentive Problems," *Social Science Quarterly*, 1992, *73*, 227–37.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Rousseau, Jean-Jacques, *Discours sur l'Origine et les Fondements de l'Ine'galite' parmi les Hommes*, Amsterdam: Marc-Michel Rey, 1755.

Short, Neil and Chris Taylor, *German Defences in Italy in World War II*, Osprey Publishing, 2006.

Solzhenitsyn, Aleksandr, *Prussian Nights: A Narrative Poem*, London: Collins Harvill Press, 1977.

Stein, Marcel, "A Flawed Genius: Field Marshal Walter Model. A Critical Biography," 2010.

Valentino, Benjamin, Paul Huth, and Dylan Balch-Lindsay, "'Draining the Sea': Mass Killing and Guerrilla Warfare," *International Organization*, 2004, *58* (2), 375–407.

War Crimes Commission, "The trial of Albert Kesselring," in "Law Reports of Trials of War Criminals. Volume VIII," London: War Crimes Commission, 1949.

Weinstein, Jeremy, *Inside Rebellion: The Politics of Insurgent Violence*, Cambridge: Cambridge University Press, 2007.

Wray, Timothy Allen, *Standing fast: German defensive doctrine on the Russian Front during World War II: prewar to March 1943*, Vol. 5, US Army Command and General Staff College, 1987.