Galletta, Sergio; Giommoni, Tommaso

Working Paper

War Violence Exposure and Tax Compliance

CESifo Working Paper, No. 11230

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Galletta, Sergio; Giommoni, Tommaso (2024) : War Violence Exposure and Tax Compliance, CESifo Working Paper, No. 11230, CESifo GmbH, Munich

This Version is available at:
https://hdl.handle.net/10419/302715

CESifo WORKING PAPERS

11230
2024

July 2024

War Violence Exposure and Tax Compliance

Sergio Galletta, Tommaso Giommoni

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 11230*

War Violence Exposure and Tax Compliance

Abstract

We explore the effects of exposure to conflict and violence on civic compliance with the law. Using newly digitized historical records of income declarations and tax audits from post-World War I Italy, we show that losing a relative as a direct result of the war reduces tax compliance. To account for the potential endogeneity of the treatment, we use an instrumental variable strategy exploiting the exogenous allocation of soldiers to more/less dangerous military units. Our results show that the effect of reduced tax compliance remains consistent across different measures of compliance and is not due to economic reasons. We also find that this negative impact is lessened when the state acknowledges the sacrifice of the deceased, in communities that suffered many casualties, or in areas with high levels of social cooperation before the war. Overall, our findings suggest that war can erode social norms, leading to a lower willingness to contribute to public goods, such as paying taxes.

Sergio Galletta
ETH Zurich / Switzerland
sergio.galletta@gess.ethz.ch

Tommaso Giommoni
ETH Zurich / Switzerland
giommoni@ethz.ch

We wish to thank Daron Acemoglu, Elliott Ash, Sergio Correia, Camilo Garcia Gimeno, Evelina Gavrilova, Leander Heldring, Marko Köthenbürger, Krzysztof Krakowski, Federico Masera, Andrea Moro, Saresh Naidu, Aurélie Ouss, Joachim Voth for the insightful suggestions. We have also benefited from comments by participants at the "Petralia workshop" (Petralia), Workshop on Political Economy (Bruneck), Zurich Conference on Public Finance in Developing Countries, CAE Conference (Toronto) and seminar at the NHH Norwegian School of Economics (Bergen), ETH Zurich, Universitá della Svizzera Italiana (Lugano), ifo Institute (Munich), University of Bergamo, University of Cagliari, ZEW-ifo Young Scholar Political Economy Workshop. This version: July 2024 (First version: January 2023).

1. Introduction

Wars have recurrently shaped human history, causing clear negative impacts on the economy and society, such as destruction and deaths (Collier et al., 2003). However, evidence also suggests that wars can positively affect state development by prompting investments in better institutions and increased fiscal capacity (Besley and Persson, 2009, 2010; Gennaioli and Voth, 2015). The success of these changes relies on citizens' willingness to comply with the law (Besley, 2020; Levi, 1988), which can be influenced by their war experiences. Some studies find war exposure enhances cooperation and societal engagement post-war (Bauer et al., 2016), while others indicate it can erode confidence in the state and reduce compliance (Conzo and Salustri, 2019; De Juan and Pierskalla, 2016; Voors and Bulte, 2014). These effects vary based on historical and conflict contexts (Besley and Persson, 2009; Bowles, 2006; Cassar et al., 2013; Rohner et al., 2013), reflecting potential breaks in the social contract between the state and its citizens (Caprettini and Voth, 2023).

In this paper, we examine whether exposure to war violence affects individuals' willingness to obey the law, focusing on tax compliance. Tax compliance serves as a proxy for civic culture, reflecting prosocial preferences and trust in the state (Putnam, 1995; Luttmer and Singhal, 2014), and is directly linked to fiscal revenues, a key measure of state capacity. The causal effect of war exposure on tax compliance remains an open empirical question we aim to address.

To investigate this question, we provide novel evidence on how exposure to war violence affects individual tax compliance in post-World War I Italy. Using newly digitized historical records, we analyzed income declarations from a large sample of Italian taxpayers in the early 1920s, including details on their compliance with tax laws as determined by audits. We linked tax noncompliance to indirect war violence exposure by identifying whether a taxpayer had a relative who died in the war, distinguishing between battlefield deaths and other causes. This was achieved by matching taxpayers by surname and municipality of origin with the Italian "Honor Roll of World War I Dead" (*Albo dei Caduti Italiani della Grande Guerra*), which lists biographical information on identified military casualties.[1]

[1]This approach overcomes the lack of specific census data on family ties. We discuss in the paper the potential limitations and biases introduced by this method and show that they do not significantly affect the overall results.

Providing causal evidence on the effect of war violence exposure on tax compliance is challenging due to potential unobserved family characteristics that influence both tax compliance and the probability of war casualties. To address endogeneity, we use multiple empirical strategies. First, we estimate a model where treatment is defined by whether a taxpayer had a family member who died during the war, including several fixed effects to capture unobservable characteristics. Next, we narrow our focus to taxpayers who experienced a loss, distinguishing between battlefield and non-battlefield deaths. This helps us exclude biases from selection into the war by comparing families who all experienced a relative's death and focusing on a treatment that is directly linked to the violent and traumatic aspects of war. Finally, we employ an instrumental variable (IV) strategy, where the instrument measures the average probability of having a relative dying on the battlefield compared to other causes, which is mostly determined by the exogenous allocation of soldiers to riskier military units.[2]

Our first set of results suggests that tax noncompliance is 0.4 percentage points (p.p.), or 3.4% of the average tax noncompliance, higher among families that experience a loss during the war, and this effect is entirely driven by death on the battlefield. The results are robust to the inclusion of a set of relevant covariates such as income, municipality fixed effects, surname fixed effects, and surname $\times$ province fixed effects. Importantly, the evidence is also confirmed if we constrain the analysis to the sample of taxpayers who had a relative who died during the war, thereby using a treatment status defined by the reason for death. Families that experienced the loss of a relative on the battlefield are 1.1 p.p. (8.1% of the average tax noncompliance) less likely to comply with the tax law compared to families where the loss was due to other reasons. When using the IV strategy, the effect is 6.6 p.p (48.5% of the average tax noncompliance).

To validate the IV strategy, we show that the instrument strongly predicts a soldier's cause of death, but it is not correlated with his family's area of origin, income, education, or profession and that there is no systematic sorting of soldiers from a given family into high/low-risk military units. Indeed, the crucial identifying assumption is that military unit assignment is not affected by the pre-existing chance of one's family obeying the tax law, and the exclusion restriction assumption only applies if the riskiness of a soldier's

[2]With a similar intuition Jha and Wilkinson (2012) use the arbitrary assignment of Indian battalions to different fronts and to different periods of time at the front during the World War II to study the link between combat experience and organizational skills driving political action.

military unit affects their family members' tax compliance through the probability of the soldier's death. We add some evidence on this showing that the instrument is orthogonal to the probability of tax noncompliance as predicted by a regularized logistic regression trained on the surname-municipality of origin of the taxpayer.

We complement our main evidence with additional analysis. First, we address the possibility that our results are driven by auditors' behavior rather than increased taxpayer wrongdoing. Political authorities might have directed auditors to scrutinize taxpayers believed to oppose the regime, potentially correlating with our treatment. To test this, we use an alternative measure of tax noncompliance: declaring an income just below the minimum taxable level. This measure is also statistically significantly affected by war violence exposure in a similar direction as in the main analysis. Second, we test for extrinsic motives by showing no statistical difference in reported income between taxpayers whose family members died on the battlefield and those whose family members died from other causes. Third, we find the effect is mitigated under specific conditions. When a relative who died on the battlefield was a military official or is honored with a medal, the effect is null or reversed. The effect is also lower for families in areas more exposed to the war or in municipalities with strong pre-war preferences for redistribution, as indicated by the 1913 socialist vote share. This suggests that state recognition and pre-existing prosocial attitudes can mitigate the negative impact of war loss on tax compliance.

Additionally, we document that tax compliance declines with the number of family losses and that the effect remains similar when focusing only on foot soldiers, who are less subject to selection bias. Our main results are robust when controlling for taxpayers' professions, excluding common surnames, and excluding large cities.

The results of our study are consistent with the idea that exposure to war violence can decrease trust in the state and social cohesion, leading to a weakening of social norms and institutions, such as the rule of law. The fact that tax compliance is influenced by the cause of death suggests that intrinsic motivation plays a crucial role, rather than economic conditions or tax enforcement levels. This aligns with our findings on the heterogeneous effects, where recognition of the deceased's sacrifice and strong pre-existing prosocial attitudes mitigate the negative impact. Lower tax compliance negatively impacts fiscal revenues, a crucial determinant of state capacity, reducing the state's ability to function well.

Our analysis contributes to the broader literature on the consequences of war, focus-

ing on both economic and non-economic outcomes. On the economic side, studies have shown that wars can impact state and fiscal capacity, directly linked to tax compliance. Besley and Persson (2009) and Besley and Persson (2010) provide the theoretical basis, suggesting that external conflict is associated with higher legal and fiscal capacity. Gennaioli and Voth (2015) argue that European states formed and invested in state capacity as public revenues became crucial for military success. Besley (2020) expands on this, exploring how a strong civic culture, which includes tax compliance and public goods provision, leads to a stronger economy. Similarly, Becker et al. (2020) find that cities involved in wars developed higher democratic institutions and increased local fiscal capacity.

On the non-economic side, Bauer et al. (2016) provide a meta-analysis showing a positive link between conflict exposure and prosocial behavior, while Gilligan et al. (2014) and Bellows and Miguel (2009) find increased social capital and community engagement following violence. Conversely, Cassar et al. (2013) and De Juan and Pierskalla (2016) report that violence can undermine trust and increase ethnic identity, slowing post-conflict recovery (Rohner et al., 2013). Bühler and Madestam (2022) and Dell and Querubin (2018) also find that state repression and aggressive military tactics can weaken governance and civic engagement.

Studies focusing on world wars generally emphasize a negative effect on trust. Conzo and Salustri (2019) and Grosjean (2014) show that childhood exposure to World War II lowers trust in governmental institutions. Caprettini and Voth (2023) find that the New Deal boosted patriotism during World War II. Cagé et al. (2020) demonstrate how heroic figures influence political behavior, while Acemoglu et al. (2022) link high casualties in World War I to increased support for the Socialist and later the Fascist Party in Italy. Scheve and Stasavage (2010) show that mass mobilization in World War I led to significant and lasting increases in progressive taxation, as governments responded to wartime demands for fairness by taxing the wealthy more heavily.

Our results also contribute to research on tax compliance, which is influenced by both pecuniary (extrinsic) and non-pecuniary (intrinsic) motivations (Luttmer and Singhal, 2014). Most literature focuses on extrinsic motivations. For example, Slemrod (2019) reviews how public policies reduce tax evasion, and Kleven et al. (2011) find that audits lead to higher reported income, indicating deterrence effects. Similar findings are reported by DeBacker et al. (2018) in the U.S. and Advani et al. (2021) in the U.K.

Fewer studies address intrinsic motivations. Dwenger et al. (2016) use a field exper-

iment on a local church tax in Germany to show that, even without deterrence, 20% of individuals paid the owed taxes, highlighting the role of intrinsic motivation. Surveys and experiments further emphasize individual characteristics in compliance behaviors (Cummings et al., 2009; Guerra and Harrington, 2018; Doerrenberg and Peichl, 2013). Cross-country data from the World Value Survey show that positive attitudes toward tax compliance correlate with conflict frequency and duration, although high casualty rates reduce compliance support, aligning with our findings (Feldman and Slemrod, 2009).

Our study offers several advantages over previous research. First, by using real-world historical data on individual tax compliance, we capture actual behavior rather than relying on surveys or experimental settings that measure social preferences and trust. Tax compliance in our context serves as a proxy for the willingness to contribute to society. Second, our approach distinguishes between intrinsic and extrinsic motives for compliance; the treatment effect hinges solely on the reason for a family member's death, not on the family's economic status. Third, our focus on post-World War I Italy provides a unique historical and economic context that has been less explored, as most studies concentrate on recent conflicts in developing countries.

The paper proceeds as follows. Section 2 provides institutional and historical details. Section 3 and Section 4 describe our data and empirical strategy, respectively. Section 5 and 6 present our results. Section 7 concludes.

2. Historical background

2.1. World War I

Italy during the conflict. In August 1914, the five European Great Powers initiated the first World War. Nine months later, Italy entered the war and battled against Austria-Hungary, its pre-war ally. The decision to join the conflict can be attributed to Italy's diplomatic negotiations with both Allied and Central powers about the compensation linked to war participation. The Allies agreed to Italian territorial demands, and on April 1915, the Treaty of London was signed. Italy officially entered the war a month later. Despite this decision, the idea of entering the conflict was not popular among Italian politicians and the population at large, with the exception of a small but very active group that strongly favored involvement in the conflict. Indeed, the majority of parliamentarians, including the Socialists, the Popular Party, and most Liberals,

were opposed to entering the war. During this period, Benito Mussolini, the future leader of the fascist dictatorship that governed Italy for twenty years after the end of the War, shifted his political position moving away from the Socialist Party to join more pro-intervention/nationalist forces.

Italian military activities were mostly concentrated on the Eastern Italian front, where the army attempted to enter the Austria-Hungary territory to reach the city of Vienna. Several battles were fought over three years, from 1915 to 1918, resulting in heavy casualties on both sides. The most important battles for the Italian army during World War One were those that took place on the Isonzo river (in the territory of present-day Slovenia). After an initial gain of the ground by the Italian army, the Austro-Hungarian army won the Battle of Caporetto in October/November 1917 and moved the front line to the Piave river, 100 km inside of the Italian territories. One year later, under the command of General Armando Diaz, the Italians emerged victorious in the battle of Vittorio Veneto, and on November 1918, the armistice was signed. After the end of the war, Italy received only part of the territories initially promised by the Anglo-French alliance.

Conscription methods. From 1915 to 1918, nearly six million men born between 1874 and 1899 underwent mandatory conscription into the military. This significant conscription effort was not a novel strategy but was rooted in procedures that had been used since the pre-war era. Municipal conscription offices compiled detailed lists of young men who became eligible for military service upon turning 20. Furthermore, administrative ease was achieved by treating each province as its own military district, thereby allowing the geographical and administrative boundaries of a province to coincide with its corresponding military district.

Once these provincial conscription lists were assembled, they were handed over to the Conscription Council. This Council was a structured body, presided over by a prefect representing the Minister of War. Its key responsibilities were twofold: overseeing the lottery drawing for the conscription from each provincial military district and rigorously assessing the psychological and physical fitness of the registrants according to their year of birth.

However, not every name on the conscription lists would wear the military uniform. Various factors, such as family obligations or health conditions, might deem some men temporarily or even permanently unfit for service. Those deemed eligible were broadly

categorized into two segments. The first segment comprised those chosen through the lottery system, ensuring that the annual conscription quota set by the Ministry of War was met. The second segment, on the other hand, consisted of those who weren't immediately inducted but were kept on standby for potential future service.

Composition of the Italian army. Upon selection, men from the first segment were allocated to particular military specializations (i.e. infantry, cavalry, artillery, grenadiers, alpine, *bersaglieri* and engineers corps). This process factored in a range of criteria including physical characteristics (e.g., the grenadiers had a height benchmark of 176 cm) and personal experiences like prior horse-riding for cavalry assignments.[3] Regiment composition followed other criteria; men from five or six different districts would be grouped together, ensuring that each regiment had representation from one of the five military zones into which the peninsula was divided.[4] This emphasis on diverse regiment composition can be traced back to strategies post-Italian unification in 1861. The army endeavored to cultivate a strong Italian identity by ensuring recruits from varied regions served together in a regiment. Additionally, periodically rotating these recruits across regions was a tactic employed to dissolve linguistic and regional barriers.

It's worth noting that the World War I years brought about some adjustments to this regimented system. The overwhelming demands of the war meant that Italy had to revisit and adapt its recruitment criteria. Measures such as the relaxation of the minimum height requirement, reconsideration of health disqualifications, and sidelining the lottery drafting system were implemented. Despite these changes and some logistical challenges leaning towards regional recruitment, the overall composition of the regiments remained predominantly national. A blend of both veterans and fresh recruits ensured that the essence of national unity within military units was preserved.

Importantly, the morale of the soldiers was very low during the years of the war. Indeed, most soldiers did not identify with an overarching moral cause behind their wartime mobilization. The frustration of the soldiers, who most of the time came from the countryside, was partly due to their disappointment with the reasons for the conflict and partly due to the extremely harsh conditions of trench warfare (Bianchi, 2014). This sentiment is also reflected in the number of soldiers that tried to avoid involvement in

[3] Alpine troops, for instance, were preferably coming from municipalities in mountainous areas.

[4] Appendix Figure A.1 provide an example of such an allocation method, reporting the distribution of soldiers from each district that were supposed to join the different infantry regiments in 1880.

the war operations compared to those volunteering. While the number of volunteers was very modest (10,000), a large number of violations were recorded (470,000 resisted call-ups, 310,000 committed acts of indiscipline under arms, and 300,000 deserted). By the end of the war, roughly 600,000 Italian soldiers had died. Similar negative feelings about the war resonate with the population at large. Indeed, the major political groups opposing the war, the Socialist and the Popular party, received historically high support in the first national elections after the war.

2.2. The Italian tax system

The Italian income tax in 1920s. Since the Italian unification in 1861, the national government has worked to harmonize the tax system across the different areas that used operate under diverse tax regulations. As was common in other countries during that period, Italian tax revenue relied heavily on direct taxation (e.g., land taxes, property taxes, and labor and capital-income tax). For example, in 1921, direct taxes represented 37% of the total state revenue.[5]

The most relevant direct tax was the income tax called "*imposta di ricchezza mobile*", which was in effect between 1864 and 1973.[6] This tax was levied on income generated using capital or labor force (i.e., it excluded income from land use).[7] Taxpayers belonged to different categories depending on their source of income: derived from capital (category A), commercial and industrial activities (category B), liberal arts (category C), or from salaries and pensions of public employees (category D). Importantly for our analysis, a few years after the war ended, the Italian tax authority decided to publish several volumes reporting information about the taxable income of taxpayers from categories B and C, i.e. those individuals that had to file the tax return (more details in Section 3). Taxation was progressive, and each category of taxpayer was subject to a different tax schedule. For example, from 1918 to 1924, there were four tax brackets for category B ranging between 10% (income lower than 1,500 lire) to 16% (income higher than 5,000

[5]Specifically, 4,749,517,000 over a total of 12,770,958,000 lire. The second largest category was profit from monopolies such as tobacco and salt (3,003,421,000 lire), while indirect taxes represented the third largest category (2,186,950,000 lire) (De Stefani, 1923).

[6]The "*imposta di ricchezza mobile*" was initially introduced with the Law of July 14, 1864, No. 1830. This tax has been largely understudied. Just recently a few contributions have started exploring the vast source of historical data to economically address relevant questions (Acconcia et al., 2022; Gabbuti et al., 2022; Galletta and Giommoni, 2022).

[7]What we describe in this section is mostly based on Clementini et al. (1934).

lire), and there was a minimum taxable income of 533.4 lire, while for category C there were also four tax brackets, but they ranged between 9% (income lower than 1,667 lire) to 15% (income higher than 5,000 lire), and there was a minimum taxable income of 640 lire. There were two different methods for collecting taxes: public employees and salaried workers had a withholding tax deducted at source, "*ritenuta diretta*", while self-employed and personal businesses were required to file individual tax returns, "*ruoli nominativi*". For the purpose of the paper, it is worth describing the system that guided the latter method of tax collection.

Tax collection. The basic idea was that each municipality was required to keep an annually updated list of people living in their territory that were likely to have taxable income as requested by the law. With the support of the tax agents, local authorities, such as the municipal council, "*giunta comunale*", were in charge of identifying eligible citizens and drafting a list of potential taxpayers, among those residents in the municipality. The tax agents could also update this list if they had information helpful in identifying taxpayers that needed to be added or removed from it.

The taxpayers on the list received a letter requiring them to complete a tax return for their previous year's income and to send it back to the tax agent, typically between the 1st of May and the 31st of July. Importantly, tax agents had the task of verifying the correctness of the declaration, by having the right to access documentation about the taxpayer in the hands of the public administration, and directly from individual business accounts. For those taxpayers that did not send back the declaration but were already on the list, the information based on the income declared the previous year was imputed, unless there were suspicions that the income did change. In such cases, the agent rectified the information by correcting the declared income.[8] Similarly, if a taxpayer sent a declaration that the authority found to be incorrect, it was adjusted accordingly by the tax agent. In cases where a taxpayer deliberately omitted an item of their income or underreported it, they incurred a penalty amounting to twice the tax owed on the discrepancy between the reported income and the actual income. Every time the agent updated the income on the list, she notified the taxpayer of the change. If the taxpayer disagreed with the decision, she could appeal to a tax commission that would rule in favor or against such decision. Finally, once the list was completed, it was

[8]However, such a decision could be modified for four years, unless the taxpayers updated with a new tax declaration.

sent to the tax collector, who was responsible for the collection.

3. Data

3.1. Income declaration data

The income data we use in this paper are included in the tabulations published by the Italian tax authority in 1924, which refer to the previous fiscal years. To collect this information, we digitized the historical publications for 24 Italian provinces (see Appendix Figure A.2, panel a), with a coverage of around 2,000 municipalities.[9] Figure A.3 provides an example, showing one page of the publication for the province of Bologna.[10] These reports contain information about the taxable income of taxpayers declaring revenues from commercial and industrial activities (category B) or the liberal arts (category C). Most of these taxpayers are owners of small businesses (e.g., individual companies or small stores) and craftsmen. Table A.1 displays the number of submitted tax declarations and the total and average amount of income for each group. The categories under analysis (B and C) represent the majority of the taxpayers subject to this income tax (69.2%). Similarly, if we focus on declared income, groups B and C account for nearly 7 billion lire, more than two-thirds of all declared income in that year. The final dataset includes information on 215,180 out of 955,198 taxpayers (categories B and C), representing nearly 23% of the national sample in these two categories.

Aside from declared income, the tabulations include additional information on individuals' tax declarations. Importantly, it is reported whether the declared income has been corrected by the tax auditor. We use this information to construct a binary variable capturing whether someone's income is audit-adjusted, which is our primary measure of tax noncompliance at the individual level. Second, data on taxpayers' occupations are included, which we harmonized and aggregated, whenever possible, in larger occupation categories. Table A.2 shows the main descriptive statistics, panel a focuses on the entire sample of taxpayers, while panel b only refers to those taxpayers with at least one rela-

[9] Importantly, the selection of the provinces in the dataset is not due to any particular reason. These publications are those that we have been able to collect and digitize via online sources and historical archives. It is useful to note for the scope of the analysis that these provinces are scattered across Italian regions.

[10] These income declaration data are also used by Galletta and Giommoni (2022) to measure municipal inequality in Italy.

tive who died during the war (sample of reference for the IV analysis). In the full sample (panel a), the average (gross) declared income is 3,445 lire, and 11.9% of taxpayers report an audit-adjusted income. Moreover, 3.1% of taxpayers declares an income just below the minimum taxable income (533.4 lire for category B and 640 lire for category C). Finally, Table A.3 displays the information on individuals' professions. A minority of taxpayers (5.7%) holds a university degree and practices liberal arts (5.9%), while the most common profession (15.7%) is owner of a retail business.

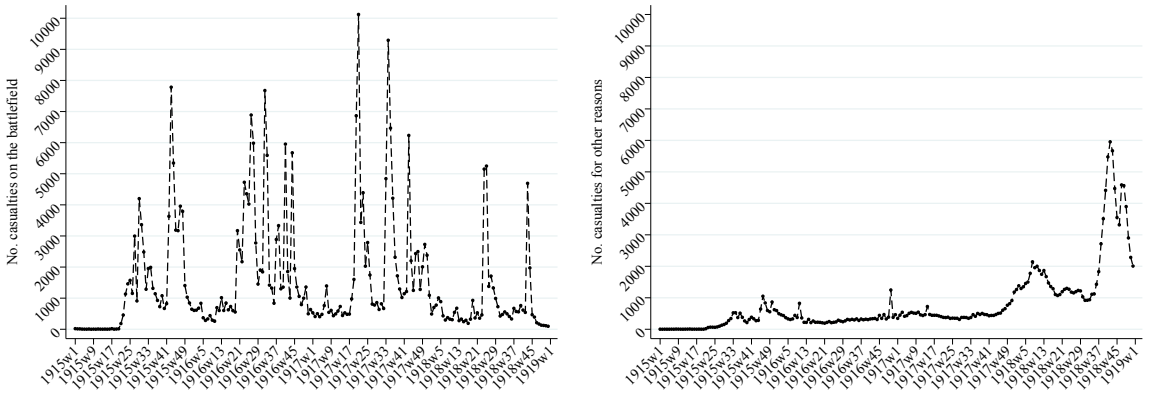
3.2. World War I casualties

The explanatory variable of the analysis captures the extent to which a family has been impacted by the war. To construct this measure, we collected information on the universe of Italian military victims in WWI. These data are contained in the *Albo dei Caduti Italiani della Grande Guerra*, a publication containing information on all the Italian military victims that had been identified from 1915 to 1920. The number of Italian military victims in WWI has been estimated at around 651,000 (Mortara, 1925), and this publication contains information on 540,401 soldiers, suggesting an excellent overall coverage.

Figure 1: Evolution of casualties over time

(a) In the battlefield

(b) For other reasons



Notes: The figures show the number of casualties over time, as reported in the *Albo dei Caduti Italiani della Grande Guerra*, with weekly aggregation. Panel a shows the number of soldiers that died on the battlefield, Panel b shows the number of soldiers that died for reasons different from fighting.

In particular, the *Albo dei Caduti* contains information on soldiers that a) died in combat or due to war injuries, b) went missing in action, c) died or went missing in captivity (except deserters), d) died of illness related to the service, e) died of an accidental

cause, and f) died by a suicide whose cause was related to war service (Zugaro, 1926). The information on military victims contained in the publication is very detailed, including names, dates of birth, cities of origin, military rank, and information on whether the soldiers were awarded a Medal of Honor (and the type of Medal). Also, information on the circumstances of death, such as the date, the place, and the cause of death, is available. This allows us to distinguish between soldiers who died on the battlefield, while fighting, from those who perished due to other reasons less directly linked with war operations, such as illness or an accident. Figure 1 shows the number of casualties on the battlefield (panel a) and for other reasons (panel b) over time. While the number of military victims in combat varies greatly over time, with spikes in correspondence with large battles (such as the battle of Caporetto in October/November 1917), the trend in victims for other reasons is quite flat, with the only exception being the two spikes in 1918, due to the Spanish flu.[11]

Our goal is to measure how families have been affected by WWI. With this aim, we aggregate these data at the *family* level: given the lack of specific data on family ties, we define a family as a surname-municipality cluster.[12] Therefore, all soldiers with the same surname from the same municipality are considered part of the same family. This approach is reasonable because Italy in the 1920s was characterized by a large number of highly locally concentrated surnames (Caffarelli and Marcato, 2008) and a high number of local geographical units as well as low geographical mobility.[13] We corroborate this evidence with our data. Our findings show that, from the list of deceased soldiers, over 53% of the surnames are found in a single municipality, and more than 85% of the surnames are concentrated in up to five municipalities.[14] Following this approach, we created a set of variables capturing the impact of WWI on Italian families. In particular, we created several variables at the family level (surname-municipality) indicating whether one or more family members 1) died during the war for any reason, 2) died on the battlefield, or 3) died for other reasons, such as an accidental cause or illness.

[11] Moreover, Figure A.2 (panel b) shows the share of military victims over the male population in 1911 for the set of provinces in the analysis.

[12] This is a common method used in the literature when specific information on family relationships is unavailable (e.g., Gagliarducci and Manacorda, 2020)

[13] In particular, in 1921, Italy was divided into 9,208 municipalities, the majority of which were small towns. The median municipal population was 2,123, and the median municipal area was 2.1 km^2.

[14] We report the results graphically in Appendix Figure A.5.

We then matched this family-specific data with the individual dataset of tax declarations using the information on surname and municipality of residence. This procedure matched each taxpayer with the information on whether his family has been affected by WWI, as previously defined. This method is naturally correlated with some measurement error; we discuss in section 4.3 the potential impact of this misclassification. Table A.2 reports some descriptive statistics: if we focus on the full sample (panel a), 31% of taxpayers belong to a family that suffered a loss due to the war, and the average number of victims is 0.79. Moreover, the share of taxpayers with a relative who died in combat is larger (25.2%) than those with a relative who died for other reasons (5.7%).

Finally, we rely on the *Albo dei Caduti* to extract information on the military unit membership of the victims, which identifies the regiment within a specialization in which a certain soldier was included. After a thorough cleaning in which we corrected for likely wrong or miss-specified units, we managed to assigned a unit membership to the 84.5% of the victims recorder in the *Albo dei Caduti*. Table A.4 details the structure of the Italian army across the different military specializations and units. This information is relevant for our analysis as we want to measure the level of risk that characterized each military unit during the war, and how this affects the probability that a military unit member died in combat. With this goal, we define the risk score of a military unit as the ratio between the number of casualties on the battlefield and the total number of casualties.[15] Figure A.4 shows the distribution of the risk score by soldier (panel a) and by family –after the matching– (panel b). Both figures display a distribution with a left tail and a larger mass around the value of 0.7. This is confirmed by Table A.2 (panel b), which shows that the average risk score in the taxpayers' dataset is 0.67.

4. Empirical Strategy

Our aim is to investigate the effect of war violence exposure on tax compliance. Simply regressing tax compliance on violence exposure via OLS would provide a biased estimate. Indeed, the empirical challenge comes from the presence of unobserved characteristics of a given family that could be associated with both their members' will-

[15]It is important to mention that for the construction of this indicator, we excluded from the sample the military units with very few casualties, whose classification is likely to be inaccurate. Moreover, the main results also hold if we only focus our analysis on the largest units (top 10%) that include the majority of casualties in the analysis.

ingness to comply with the tax law and their probability of dying during the war. Biased estimates can be determined by two clear sources of selection: first, a selection into participating in WWI; second, a selection into dying during the war. While participation in the army was based on conscription, there was still some level of noncompliance with the military draft. Therefore, the first source of selection will bias the estimates if, for example, members of families who tend to have little trust in the government will be more likely to evade the military draft and not pay taxes. About the second source of selection, it would be problematic if, for instance, members of a very nationalist family are generally compliant with the tax law and have behavior on the battlefield that makes them more at risk of dying. To reduce the endogeneity concerns, we first apply a fixed-effects model meant to capture relevant unobservable characteristics, and next, we use an instrumental variable strategy that exploits exogenous variation in the relative risk of dying on the battlefield instead of dying for other reasons. For the second empirical application, we use only the sample of taxpayers with at least one relative who died during the war while enrolled in the army.

4.1. Fixed effects

We start by estimating the following equation via OLS:

$$\text{Tax Noncompliance}_{ismp} = \beta \text{Death relative in WWI}_{smp} + \gamma_s + \lambda_m + \theta_{sp} + \eta \mathbf{X}_{idp} + \epsilon_{ismp} \quad (1)$$

where *Tax Noncompliance*$_{ismp}$ is a variable identifying whether an individual i with surname s from a municipality m in province p had an audit-adjusted tax return. The treatment is *Death of a relative in WWI*$_{smp}$ which is a dummy that identifies if there is at least one soldier that died during the war for each surname-municipality. β is the coefficient of interest, which identifies the relationship between exposure to war violence and tax compliance. To account for relevant unobservable characteristics at the geographical and family level, we introduce a set of fixed effects, γ_s, λ_m and θ_{sp}, and controls, $\mathbf{X}_{idp}$. ϵ_{ismp} is the error term. Thanks to λ_m we control for characteristics shared among individuals that live in a given municipality. Therefore, we implicitly control for predetermined political preferences and attitudes towards the government and the war at the municipality level. Moreover, we control for the geographical variation in tax authority conduct, as all individuals in a given municipality are under the control of the same local office. The inclusion of γ_s and θ_{sp}, helps us reduce concerns linked with the selection

due to unobservable at the family/surname level. In particular, θ_{sp} allows a comparison of the level of tax compliance between individuals that share the same surname in a given province. In other words, the effect is estimated by comparing someone with a given surname and municipality of origin who has a family member who died during the war with other individuals with the same surname coming from the same province but different municipalities that did not have a family member that died. This comparison would largely remove potential endogeneity issues if individuals with the same surname from a certain province belonged to the same family, and if living in a municipality or another, within a specific province, is somehow exogenous to the relationship we are studying. Finally, the controls set, $\mathbf{X}_{idp}$, includes the number of taxpayers from the same surname-municipality and the tertile of declared income.

To improve this initial specification, we introduce a narrower definition of treatment that uses information on the actual cause of death. Specifically, we use in place of the main treatment two variables: 1) *Death of a relative on the battlefield*$_{smp}$ identifies for each surname-municipality if there is at least one soldier that died for reasons linked to war operations, such as on the battlefield or as a consequence of actions on the battlefield, while 2) *Death of a relative for other reasons*$_{smp}$ is a dummy that identifies whether a certain family had (only) relative(s) in the army that died for reasons not clearly linked with military operations, for instance, illness or accidents.

This alternative treatment definition helps us on several dimensions. First, it allows for a better understanding of the forces driving the main effect. If intrinsic motives are behind the relationship we are studying, one should expect the main effect to be mostly due to cases in which the relative's death is directly linked to the war, such as death on the battlefield, as these are the only kinds of deaths that can plausibly be directly linked to the violent aspects of war. Moreover, it allows us to analyze a restricted sample of taxpayers: those with a relative who went to war and died, which completely rules out the selection bias coming from having a family member who participated in the war. This restriction also removes concerns that the effect is mainly driven by factors linked with the death of a relative in general, as in this case both the treated and the control group experienced the loss of a relative. Lastly, it introduces the basic variable definitions used in the IV strategy described in the next section.

4.2. Instrumental variables strategy

While fixed effects account for relevant unobservable features, they do not directly address potential selection issues at the municipality-family level. Therefore, we complement the previous identification strategy by applying an instrumental variables strategy that exploits the exogenous allocation of soldiers to more/less risky military units. Though we do not have detailed information on all soldiers that participated in the war, we can still compute a relative measure of riskiness for each military unit by using the probability of dying on the battlefield compared to other reasons from the universe of victims during WWI. Because the risk score is defined only for soldiers that died during the war, we constrain the analysis to the sample of taxpayers that had (at least) a relative that died during the war.[16] Panel B of Table A.2 reports the descriptive statistics for this sample of taxpayers. Importantly, given that we rely on the assumption of plausibly exogenous allocation of soldiers across different units, we always exclude from the sample all military specializations that applied a selective recruitment of soldiers (as discussed in Section 2). Therefore , we only focus on taxpayers whose relative(s) served in artillery, infantry or in one of the corps of *Bersaglieri*. Nonetheless, it is worth mentioning that these three specializations represent the great majority of the military victims, the 87.7% (as displayed in Table A.4).

We start by computing for each soldier a score for the risk of dying on the battlefield of his military unit, compared to the overall risk of dying in the military unit. Importantly, we do not include members of the military unit that come from the same municipality and have the same surname as the soldier, in order to account for potential concerns related to selection due to municipality or family of origin. Next, we create the instrument by averaging the risk score at the surname-municipality level. Our approach is similar to a standard leave-one-out instrument but differs from it in two aspects: 1) we leave out not only the soldier of interest but also individuals from the same municipality or with the same surname; 2) the instrumental variables strategy is not used to study soldiers' outcomes but to measure the compliance behaviors of their family members that did not interact directly with the military unit. These aspects should account for concerns about the validity of our strategy, in contrast with the standard leave-one-out instrument (Angrist, 2014). In particular, in our setting, the exclusion restriction

[16] Ideally, having access to the full record of Italian soldiers that participated in WWI, we would link taxpayers also with soldiers that went to the war and survived.

assumption depends on the fact that the riskiness of a soldier's military unit affects his family members' tax compliance only through its effect on the probability of the soldier's death, which is a plausible assumption.

We estimate the following first-stage equation:

$$\text{Death rel. battlefield}_{smp} = \alpha \text{Avg. military units risk}_{smp} + \\ + \gamma_s + \lambda_m + \theta_{sp} + \eta \mathbf{X}_{idp} + \epsilon_{ismp} \quad (2)$$

where, in addition to all elements previously defined, we also have *Avg. military units risk*$_{smp}$, which is the average risk score associated with the military units to which the family member(s) have been assigned. Moreover, in this second set of analyses, which only focuses on families that have at least one loss due to the war, we also control for a battery of characteristics of the military victims, such as a dummy variable indicating that at least one military victim (within the family) was an army official, a set of variables indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), and a dummy capturing whether the military victim was young (below the median age). The crucial assumption is that the assignment to a given military unit is exogenous with respect to the family's likelihood of compliance with the tax law, conditional on controls. We support this assumption with historical evidence and by providing a set of empirical checks, as discussed in section 5.2. Following the standard 2SLS approach, the tax compliance measure will be regressed on the fitted value from the first stage.

4.3. Misclassification

An important aspect to mention is that the analysis could suffer from misclassification issues due to the matching procedure between the two main datasets. In fact, there could be cases in which a taxpayer is mistakenly identified as having a family member who died during the war (either on the battlefield or for other reasons), although that was not the case and other instances where the opposite occurs. From the literature, we know that this kind of misclassification can produce nonclassical measurement error, which in the OLS are typically the source of attenuation bias (Aigner, 1973). However, the opposite occurs in the IV, thereby producing an upward bias (Bingley and Martinello, 2017; Meyer and Mittag, 2017; DiTraglia and Garcia-Jimeno, 2019). Therefore, it is plausible to think that results may provide bounds of the actual effect, and we provide

some estimates about that.[17]

5. Main results

5.1. Fixed effects

Table 1: Effect of War on Tax noncompliance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Death of a relative	0.013**	0.011*	0.006**	0.004*				
	(0.006)	(0.007)	(0.003)	(0.003)				
Death of a relative in the battlefield					0.018***	0.016**	0.008**	0.005**
					(0.007)	(0.007)	(0.004)	(0.003)
Death of a relative for other reasons					-0.005	-0.006	0.002	0.001
					(0.005)	(0.006)	(0.004)	(0.004)
N observations	214,931	214,931	186,903	169,462	214,931	214,931	186,903	169,462
Adjusted R^2	0.000	0.015	0.172	0.265	0.001	0.015	0.172	0.276
Mean dep. var.	.120	.120	.122	.125	.120	.120	.122	.125
Baseline controls		✓	✓	✓		✓	✓	✓
Surname FE			✓				✓	
Province FE			✓				✓	
Municipality FE				✓				✓
Surname FE × Province FE				✓				✓

Notes: The sample includes all taxpayers in the main dataset. The dependent variable is *tax noncompliance*, a binary variable indicating whether the taxpayer's income has been adjusted by the auditor. *Death of a relative* is a dummy variable indicating whether the family of the taxpayer has suffered one or more losses due to the war. *Death of a relative in the battlefield* and *Death of a relative for other reasons* are dummy variables indicating, respectively, whether at least one family member died on the battlefield, or whether one or more family members died only for other reasons. Baseline controls include: the number of taxpayers from the same surname-municipality, and tertile of income FE. Standard errors are clustered at the municipality level in parenthesis. * p < 0.1, ** p < 0.05 and *** p < 0.01

Our initial set of results, reported in Table 1, provide estimates on the difference in the level of tax noncompliance between taxpayers who have (at least) a relative who died during the war compared to those who did not have any loss. Columns 1 to 4 focus on our broader definition of treatment. When no control is included, we find that having a relative that died during the war increases the probability of being noncompliant with the tax law by 1.3 percentage points (p.p.). The coefficient is statistically significant at the 5% level. When including baseline controls, the effect is reduced to 1.1 p.p., and adding surname and province fixed effects leads to an effect of 0.6 p.p. In both cases, the coefficients are statistically different from zero. In Column 4, we add municipality and surname × province fixed effects to the specification and find a 0.4 p.p. higher probability of being identified as tax noncompliant among those taxpayers that had a

[17] A more formal explanation of the consequences of misclassification in our scenario is reported in Appendix B

relative that died during the war, compared to those who did not. This effect corresponds to 3.4% of the average tax noncompliance. The goodness of fit of this model increases significantly as we enrich the specification. In particular, it moves from 0.0003 to 0.015 from the first to the second Column, it amounts to 0.172 in the third Column, and to 0.265 in the fourth Column.

In Columns 5 to 8 of Table 1, we report coefficient estimates that account for the cause of the death of the taxpayer's relative in the army.[18] The estimates suggest that the overall effect is mainly explained by the death of (one or more) family member(s) that occurred on the battlefield rather than for other reasons. The size of the coefficients of *death of a relative on the battlefield* are slightly larger than those of the broader definition of the treatment across the different specifications and have higher levels of significance. Instead, *death of a relative for other reasons* has a smaller size, and it is never statistically significant. We interpret this result as initial evidence that the effect of violence exposure on tax compliance is not a mere consequence of the loss of a relative, that could have a strong negative economic impact making taxpayers more likely to under-report or not report their income.

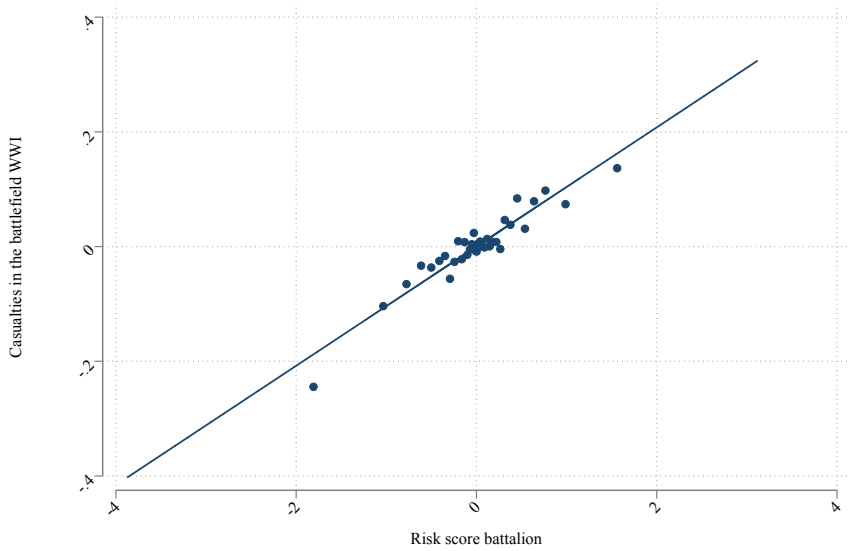
5.2. Instrumental Variables

The set of fixed effects included in the initial specification controls for relevant unobservable factors, partially reducing endogeneity concerns. Yet, this initial strategy does not directly account for potential selection issues at the municipality-surname of origin level. This section provides results that directly address endogeneity issues using an IV strategy. Our instrument is a measure that captures the risk of having a relative that died on the battlefield compared to other reasons. Therefore, these results are specific to the *death of a relative on the battlefield* treatment. Before discussing the results, we report on the instrument's relevance and the exogeneity of the assignment of soldiers to military units.

As a first step, we graphically document the positive relationship between our instrument and the treatment. In Figure 2, we report the relationship between the treatment and the instrument (with the complete specification, as discussed in section 4.2). As ex-

[18] An important assumption behind this analysis is that the family of a soldier that died while in service was aware of the cause of death. This is true in this setting, as the communication to the family reported some key information, including date, place, and cause of death, as reported by several historical documents (e.g., Storia e memoria di Bologna).

Figure 2: First stage – Binscatter



Notes: The figure plots the regression of the instrument (risk score of military units) on the main explanatory variable of the analysis (variable capturing whether a family member died on the battlefield). The specification of the analysis is as in Table 2 (Column 4), and it includes Municipality fixed effects, Surname fixed effects interacted with Province fixed effects and baseline controls which include: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor, the average age of the military victims within the family, number of taxpayers from the same surname-municipality, tertiles of income FE. The gray shaded area indicates the confidence bands. Coefficient=0.104, F-Test=219.

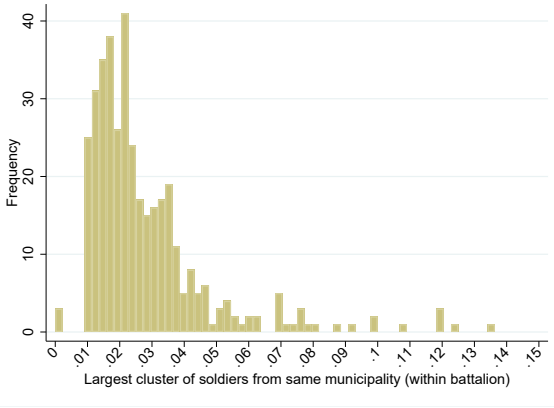
pected, there is a positive relationship with a coefficient of 0.104 and a level of statistical significance that suggests the relevance of the instrument (F-test=219).

As already discussed in section 4.2, a crucial assumption for the validity of our strategy is that the assignment to a given military unit, and therefore the riskiness associated with entering the war, is exogenous with respect to the family's likelihood of compliance with the tax law. We begin by providing a set of descriptive tests meant to highlight the absence of systematic selection of soldiers from a particular geographical area or family into specific military units. Importantly, as already mentioned, in this analysis we only focus on the military victims members of those specializations that were not applying any type of selective recruitment, i.e., artillery, infantry, and the corps of *Bersaglieri*. This may further attenuate the risk of endogenous recruitment.

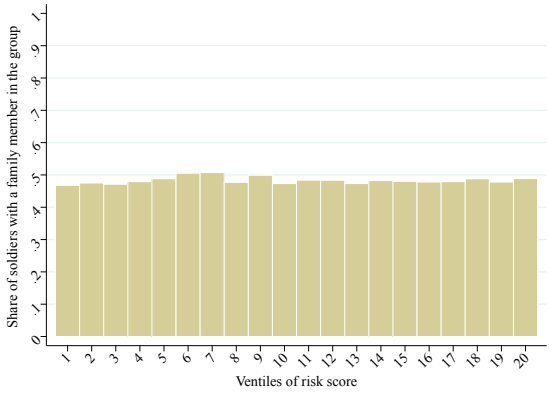
First, soldiers were not assigned to military units because of their area of origin. As we suggest in Section 2, the probability of being assigned to a specific battalion (and military unit) was, by law, not related with the area of origin, as soldiers were intentionally grouped in military units with fellow soldiers from other regions. Figure 3, Panel (a), shows that the largest cluster of soldiers from the same municipality (as a

Figure 3: Selection into military unit

(a) Municipality



(b) Cluster by risk level



Notes: Panel a, shows the distribution of the cluster of soldiers from the same municipalities (as a share of military unit size). The median value is 0.022. Panel b shows the share of soldier with a family member by risk groups (ventiles of risk).

share of the military unit size) was very small in the great majority of battalion. This evidence suggests that most of the soldiers were located in military units where none (or very few) of their peers came from the same municipality.

Further, Figure 3, Panel (b), shows that there is no clear indication that soldiers with a given surname-municipality are assigned to military units with a given level of risk. If there was a systematic selection based on the family of origin, we should have expected some systematic evidence that specific surnames were more often assigned to either low or high-risk military units.

Next, we conduct a detailed examination to determine whether belonging to a specific family (identified by surname and place of origin), which we can consider predetermined with respect to the war, can predict tax compliance. Additionally, we investigate how these predictions correlate with the instrumental variable we are using in our study. We generate predictions of tax compliance based on machine learning techniques (elastic net with 5-fold cross-validation) where the predictors are dummies identifying surname and municipality of origin.[19] Interestingly, we find that our model performs well (mean ROC-

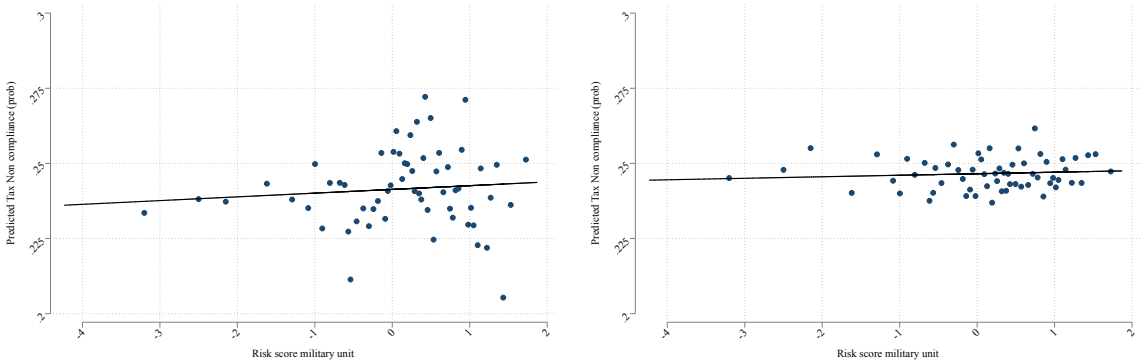
[19]We tuned the hyperparameters via cross-validation and get an $\alpha = 0.0001$, and an L1 ratio=0.9. In other words, we have a model with a small penalty (alpha) and a mix that is heavily biased towards L1 regularization (Lasso).

AUC of 0.780) when predicting tax compliance.[20] This result implies that belonging to a certain family is a good predictor of tax behavior. For identification purposes, it is crucial to exclude that families with a low/high probability of not being compliant with the tax law are systematically assigned to low/high-risk military units.

Figure 4: Predicted tax compliance and military units risk

(a) Only Province FE

(b) All Controls



Notes: Panel (a) shows the analysis conducted with the inclusion of province fixed effects, while panel (b) displays the analysis with the inclusion of the complete set of fixed effects and controls (as in Table 2, Column 4). The analysis of panel (a) shows a coefficient of .0012, with a standard error of .0016 (p-value of 0.428). The analysis of panel (b) shows a coefficient of .0005, with a standard error of .0007 (p-value of 0.455).

Therefore, in Figure 4, we display two binscatter plots, panels A and B, where The y-axis represents the predicted probability of tax noncompliance, while the x-axis corresponds to our instrumental variable. Panel A shows the analysis conducted with a minimum control set (only province fixed effects), while panel B displays the analysis done on the full control set (as discussed in Section 4.2). In both cases, the correlation between the predicted probability of tax noncompliance and our instrument is zero, with a small and statistically insignificant coefficient. This test further supports the exogeneity of our instrumental variable.

Next, the worries about selecting soldiers into specific low-risk military units hinge on the idea that unit risk was predictable. However, war operations followed needs that would change depending on the outcomes of the military confrontation on the battlefield,

[20] Appendix Figure A.6 report a calibration plot that confirms the high quality of the predictions. A closer alignment of the curve with the diagonal indicates better calibration. Discrepancies between the curve and the diagonal line suggest regions of probability where the model might be over- or underestimating the risk.

making it difficult to identify low-risk military units in advance. As proof of this, we find that only 11.8% of military units are in the lowest quintile of the risk distribution for all the years of war, while only 3.8% of them are in the top quintile of the risk distribution for all the years of war. Moreover, as an additional evidence on this, Figure A.7 reports the number of victims over time for the largest 100 military units, which represent more than half of all military victims of WWI (specifically, the 54.8%). Panel a shows the total number of victims, while panel b focuses on the victims in the battlefield. The figure clearly shows that the unit risk is very difficult to predict as, beyond the usual seasonality in military operations, each unit is usually exposed to low and high risk periods over time.

Finally, to provide more specific evidence on the exogeneity of the instrument, we rely on balance checks. Namely, Table A.5 reports coefficient estimates from a number of regressions where we regress the instrument on several taxpayers' observable characteristics. These results show that there is no correlation between the instrument and 1) the taxpayer's income level, 2) their level of education –proxied by whether someone is in a profession that requires specific formal education, and 3) the profession.[21]

Table 2 presents the main results with this IV approach. We begin by reporting OLS coefficients in Panel A (which differs from the analysis reported in Table 1 as we now focus only on families with at least one military victim). The coefficients suggest that having a family member that died on the battlefield increases tax noncompliance. The size of the coefficients varies between 0.028, statistically significant at the 1% level, and 0.011, statistically significant at the 10% level. This effect corresponds to 8.1% of the average tax noncompliance. Interestingly, these coefficients are larger than those reported in Columns 5-8 of Table 1. As the IV results are based only on a sample of families that experienced a loss from the war, a sample that should not suffer from bias due to the first source of selection mentioned in Section 4, having larger coefficients could be due to an attenuation bias in the estimates reported in Table 1. Panel B of Table 2 reports the estimates of the first stage. As already reported, we find a positive and highly significant relationship between our instrument and the treatment. Importantly, the point estimates of the first stage are almost not affected by the inclusion of controls and fixed effects. Moreover, Panel C reports a positive and significant reduced form

[21] We report the same analysis conducted with "family" (surname-municipality) in Table A.6 and the results are similar, with no correlation between the instrument and family characteristics.

Table 2: Effect of War on Tax noncompliance - IV

	(1)	(2)	(3)	(4)
***Panel A*: OLS**				
Death of a relative in the battlefield	0.028***	0.036***	0.011*	0.011*
	(0.007)	(0.006)	(0.006)	(0.006)
Adjusted R^2	0.001	0.019	0.214	0.300
***Panel B*: First stage**				
Risk of military unit	0.102***	0.105***	0.098***	0.104***
	(0.004)	(0.004)	(0.005)	(0.007)
Adjusted R^2	0.062	0.100	0.486	0.735
F-stat	634.13	835.37	332.71	219.19
Mean dep. var.	.808	.808	.812	.815
***Panel C*: Reduced form**				
Risk of military unit	0.019***	0.021***	0.004*	0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
Adjusted R^2	0.003	0.021	0.214	0.300
***Panel D*: IV**				
Death of a relative in the battlefield	0.186***	0.203***	0.045*	0.066**
	(0.031)	(0.028)	(0.026)	(0.030)
N Observations	54,771	54,771	51,486	49,472
Mean dep. var.	.136	.136	.138	.138
Baseline controls		✓	✓	✓
Surname FE			✓	
Province FE			✓	
Municipality FE				✓
Surname FE × Province FE				✓

Notes: The sample includes taxpayers whose families suffered a loss due WWI and whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. The dependent variable is *tax noncompliance*, a binary variable indicating whether the taxpayer's income has been adjusted by the auditor. *Death of a relative on the battlefield* is a binary variable indicating whether at least one family member died on the battlefield. Baseline controls include: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), dummy capturing whether the military victim was young (below the median age), number of taxpayers from the same surname-municipality, terciles of income FE. Standard errors are clustered at the municipality level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$

effect, ranging between 0.007 and 0.019, which is always statistically significant. Finally, Panel D of Table 2 shows the IV coefficients. Consistent with the effect reported in the first stage and the reduced form, we find a positive and significant effect of having the death of a relative on the battlefield on tax noncompliance.[22] The size of the coefficient in the most complete specification is 0.066. This effect corresponds to 48.5% of the average tax noncompliance.[23] Overall, we confirm that the presence of a family member's death on the battlefield has a substantial and statistically significant impact on tax noncompliance.[24] The findings from the IV approach confirm and extend the OLS results, indicating that the observed effect is not due to sample selection bias. The rise in tax noncompliance likely reflects a behavioral response to the loss, possibly indicating a broader dissatisfaction with the state or society after such a traumatic event.

6. Additional results

6.1. Alternative measure of noncompliance

Although we control for location-specific tax authority behavior by introducing municipality fixed effect, there is still concern that auditors are potentially adopting different conducts depending on the exposure of a certain family to the war. For instance, auditors could focus more often on families that suffered a loss from the war, as they might show clear anti-government attitudes. In such cases, assuming that the identification of someone's misbehavior is largely determined by receiving the audit, this could partially explain our main evidence.

To account for this issue, we created a new measure of tax compliance by looking at whether the amount declared by taxpayers is close to specific kink points in the tax

[22] In unreported results, we find that the coefficients are statistically significant also with different cluster specifications (i.e., municipality and surname, surname, province and surname, province).

[23] The significant difference between OLS and IV estimates can be attributed to misclassification bias and the nature of IV estimates. Misclassification biases results, with IV likely overestimating and OLS underestimating the effect (see Section 4, 6.4, and Appendix Figure A.9). Additionally, the IV method provides the Local Average Treatment Effect (LATE), reflecting the average effect only for those who comply with the treatment, who may be more sensitive to traumatic events.

[24] We report in section 6.4 several robustness checks: (i) using the number of battlefield casualties as the treatment variable; (ii) considering as treatment group the families with casualties only from the battlefield; (iii) excluding taxpayers affected by the Spanish flu; (iv) excluding families with army officials; and (v) controlling for taxpayers' professions, frequent surnames, and large cities. Each of these checks confirms the main results.

Table 3: Effect of War on Tax Avoidance

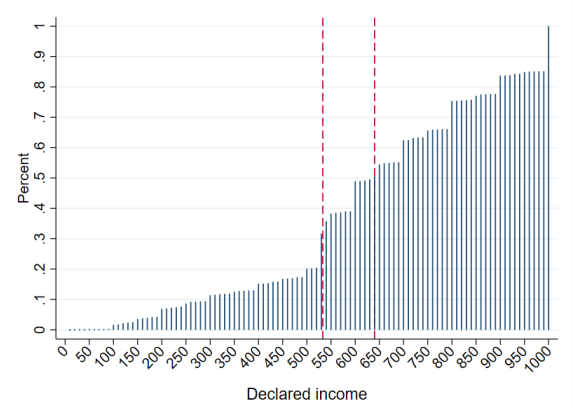
	OLS	First stage	Reduced form	IV
	(1)	(2)	(3)	(4)
Death of a relative in the battlefield	0.021**			0.135*
	(0.011)			(0.069)
Risk of military unit		0.090***	0.012**	
		(0.021)	(0.005)	
N observations	4,297	4,297	4,297	4,297
Mean dep. var.	.020	.020	.020	.020
Baseline controls	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Surname FE × Province FE	✓	✓	✓	✓

Notes: The sample includes taxpayers whose families suffered a loss due to WWI with an income ranging between 500 and 700 lire, and whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. The dependent variable is *tax avoidance*, a binary variable indicating whether the taxpayer's income is just below the exemption thresholds (- 10 lire). *Death of a relative on the battlefield* is a binary variable indicating whether at least one family member died on the battlefield. Additional controls include: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), dummy capturing whether the military victim was young (below the median age), number of taxpayers from the same surname-municipality, dummy indicating incomes rounded at hundred lire, differential for type of taxpayer. Standard errors are clustered at the municipality level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

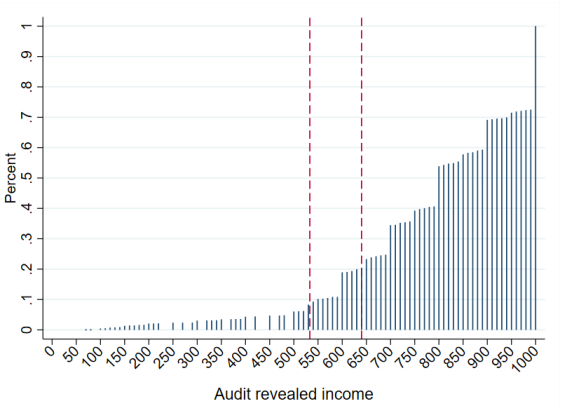
schedule. In particular, there are two important mass points distinguishable by looking at the distribution of declared income: 534 and 640 lire. These are the amounts that define the no-tax area for taxpayers belonging to the categories under analysis. We exploit this feature of the tax system to explore whether taxpayers bunch around these kink points. There is, indeed, historical evidence suggesting that this was a common practice to elude taxes (see, for example, Einaudi, 1980). This is visible in Figure 5, which shows the cumulative income distribution –up to 1,000 lire– of taxpayers that did not receive any audit (panel a) and of those whose income has been audit-adjusted. On the one hand, we observe an increase in density in correspondence with the two kinks in the distribution of non-audited incomes, aside from the standard jumps at certain rounded numbers. On the other hand, no rise is visible in the other distribution, plotting the audit-adjusted incomes. This suggests that the "true" income distribution –corrected by the auditors– is smooth around these values and that any excess mass is likely to be linked with some form of misbehavior. Therefore, we define an alternative measure of noncompliance that identifies whether a taxpayer did declare an income which is in a window of 10 lire below each threshold. Finally, it is important to mention that auditors did not seem to treat taxpayers with a declared income close to these thresholds

Figure 5: Cumulative income distribution

(a) Declared incomes



(b) Audit-adjusted incomes



Notes: Panel a shows the cumulative income distribution for the taxpayers that didn't receive any audit. Panel b shows the cumulative income distribution for the taxpayers whose income has been audit-adjusted. Dotted red lines indicate the bins containing the values 534 and 640. The figure only focuses on declared incomes below 1,000 Lire and shows bins of 25 Lire.

differently from the others.

Table 3 shows the results of the analysis using this alternative measure of tax noncompliance. In this case, we focus only on the sample of families that suffered a loss due to war, and on taxpayers with declared income between 500 and 700 lire. The results confirm the main findings that the main treatment has a positive and significant effect on this additional measure of tax noncompliance in the OLS analysis (Column 1), in the reduced form (Column 3) and in the IV analysis (Column 4). In particular, if we focus on the IV analysis, having a relative that died on the battlefield increases the probability of tax noncompliance by 13.5 p.p. These results confirm that the main effects, that we discussed in section 5.2, are not due to a differential behavior of tax auditors with respect to the family status. Finally, as a robustness check, Figure A.8 plots the coefficient of the IV analysis changing the definition of the tax avoidance variable, using a progressively smaller window of income (reducing the bandwidth from 10 to 2 lire). The results remain robust for these alternative windows of income.

6.2. Effect on income

One simple explanation of our main result is that war represents an economic shock for the families involved, inducing them to evade their taxes more often. For instance, we can expect that members of families that have lost a member during the war might experience a general reduction in income, which eventually affects their propensity to

Table 4: Effect of War on Income

	Full sample		Reduced sample	
	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)
Death of a relative	-0.023***			
	(0.008)			
Death of a relative in the battlefield		-0.022**	-0.006	0.064
		(0.008)	(0.017)	(0.082)
Death of a relative for other reasons		-0.027**		
		(0.013)		
N Obseravations	169,462	169,462	49,472	49,472
Mean dep. var.	7.49	7.49	7.50	7.50
Baseline controls	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Surname FE × Province FE	✓	✓	✓	✓

Notes: The sample includes all taxpayers in columns 1 and 2, and taxpayers whose families suffered a loss due to WWI in columns 3 and 4. The sample always focuses on taxpayers whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. In columns 1 to 3, OLS estimates are reported, while in column 4 IV estimate is reported. The dependent variable is *declared income* (expressed in logarithm). *Death of a relative in the battlefield* and *Death of a relative for other reasons* are dummy variables indicating, respectively, whether at least one family member died on the battlefield, or whether one or more family members died only for other reasons. For columns 1 and 2 baseline controls include: the number of taxpayers from the same surname-municipality. For columns 3 and 4 baseline controls include: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), dummy capturing whether the military victim was young (below the median age), number of taxpayers from the same surname-municipality. Standard errors are clustered at the municipality level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

pay their taxes correctly.[25]

We test this hypothesis by checking whether our alternative definitions of treatment correlate with the reported income, accounting for the different identification strategies and sample compositions. These results are reported in Table 4. In the first two Columns, we show the coefficient estimates of the fixed effect model when applied to the full sample of taxpayers. Coherently with previous evidence, we find a significantly different income level for families that suffered a loss from the war, compared to families that did not. Interestingly, we also find that when splitting the main treatment depending on the reason for death, the coefficients are still significant and very similar in size. In the last two Columns, we focus on the sample of taxpayers with a relative that died during the war, therefore exploiting the reason for death as treatment. Consistent with the

[25]For instance, Dupraz et al. (2021) examine the consequences of father loss for children during the U.S. Civil War, and find that children who lost a father in the war experienced a negative effect on income.

aforementioned results, we show that the declared income is not statistically different for families with a casualty on the battlefield compared to families with a casualty for other reasons, both in the OLS with fixed effects, and the IV.[26]

These results provide interesting insights into the potential channels behind the effect of wartime violence on tax compliance. In particular, we argue that they are generally consistent with an intrinsic motivation rather than an extrinsic motivation channel. First, using the narrow definition of treatment that discriminates only on the reason for death, we did not find differences in income in the restricted sample. Second, analyzing the full sample, we find that for both reasons of death, the income is lower compared to a family that did not suffer any loss. However, as we show in Table 1, Columns (5-8), only those families that lost a family member on the battlefield are significantly less likely to be compliant with the tax law.

6.3. Effect heterogeneity

In this section, we explore the heterogeneous effects of war on tax noncompliance, focusing on how these effects vary based on local, individual, and family characteristics. Our objective is to provide suggestive evidence of the roles played by the state and local communities in mediating these effects. Table 5 presents the results of our reduced form analysis, where the dependent variable in all columns is *tax noncompliance*, defined as a binary indicator of whether the taxpayer's income has been adjusted by the auditor. The primary variable of interest is the risk associated with the military unit, which interacts with several variables: *Army official*, *Medal of Honor*, *municipal deaths above the third tercile*, and *Socialist vote share in 1913 above the third tercile*. We argue these factors may shape social norms, trust in government, and community cohesion, thereby influencing tax compliance behavior.

Columns (1) and (2) focus on family and individual-level outcomes, examining the interactions between the risk associated with military units and the roles within the military or specific honors. In both columns, the coefficient for the risk of military unit is positive and significant, indicating a similar increase in tax noncompliance (0.009 in Column 1 and 0.008 in Column 2), confirming the main evidence in case the interacted

[26]Importantly, the result is the same if we run this test on the reduced sample of the audit-adjusted incomes. The fact that the null result also emerges when looking only at the "correct" incomes (audit-adjusted) reassures that this is not somehow driven by misreporting in the sample of declared incomes, that have not been corrected by the auditors.

Table 5: Effect of War on Tax noncompliance - Mechanisms (Reduced form)

Interacted term:	Dep. var.: Tax noncompliance			
	Army official	Medal of honor	Mun. deaths (third tercile)	Socialist vote (1913) (third tercile)
	(1)	(2)	(3)	(4)
Risk of military unit	0.009***	0.008**	0.010**	0.010**
	(0.003)	(0.003)	(0.004)	(0.004)
Interacted term	-0.011	-0.015	-	-
	(0.009)	(0.013)		
Risk of military unit ×	-0.021**	-0.012	-0.010*	-0.010
Interacted term	(0.009)	(0.011)	(0.006)	(0.006)
N observations	49,472	49,472	49,472	49,472
Mean dep. var.	.138	.138	.138	.138
Baseline controls	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓
Surname FE × Province FE	✓	✓	✓	✓
F-stat Risk + *Inter. term*	4.50	3.08	3.10	2.96

Notes: The sample includes taxpayers whose families suffered a loss due WWI and whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. The dependent variable is *tax noncompliance*, a binary variable indicating whether the taxpayer's income has been adjusted by the auditor. *Risk of military unit* is the average risk score associated with the military units to which the family member(s) have been assigned. Baseline controls include: dummy variable indicating that at least one military victim (within the family) was an army official [omitted in column 1], dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver, or gold medal) [omitted in column 2], dummy capturing whether the military victim was young (below the median age), number of taxpayers from the same surname-municipality, terciles of income FE. Standard errors are clustered at the municipality level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$

terms have value 0 (i.e., no official in the military and no medals received). However, the interaction terms in both columns show negative coefficients that are larger in absolute value than the main effects. Specifically, the interaction term for Army Official (1) is -0.021, statistically significant, and for Medal of Honor in Column in Column (2), it is -0.012, not statistically significant, indicating a moderating effect. For families without an army official or a Medal of Honor recipient, we find results in line with the main effect; therefore, the risk associated with a military unit increases tax noncompliance. However, for families with an army official, the overall effect of military unit risk on tax noncompliance is $0.009 - 0.021 = -0.012$. The significant mitigating effect of the interaction term suggests that these families are less likely to evade taxes despite the increased military unit risk. Families of army officials might feel proud of their military role and have a strong sense of duty and loyalty to the state. The prestige and honor associated with an official who died in the war may foster greater commitment to civic responsibilities, including tax compliance. In contrast, families of regular soldiers, often conscripted unwillingly, may not share the same level of voluntary engagement or con-

nection to the state's ideals, potentially leading to less emphasis on civic duties like tax compliance. For families with a Medal of Honor recipient, the negative interaction terms suggest a reduction in the positive effect of military unit risk on tax noncompliance, $0.008 - 0.012 = -0.004$, which is jointly statistically significant. This indicates that the presence of a Medal of Honor recipient might actually reduce the impact of military unit risk on tax noncompliance. Such recognition can strengthen the family's positive connection with governmental institutions, enhancing trust and compliance with state policies.

Columns (3) and (4) extend the analysis by considering community-level variables, specifically the number of municipal deaths above the third tercile and the percentage of Socialist votes in the 1913 elections above the third tercile. In Column (3), the coefficient for the risk of military units is again positive and significant. This indicates that higher military unit risk is associated with increased tax noncompliance in areas with a low number of military victims. The interaction term with municipal deaths in the top tercile is -0.010 and statistically significant at the ten percentage point level. The F-statistic for the overall effect (Risk of military unit + Interaction term) is 3.10, indicating a marginally significant level of the combined effects. Overall, when the losses are widespread in the community, the shared negative experience slightly mitigates the effect on tax compliance. In Column (4), the risk of military unit coefficient is 0.010, and the interaction term with Socialist vote percentage in the top tercile is -0.010, which is barely insignificant. This negative interaction term indicates a mitigating effect of higher Socialist support on the positive relationship between military unit risk and tax non-compliance. Specifically, for municipalities with high number of Socialist votes, the overall effect of military unit risk on tax noncompliance becomes $0.010 - 0.010 = 0.000$, effectively neutralizing the baseline positive effect. This finding suggests that stronger pre-existing preferences for redistribution may counteract the tendency towards tax noncompliance among those experiencing higher military unit risks.

These findings suggest a narrative of state recognition and reciprocation for loss. The formal acknowledgment of sacrifices, whether through recognition of relatives' military status or medals, indicates the state's appreciation and understanding of personal loss, potentially providing consolation and mitigating negative impacts. Furthermore, communities with significant exposure to war and a pre-existing preference for redistribution likely develop stronger social cohesion and support networks, enhancing reciprocity within the community.

6.4. Robustness checks

In this section, we discuss a series of additional results and robustness checks that we conducted to confirm the validity of the main evidence.

Initially, we examine the impact of misclassification on our findings. As previously noted, such discrepancies will likely diminish the OLS analysis coefficient while inflating the IV estimation coefficient. We use the formal derivation in Appendix Section B to provide some initial evaluation. For example, the estimate of the OLS coefficients in Table 1 are all supposed to be downward biased. Assuming a high measure of total misclassification (FPR + FNR) equal to 40%, we would have that the true estimate of *Death of a relative* in Column 4 is supposed to be equal to $0.007 = 0.004/(1 - 0.40)$. Instead, when we look at the estimate in Column 8, we would have to account only for the level of false positive for both *Death of a relative in the battlefield* (π_{13}) and *Death of a relative for other reasons* (π_{23}). Assuming a level of 20%, we would have the real estimate for *Death of a relative in the battlefield* is $0.006 = 0.005/(1 - 0.20)$ and *Death of a relative for other reasons* $0.0012 = 0.001/(1 - 0.20)$. Appendix Figure A.9 provides a flexible evaluation of the effect of misclassification on the results reported in Table 2. In this case, we observe that even with a high false positive rate (e.g., 40%), the corrected IV estimate (0.04) remains higher than the OLS estimate (0.018). However, the difference between the two β coefficients would persists despite the correction.

Next, in the previous analyses we have always considered a binary treatment that indicated whether the taxpayer's family suffered *at least* one loss on the battlefield. We conduct the analysis here by adopting a different treatment definition. As families may lose more than one member due to the war, it may be important to explore the intensive margin, studying the impact of the number of casualties on tax noncompliance. Table 6 (Column 1) shows the IV analysis using the number of relatives dying on the battlefield as treatment. The results show that having an additional casualty increases the probability of misbehavior by an additional 3.9 p.p., suggesting that the effect increases with the number of casualties.

Second, we adopt a treatment variable capturing whether the family had casualties *only* on the battlefield (one or more), in order to focus on a more specific group of families that didn't experience any loss due to reasons other than fighting. Also in this case, the results are confirmed, as displayed in Table 6 (Column 2).

Third, as displayed in Figure 1, the trend in casualties for reasons other than fighting

is mainly flat, with the exception of the Spanish flu wave in 1918. To account for this shock, we rerun the main IV analysis, excluding the taxpayers whose relative(s) died because of the Spanish flu during the worst pandemic wave (September 1918-February 1919). The results are reported in Table 6 (Column 3) and confirm the previous findings.

Fourth, we may be concerned that army officials may have received privileged treatment while at the front, compared to footsoldiers, affecting their likelihood of dying or their cause of death. Since this could also be correlated with some specific characteristics of their family, we conduct the main IV analysis excluding families with at least one army official. The results are even stronger and more precisely estimated compared the main ones, as displayed in Table 6 (Column 4).

Table 6: Additional results and Robustness checks - IV

	Inten. Margin	Strict treatment definition	Excluding Spanish flu	Only footsoldiers	Controlling for professions	Excluding frequent surnames	Excluding big cities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Death of a relative in the battlefield (number)	0.039**						
	(0.018)						
Death of a relative in the battlefield		0.055**	0.066*	0.084***	0.064**	0.073**	0.066**
		(0.025)	(0.040)	(0.032)	(0.030)	(0.033)	(0.030)
N Observations	49,472	49,472	46,484	40,559	49,472	38,183	44,881
Mean dep. var.	.138	.138	.138	.147	.138	.142	.150
Baseline controls	✓	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓	✓
Surname FE	✓	✓	✓	✓	✓	✓	✓
Surname FE × Province FE	✓	✓	✓	✓	✓	✓	✓

Notes: The sample includes taxpayers whose families suffered a loss due to WWI and whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. The dependent variable is *tax noncompliance*, a binary variable indicating whether the taxpayer's income has been adjusted by the auditor. *Death of a relative on the battlefield* is a binary variable indicating whether at least one family member died on the battlefield. In column 2 it captures families that have only members who died on the battlefield. *Death of a relative in the battlefield (number)* is a variable indicating the number of family members that died on the battlefield. In column 4 the sample only includes footsoldiers. In column 6 frequent surnames, defined as the top 10 surnames (accounting for 45.7% of the sample), are excluded from the sample. In column 7 the cities in the top deciles are excluding from the sample. Baseline controls include: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), dummy capturing whether the military victim was young (below the median age), number of taxpayers from the same surname-municipality, terciles of income FE.. The analysis of column 6 also includes the dummy for a university degree and for the profession of liberal arts as well as thirty dummy variables indicating the different professions. Standard errors are clustered at the municipality level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Lastly, as an additional robustness test, we estimate the main IV model controlling for all taxpayers' professions (Column 5 of Table 6), excluding frequent surnames, those in the top 10% of surname distribution (Column 6 of Table 6), and large cities, top 10% of cities in terms of population (Column 7 of Table 6). The main results are always confirmed.

7. Conclusion

This paper studies the effect of war violence exposure on individuals' willingness to comply with the law in the context of tax compliance. Using digitized historical administrative records from Italy following World War I, we found that exposure to war violence has a negative effect on tax compliance. In order to reduce potential bias in the estimates, a fixed effects model and an instrumental variables strategy were applied. The fixed effects model controlled for unobservable characteristics at the geographical and family level, while the instrumental variables strategy exploited exogenous variation in the relative risk of dying on the battlefield versus dying for other reasons, as determined by the assigned military unit. Moreover, this negative effect on tax compliance seems to be unrelated to auditors' behavior and seems to be driven mostly by taxpayers' intrinsic motives. The results of this study diverge from existing empirical findings, which often suggest a positive link between war exposure and willingness to contribute to the state based on survey or experimental evidence. Our research adds to the tax compliance literature by highlighting the significant role of intrinsic motives, showing that trust and respect for the state are crucial for compliance.

Our findings imply that war violence negatively impacts tax compliance, potentially undermining the effectiveness of wartime institutional reforms. We also explore heterogeneity in our results, showing that the negative impact on tax compliance is mitigated in specific circumstances. For instance, when a deceased relative was honored with a medal or was a military official, the effect is null or reversed. Similarly, families in areas with high pre-war prosocial attitudes or significant war exposure experienced less negative impact, suggesting that state recognition and pre-existing community values can buffer against the erosion of trust.

The historical context and type of conflict are critical for interpreting and generalizing these results. Therefore, future research should explore how war violence exposure affects behavior and attitudes in various contexts and times. This continued investigation will

help to provide a better understanding of the long-term impacts of war on state functions and the factors that mediate these effects.

List of Figures

List of Tables

References

Acconcia, A., M. D'Amato, R. Martina, and M. Ratto (2022). The response of taxpayer compliance to the large shock of italian unification. *European Journal of Political Economy 73*, 102158.

Acemoglu, D., G. De Feo, G. De Luca, and G. Russo (2022). War, socialism, and the rise of fascism: an empirical exploration. *The Quarterly Journal of Economics 137*(2), 1233–1296.

Advani, A., W. Elming, and J. Shaw (2021, 09). The Dynamic Effects of Tax Audits. *The Review of Economics and Statistics*, 1–45.

Aigner, D. J. (1973). Regression with a binary independent variable subject to errors of observation. *Journal of Econometrics 1*(1), 49–59.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics 30*, 98–108.

Bauer, M., C. Blattman, J. Chytilová, J. Henrich, E. Miguel, and T. Mitts (2016). Can war foster cooperation? *Journal of Economic Perspectives 30*(3), 249–74.

Becker, S. O., A. Ferrara, E. Melander, L. Pascali, et al. (2020). *Wars, taxation and representation: Evidence from five centuries of German history*. Centre for Economic Policy Research.

Bellows, J. and E. Miguel (2009). War and local collective action in sierra leone. *Journal of public Economics 93*(11-12), 1144–1157.

Besley, T. (2020). State capacity, reciprocity, and the social contract. *Econometrica 88*(4), 1307–1335.

Besley, T. and T. Persson (2009). The origins of state capacity: Property rights, taxation, and politics. *American economic review 99*(4), 1218–44.

Besley, T. and T. Persson (2010). State capacity, conflict, and development. *Econometrica 78*(1), 1–34.

Bianchi, R. (2014). Social conflict and control, protest and repression (italy). In D. Ute, P. Gatrell, O. Janz, H. Jones, J. Keene, A. Kramer, and B. Nasson (Eds.), *1914-1918-online. International Encyclopedia of the First World War*. Freie Universität Berlin, Berlin.

Bingley, P. and A. Martinello (2017). Measurement error in income and schooling and the bias of linear estimators. *Journal of labor economics 35*(4), 1117–1148.

Bowles, S. (2006). Group competition, reproductive leveling, and the evolution of human altruism. *science 314*(5805), 1569–1572.

Bühler, M. and A. Madestam (2022). State repression, exit, and voice: Living in the shadow of cambodia's killing fields. *Stockholm University*.

Caffarelli, E. and C. Marcato (2008). I cognomi d'italia: Dizionario storico ed etimologico, utet.

Cagé, J., A. Dagorret, P. Grosjean, and S. Jha (2020). Heroes and villains: The effects of combat heroism on autocratic values and nazi collaboration in france. Technical report, Sciences Po.

Caprettini, B. and H.-J. Voth (2023). New deal, new patriots: How 1930s government spending boosted patriotism during world war ii. *The Quarterly Journal of Economics 138*(1), 465–513.

Cassar, A., P. Grosjean, and S. Whitt (2013). Legacies of violence: trust and market development. *Journal of Economic Growth 18*(3), 285–318.

Clementini, P., C. Bertelli, and E. Scandale (1934). *Commento alle leggi sulla imposta di ricchezza mobile*. Number Bd. 2 in Commento alle leggi sulla imposta di ricchezza mobile. Unione Tipografico-Editrice Torinese.

Collier, P. et al. (2003). *Breaking the conflict trap: Civil war and development policy*. World Bank Publications.

Conzo, P. and F. Salustri (2019). A war is forever: The long-run effects of early exposure to world war ii on trust. *European Economic Review 120*, 103313.

Cummings, R. G., J. Martinez-Vazquez, M. McKee, and B. Torgler (2009). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization 70*(3), 447–457.

De Juan, A. and J. H. Pierskalla (2016). Civil war violence and political trust: Microlevel evidence from nepal. *Conflict Management and Peace Science 33*(1), 67–88.

De Stefani, A. (1923). *Documenti sulla condizione finanziaria ed economica dell'Italia*. Libreria dello stato.

DeBacker, J., B. T. Heim, A. Tran, and A. Yuskavage (2018). Once bitten, twice shy? the lasting impact of enforcement on tax compliance. *The Journal of Law and Economics 61*(1), 1–35.

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

DiTraglia, F. J. and C. Garcia-Jimeno (2019). Identifying the effect of a mis-classified, binary, endogenous regressor. *Journal of Econometrics 209*(2), 376–390.

Doerrenberg, P. and A. Peichl (2013). Progressive taxation and tax morale. *Public Choice 155*(3), 293–316.

Dupraz, Y., A. Ferrara, et al. (2021). Fatherless: The long-term effects of losing a father in the us civil war. Technical report, Competitive Advantage in the Global Economy (CAGE).

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2016). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in germany. *American Economic Journal: Economic Policy 8*(3), 203–32.

Einaudi, L. (1980). *Interventi e Relazioni parlamentari, Volume I,*. Fondazione Luigi Einaudi.

Feldman, N. and J. Slemrod (2009). War and taxation: When does patriotism overcome the free-rider impulse? *The new fiscal sociology: Taxation in comparative and historical perspective*, 138–154.

Gabbuti, G. et al. (2022). Those who were better off: Capital and top incomes in fascist italy. Technical report, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa Italy.

Gagliarducci, S. and M. Manacorda (2020). Politics in the family: Nepotism and the hiring decisions of italian firms. *American Economic Journal: Applied Economics 12*(2), 67–95.

Galletta, S. and T. Giommoni (2022, 01). The Effect of the 1918 Influenza Pandemic on Income Inequality: Evidence from Italy. *The Review of Economics and Statistics 104*(1), 187–203.

Gennaioli, N. and H.-J. Voth (2015). State capacity and military conflict. *The Review of Economic Studies 82*(4), 1409–1448.

Gilligan, M. J., B. J. Pasquale, and C. Samii (2014). Civil war and social cohesion: Lab-in-the-field evidence from nepal. *American Journal of Political Science 58*(3), 604–619.

Grosjean, P. (2014). Conflict and social and political preferences: Evidence from world war ii and civil conflict in 35 european countries. *Comparative Economic Studies 56*(3), 424–451.

Guerra, A. and B. Harrington (2018). Attitude–behavior consistency in tax compliance: A cross-national comparison. *Journal of Economic Behavior & Organization 156*, 184–205.

Jha, S. and S. Wilkinson (2012). Does combat experience foster organizational skill? evidence from ethnic cleansing during the partition of south asia. *American Political Science Review 106*(4), 883–907.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica 79*(3), 651–692.

Levi, M. (1988). *Of rule and revenue*. Univ of California Press.

Luttmer, E. F. and M. Singhal (2014). Tax morale. *Journal of economic perspectives 28*(4), 149–168.

Meyer, B. D. and N. Mittag (2017). Misclassification in binary choice models. *Journal of Econometrics 200*(2), 295–311.

Mortara, G. (1925). *La salute pubblica in Italia durante e dopo la guerra*. Number v. 9. G. Laterza & figli.

Putnam, R. D. (1995). Tuning in, tuning out: The strange disappearance of social capital in america. *PS: Political science & politics 28*(4), 664–683.

Rohner, D., M. Thoenig, and F. Zilibotti (2013). Seeds of distrust: Conflict in uganda. *Journal of Economic Growth 18*(3), 217–252.

Scheve, K. and D. Stasavage (2010). The conscription of wealth: mass warfare and the demand for progressive taxation. *International organization 64*(4), 529–561.

Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature 57*(4), 904–54.

Voors, M. J. and E. H. Bulte (2014). Conflict and the evolution of institutions: Unbundling institutions at the local level in burundi. *Journal of Peace Research 51*(4), 455–469.

Zugaro, F. (1926). *BOLLETTINO DELL' UFFICIO STORICO. L'Albo d'Oro dei Caduti per l'Italia nella Guerra Mondiale*. Stato Maggiore del Real Esercito.

A. Appendix - Additional Figures and Tables

Figure A.1: Assignment military recruits across regiment

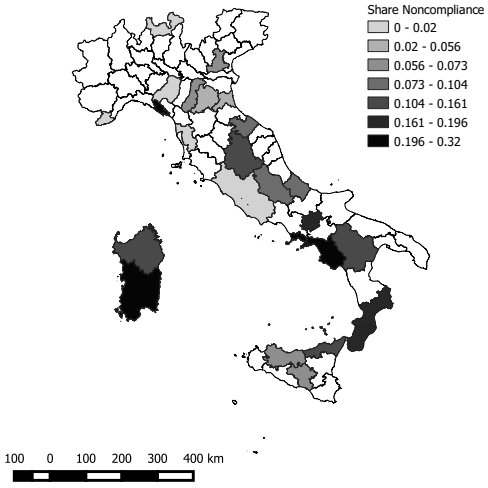
TABELLA *di riparto degli uomini assegnati ai reggimenti di fanteria di linea eccettuato i granatieri.*

REGGIMENTI di fanteria	DISTRETTI CHE DEVONO DARE LE RECLUTE E QUOTA RISPETTIVA																TOTALE per reggim.
1° fanteria	Alessand.	81	Bergamo	58	Siena	69	Napoli	40	Palermo	89	Arezzo	10	Cagliari	13			360
2° id.	Alessand.	81	Verona	57	Perugia	58	Avellino	71	Trapani	76	Arezzo	10	Cagliari	7			380
3° id.	Milano	73	Padova	80	Livorno	63	Napoli	47	Regg. Cal.	72	Arezzo	15	Cagliari	10			360
4° id.	Genova	54	Bergamo	62	Bologna	64	Avellino	71	Palermo	90	Arezzo	10	Cagliari	9			360
5° id.	Genova	60	Parma	65	Macerata	60	Avellino	67	Palermo	84	Lucca	15	Cagliari	13			364
6° id.	Genova	56	Parma	65	Siena	68	Avellino	66	Lecce	85	Arezzo	10	Cagliari	10			360
7° id.	Alessand.	79	Venezia	78	Livorno	62	Napoli	40	Palermo	90	Cagliari	11	—				360
8° id.	Alessand.	79	Udine	99	Pesaro	61	Napoli	40	Catanzaro	63	Arezzo	10	Cagliari	8			360
9° id.	Torino	50	Modena	80	Macerata	60	Salerno	62	Palermo	86	Arezzo	10	Cagliari	12			360
10° id.	Novara	79	Bergamo	50	Ancona	54	Foggia	72	Palermo	84	Arezzo	10	Cagliari	11			360
11° id.	Torino	57	Brescia	70	Roma	66	Salerno	63	Bari	86	Arezzo	10	Cagliari	8			360
12° id.	Cuneo	70	Cremona	71	Siena	70	Campobas.	60	Cosenza	69	Lucca	20	Cagliari	13			373
13° id.	Alessand.	79	Padova	81	Roma	65	Napoli	40	Trapani	75	Arezzo	10	Cagliari	10			360
14° id.	Alessand.	78	Piacenza	58	Firenze	84	Caserta	74	Caltaniss.	57	Cagliari	9	—				360
15° id.	Genova	56	Piacenza	87	Perugia	58	Teramo	76	Regg. Cal.	68	Arezzo	7	Cagliari	8			360
16° id.	Torino	55	Piacenza	86	Perugia	60	Teramo	84	Cosenza	50	Arezzo	15	Cagliari	10			360
17° id.	Milano	60	Treviso	70	Ravenna	79	Benevento	56	Bari	82	Cagliari	13	—				360
18° id.	Milano	63	Parma	57	Ravenna	79	Teramo	73	Bari	82	Cagliari	6	—				360
19° id.	Torino	63	Pavia	85	Ancona	61	Benevento	64	Catanzaro	66	Arezzo	10	Cagliari	11			360
20° id.	Cuneo	72	Brescia	55	Pesaro	81	Avellino	70	Catanzaro	62	Arezzo	15	Cagliari	13			368
21° id.	Cuneo	65	Padova	80	Livorno	60	Campobas	50	Bari	82	Arezzo	10	Cagliari	13			360
22° id.	Cuneo	72	Bergamo	48	Livorno	67	Potenza	78	Trapani	75	Lucca	10	Cagliari	10			360
23° id.	Torino	49	Verona	56	Firenze	88	Salerno	60	Caltaniss.	79	Lucca	50	Cagliari	10			392
24° id.	Alessand.	84	Verona	55	Roma	80	Campobas.	57	Messina	61	Lucca	20	Sassari	14			371
25° id.	Alessand.	72	Pavia	62	Firenze	84	Avellino	58	Bari	78	Arezzo	10	Sassari	9			373
26° id.	Novara	95	Brescia	61	Roma	76	Campobas.	57	Messina	54	Arezzo	10	Cagliari	9			362
27° id.	Cuneo	74	Verona	61	Livorno	62	Caserta	86	Messina	50	Lucca	40	Cagliari	12			385
28° id.	Como	98	Brescia	62	Livorno	61	Salerno	68	Messina	53	Lucca	40	Cagliari	16			398
29° id.	Cuneo	66	Cremona	61	Perugia	55	Potenza	73	Palermo	82	Arezzo	15	Cagliari	8			360
30° id.	Torino	44	Parma	61	Aquila	98	Salerno	63	Palermo	80	Cagliari	14	—				360
31° id.	Cuneo	65	Cremona	68	Roma	70	Chieti	97	Messina	44	Arezzo	10	Cagliari	10			364
32° id.	Torino	58	Regg. Em.	68	Aquila	98	Napoli	42	Caltaniss.	71	Arezzo	15	Cagliari	9			360
33° id.	Genova	68	Venezia	88	Perugia	55	Campobas	54	Catanzaro	65	Lucca	20	Sassari	10			360
34° id.	Genova	54	Modena	81	Bologna	74	Salerno	63	Caltaniss	68	Arezzo	10	Sassari	10			360
35° id.	Genova	46	Padova	88	Siena	69	Napoli	36	Lecce	95	Arezzo	15	Sassari	11			360
36° id.	Novara	83	Regg Em.	44	Macerata	62	Avellino	66	Palermo	80	Arezzo	15	Sassari	10			360
37° id.	Milano	75	Venezia	78	Ravenna	80	Napoli	40	Girgenti	75	Arezzo	6	Sassari	6			360

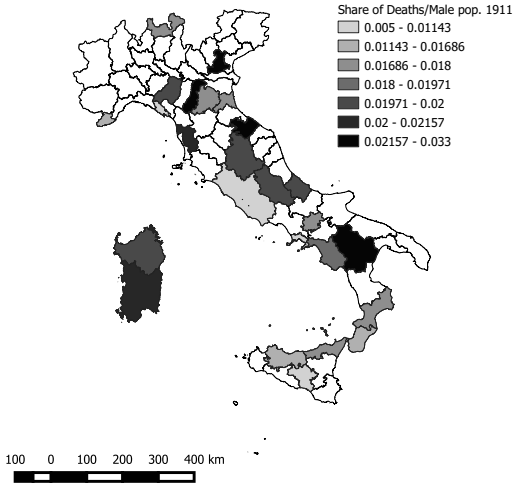
Notes: This figure reports a table about the distribution of the infantry recruits by military district assigned to each regiment in 1880 (Giornale Militare - parte seconda, 1880)

Figure A.2: Geographical distribution Tax noncompliance and WWI Casualties (Province-level)

(a) Tax Noncompliance

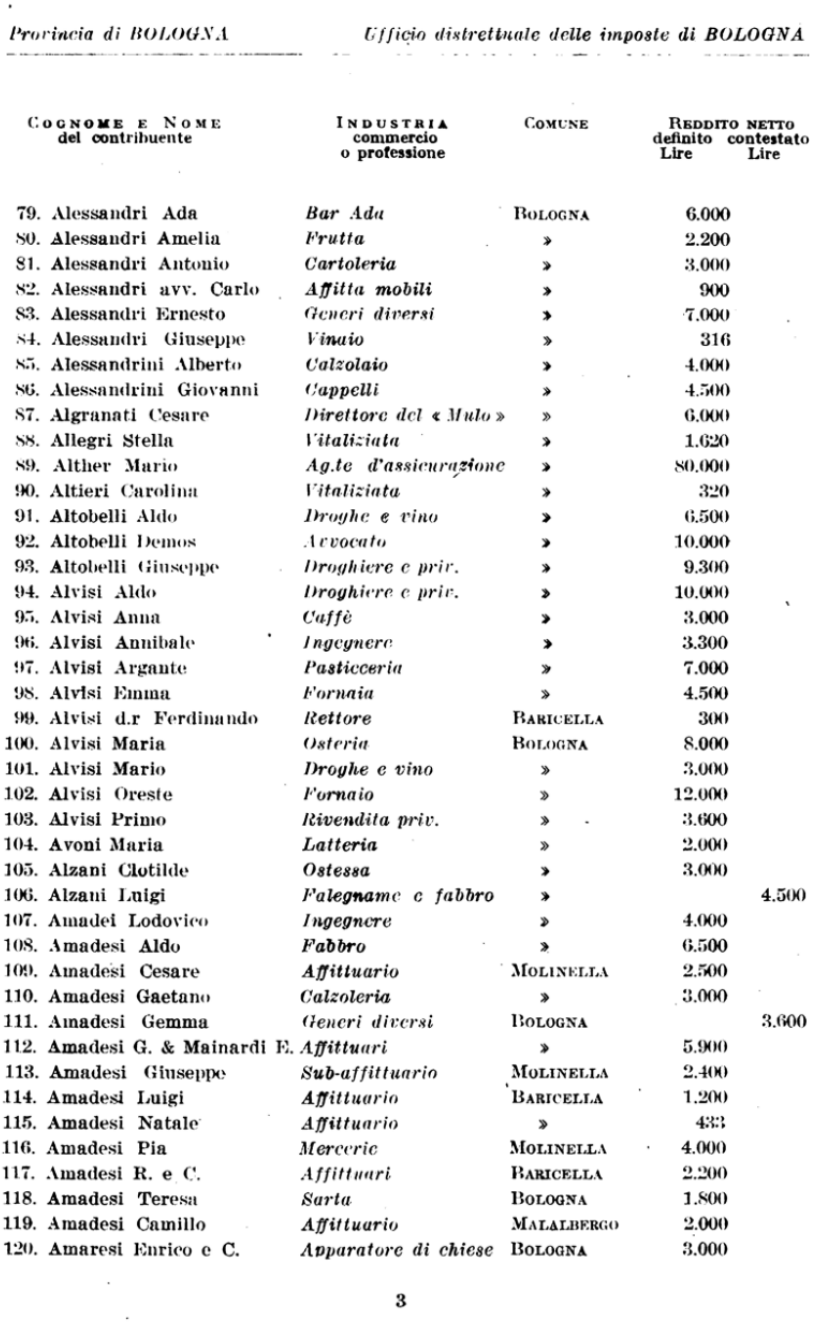


(b) WWI Deaths



Notes: Panel (a) displays the average level of tax noncompliance (i.e., Audited-adjusted Income Declaration) at the province level, while panel (b) show the geographical distribution of death in WWI.

Figure A.3: Example of tabulation - Province of Bologna



Provincia di BOLOGNA — *Ufficio distrettuale delle imposte di BOLOGNA*

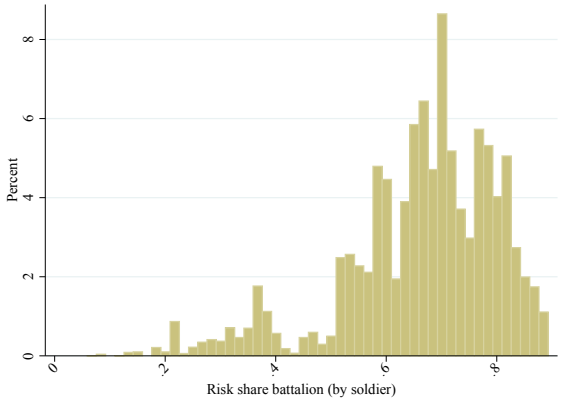
COGNOME E NOME del contribuente	INDUSTRIA commercio o professione	COMUNE	REDDITO NETTO definito Lire	contestato Lire
79. Alessandri Ada	*Bar Ada*	BOLOGNA	6.000	
80. Alessandri Amelia	*Frutta*	»	2.200	
81. Alessandri Antonio	*Cartoleria*	»	3.000	
82. Alessandri avv. Carlo	*Affitta mobili*	»	900	
83. Alessandri Ernesto	*Generi diversi*	»	7.000	
84. Alessandri Giuseppe	*Vinaio*	»	316	
85. Alessandrini Alberto	*Calzolaio*	»	4.000	
86. Alessandrini Giovanni	*Cappelli*	»	4.500	
87. Algranati Cesare	*Direttore del « Mulo »*	»	6.000	
88. Allegri Stella	*Vitaliziata*	»	1.620	
89. Alther Mario	*Ag.te d'assicurazione*	»	80.000	
90. Altieri Carolina	*Vitaliziata*	»	320	
91. Altobelli Aldo	*Droghe e vino*	»	6.500	
92. Altobelli Demos	*Avvocato*	»	10.000	
93. Altobelli Giuseppe	*Droghiere e priv.*	»	9.300	
94. Alvisi Aldo	*Droghiere e priv.*	»	10.000	
95. Alvisi Anna	*Caffè*	»	3.000	
96. Alvisi Annibale	*Ingegnere*	»	3.300	
97. Alvisi Argante	*Pasticceria*	»	7.000	
98. Alvisi Emma	*Fornaia*	»	4.500	
99. Alvisi d.r Ferdinando	*Rettore*	BARICELLA	300	
100. Alvisi Maria	*Osteria*	BOLOGNA	8.000	
101. Alvisi Mario	*Droghe e vino*	»	3.000	
102. Alvisi Oreste	*Fornaio*	»	12.000	
103. Alvisi Primo	*Rivendita priv.*	»	3.600	
104. Avoni Maria	*Latteria*	»	2.000	
105. Alzani Clotilde	*Ostessa*	»	3.000	
106. Alzani Luigi	*Falegname e fabbro*	»		4.500
107. Amadei Lodovico	*Ingegnere*	»	4.000	
108. Amadesi Aldo	*Fabbro*	»	6.500	
109. Amadesi Cesare	*Affittuario*	MOLINELLA	2.500	
110. Amadesi Gaetano	*Calzoleria*	»	3.000	
111. Amadesi Gemma	*Generi diversi*	BOLOGNA		3.600
112. Amadesi G. & Mainardi E.	*Affittuari*	»	5.900	
113. Amadesi Giuseppe	*Sub-affittuario*	MOLINELLA	2.400	
114. Amadesi Luigi	*Affittuario*	BARICELLA	1.200	
115. Amadesi Natale	*Affittuario*	»	433	
116. Amadesi Pia	*Merceria*	MOLINELLA	4.000	
117. Amadesi R. e C.	*Affittuari*	BARICELLA	2.200	
118. Amadesi Teresa	*Sarta*	BOLOGNA	1.800	
119. Amadesi Camillo	*Affittuario*	MALALBERGO	2.000	
120. Amaresi Enrico e C.	*Apparatore di chiese*	BOLOGNA	3.000	

3

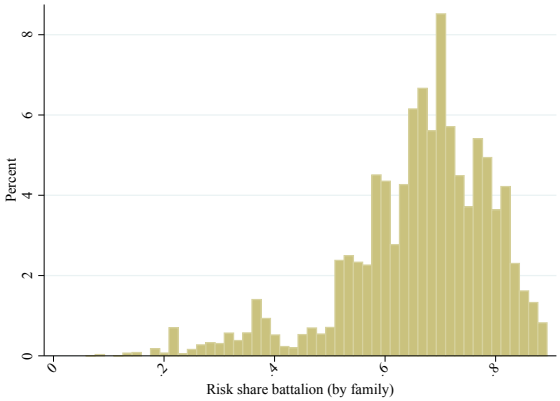
Notes: This figure reports an example of tabulation containing the income declarations of the taxpayers in the fiscal district of Bologna – Province of Bologna.

Figure A.4: Distribution risk score

(a) Risk score military unit (by individual soldier)



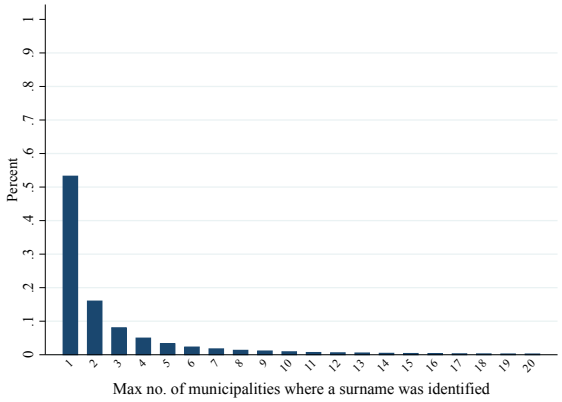
(b) Risk score military unit (by family)



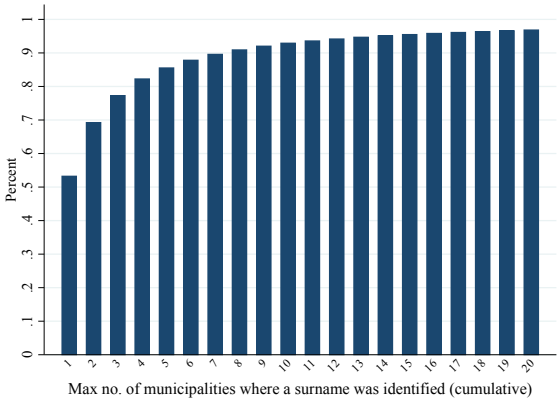
Notes: Panel a shows the distribution of the military unit risk score for the dataset of the military casualties of WWI (universe of victims). Panel b shows the distribution of the military unit risk score for the dataset of taxpayers' families (with at least a casualty).

Figure A.5: Surname distribution

(a) Distribution

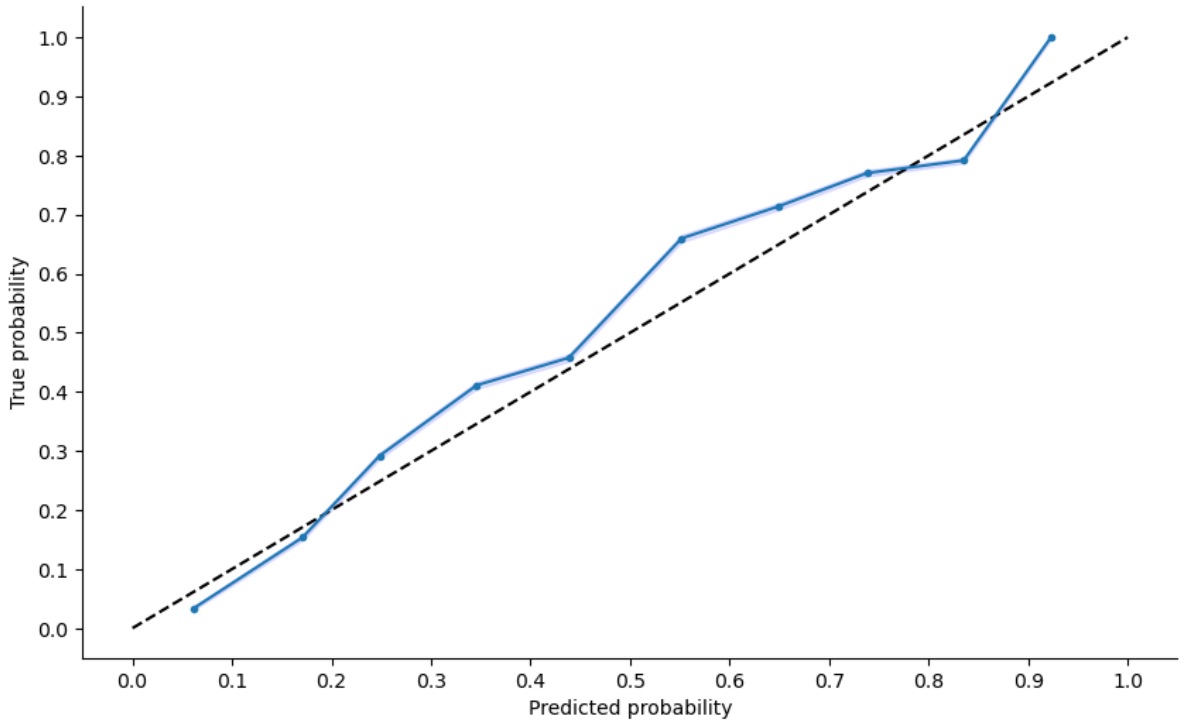


(b) Cumulative



Notes: Panel a shows the distribution of the maximum number of municipalities in which a surname is identified, according to the *Albo dei Caduti*. Panel b shows the corresponding cumulative distribution.

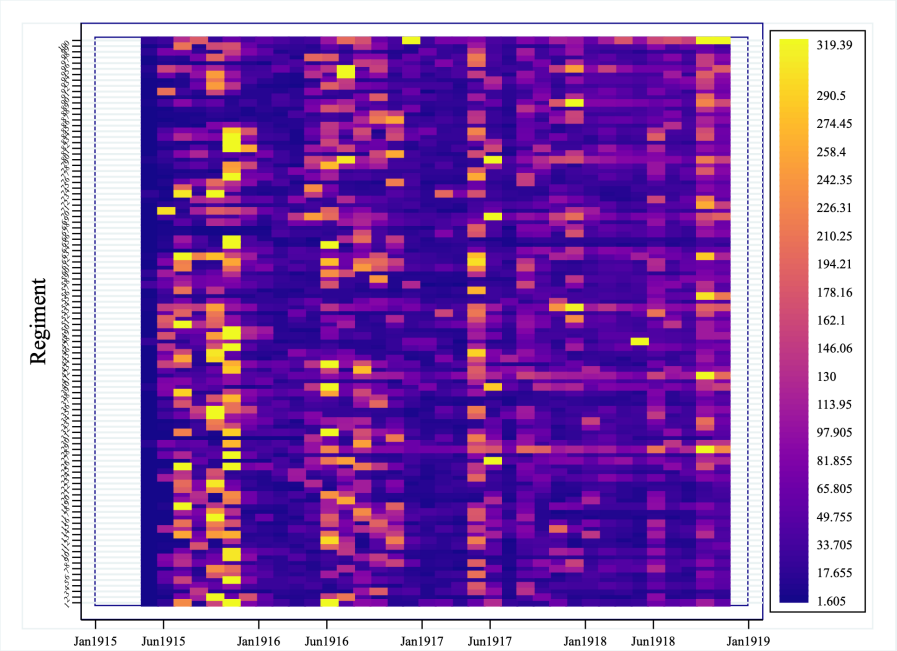
Figure A.6: Calibration plot



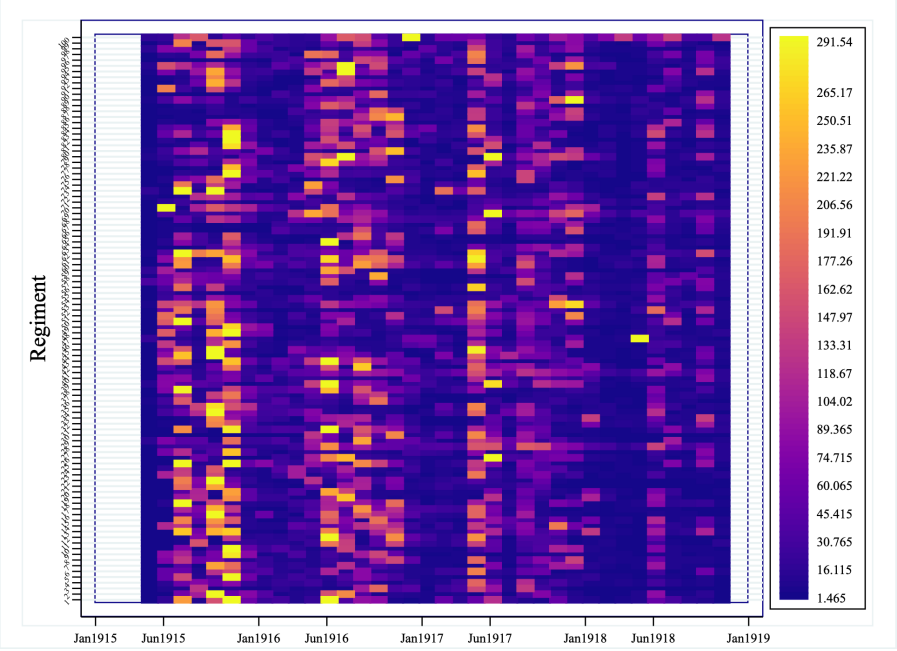
Notes: The figure plots the calibration curve for our predictive model on the test set. The diagonal dashed line represents a perfectly calibrated model, where the predicted probabilities would match the true outcomes. The curve plotted with markers represents our model's performance.

Figure A.7: Casualties by unit (regiment) over time - Largest 100

(a) All casualties

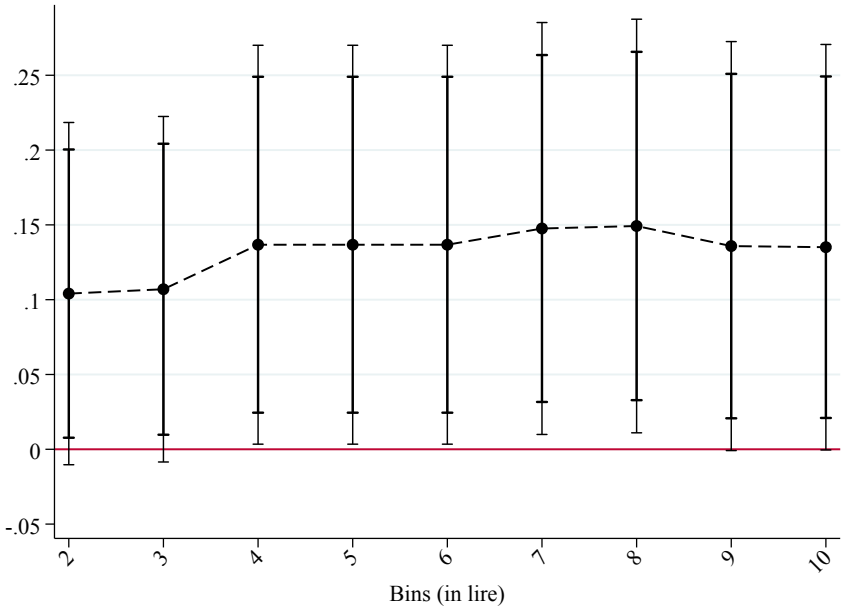


(b) Casualties in the battlefield



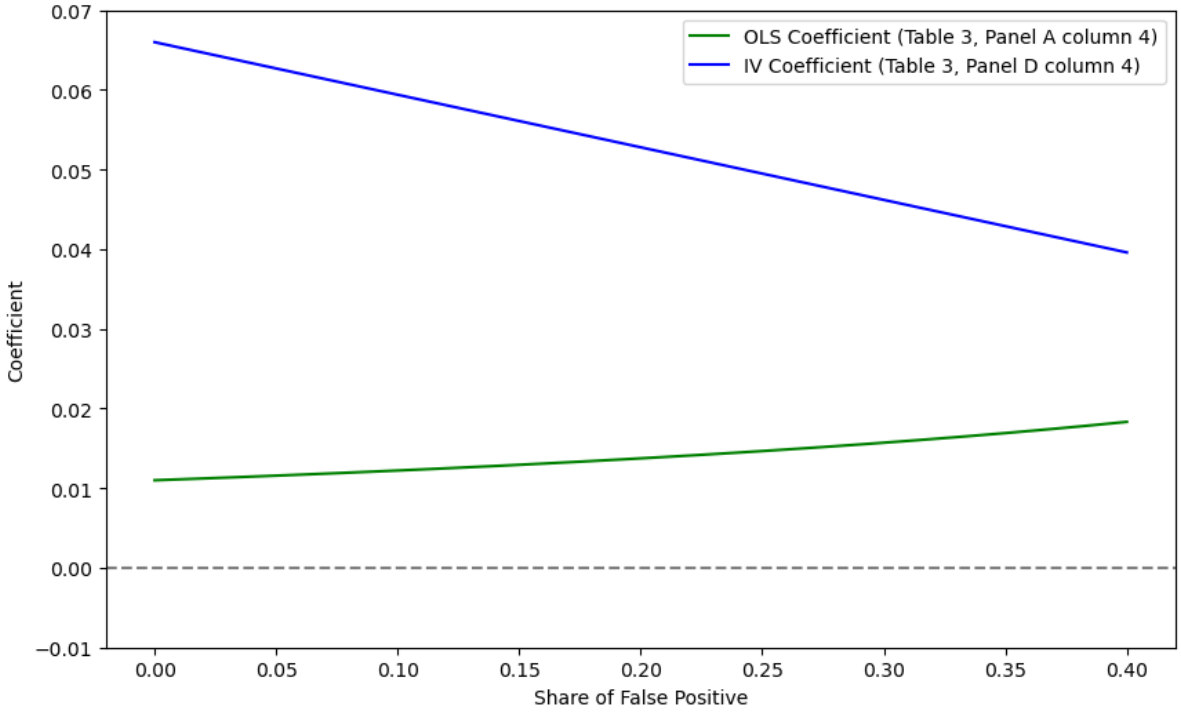
Notes: The figure shows the number of casualties over time (year-month) for the largest 100 regiments, panel a shows the total number of casualties, panel b shows the number of casualties on the battlefield. These regiments represent the 26.4% of all regiments and include the 54.8% of total casualties. The period of analysis is May 1915 - November 1918.

Figure A.8: Effect on tax avoidance - Different bandwidth definitions (income bins in *lire*)



Notes: The figure plots the coefficients of the regression according to equation 2, where the dependent variable is the measure of tax avoidance (binary variable indicating whether the taxpayer's income is just below the exemption thresholds, with a bandwidth of -10 lire). In the different specifications we modify the definition of the tax avoidance indicator progressively reducing the income window that defines it. The specification includes Municipality fixed effects and Surname fixed effects interacted with Province fixed effects as well as a control set including: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), dummy capturing whether the military victim was young (below the median age), number of taxpayers from the same surname-municipality, dummy indicating incomes rounded at hundred lire, differential for type of taxpayer.

Figure A.9: Misclassification checks



Notes: This figure reports the true coefficient estimate (y-axis) under different hypothetical misclassification levels (x-axis).

Table A.1: Taxpayers category, 1924

Category	Type of income	N. Tax payers	Sum income declared	Average Income
A	Capital	487,057	936,057,618	1,921
B	**Commercial and industrial**	**790,697**	**4,637,308,934**	**5,864**
C	**Profession of liberal arts**	**164,501**	**2,300,934,863**	**13,987**
D	Salaries and pensions of public employees	21,077	1,342,912,893	63,714
Total		1,379,871	9,217,214,308	6,679

Notes: This table reports information about the different categories of taxpayers that were subject to the income tax (Italian Ministry of Finance). The amounts are expressd in Lire.

Table A.2: Descriptive statistics

	Panel A: Full sample (OLS sample)			Panel B: Only with casualties (IV sample)		
	Average value	Standard deviation	N obs.	Average value	Standard deviation	N obs.
Audit-adjusted tax return	.119	.324	215,179	.136	.343	54,990
Death of a relative (dummy)	.310	.462	215,179	1	0	54,990
Death of a relative on the battlefield (dummy)	.252	.434	215,179	.808	.393	54,990
Death of a relative for other reasons (dummy)	.057	.232	215,179	.191	.393	54,990
Death of a relative (number)	.793	1.86	215,179	2.17	1.95	54,990
Death of a relative on the battlefield (number)	.527	1.30	215,179	1.49	1.53	54,990
Gross declared income	3,445.18	6912.90	214,931	3,184.65	5,859.85	54,928
Minimum taxable income	.0031	.056	215,179	.0032	.056	54,990
Num. of taxpayers from the same surname-municipality	4.00	6.36	215,179	6.56	7.68	54,990
Avg. military unit risk	.	.	.	.678	.118	54,990
Casualty of an army official	.	.	.	.169	.375	54,990
Casualty awarded a Medal of Honor - bronze	.	.	.	.045	.208	54,990
Casualty awarded a Medal of Honor - silver	.	.	.	.052	.222	54,990
Casualty awarded a Medal of Honor - gold	.	.	.	.0015	.0390	54,990
Casualty young (age below median)	.	.	.	.564	.404	54,833

Notes: the variables in "OLS sample" show the descriptive statistics for the sample of all taxpayers. The variable in "IV sample" show the descriptive statistics for the taxpayers with at least one casualty and whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. *Audit-adjusted tax return* is a dummy variable that captures the taxpayers that experienced a tax return adjustment. *Death of a relative (dummy)* indicates the families with at least a casualty, while *Death of a relative (number)* is the number of casualties for each family. *Death of a relative on the battlefield (dummy)* indicates the families with at least a casualty on the battlefield, while *Death of a relative on the battlefield (number)* is the number of casualties on the battlefield for each family. *Death of a relative for other reasons (dummy)* indicates families with at least one relative that died but not on the battlefield. *Gross declared income* is expressed in 1924 *Lire* and considers the corrected income for taxpayers reporting an audit-adjusted income. *Avg. military unit risk* captures the average risk of the military unit where the military victims of a certain family fought. *Minimum declared income* is a dummy variable that indicates declared incomes that are below to the minimum taxable amount (with a bandwidth of 10 lire). *Number of taxpayers from the same surname-municipality* is the number of taxpayers from a given municipality with the same surname of the taxpayer in consideration. *Casualty of an army official* indicates if the relative/s that died was an army official. *Casualty awarded a Medal of Honor- bronze/silver/gold* indicates if the relative/s that died received at least a bronze/silver/gold medal of honor, respectively. *Casualty young (age below median)* indicates whether the average age of the relative(s) who died is below the average age of death in the sample of war victims.

Table A.3: Descriptive statistics: taxpayers profession

	Average value	Standard deviation	N observations
High education	.0572	.2323	215,179
Liberal arts	.0591	.2358	215,179
Notary	.0056	.0746	215,179
Lawyer	.0149	.1211	215,179
Accountant	.0007	.0276	215,179
Physician	.0196	.1389	215,179
Veterinarian	.0017	.0415	215,179
Pharmacist	.0110	.1047	215,179
Chemist	.0000	.0052	215,179
Engineer	.0036	.0599	215,179
Architect	.0003	.0176	215,179
Surveyor	.0010	.0328	215,179
Agronomist	.0005	.0243	215,179
Actuary	.0000	.0037	215,179
Business owner: hotels	.0060	.0776	215,179
Business owner: restaurant	.0680	.2518	215,179
Business owner: retail	.1571	.3639	215,179
Business owner: firm	.0361	.1867	215,179
Business owner: laboratory	.0021	.0458	215,179
Craftsman	.0446	.2064	215,179
Landlord	.1516	.3587	215,179
Tenant	.1447	.3518	215,179
Rentier	.0062	.0787	215,179
Transport	.0081	.0898	215,179
Tailor/Fabric sales	.0309	.1733	215,179
Barber shop	.0068	.0825	215,179
Timber sales	.0166	.1280	215,179
Miller	.0082	.0905	215,179
Butcher	.0167	.1284	215,179
Bakery	.0051	.0717	215,179
Shepherd	.0474	.2126	215,179
Bricklayer	.0057	.0757	215,179

Notes: *High education* includes the professions that require a university degree. *Liberal arts* includes the following professions notary, lawyer, accountant, physician, veterinarian, pharmacist, chemist, engineer, architect, surveyor, agronomist and actuary. We have been able to categorized the professions for the the 68.4% of taxpayers.

Table A.4: Structure of Italian army during WWI (*Albo dei caduti*)

Military specialization	Number of units (regiments)	Number of military victims
Alpine troopers	8	31,370
Artillery - farmland	50	16,320
Artillery - fortress	10	10,106
Artillery - mountain	3	4,759
Bersaglieri	21	28,263
Cavalry	30	3,756
Infantry	238	340,940
Engineers corps	8	15,765
Granadiers	2	6,605
Total	370	457,884

Notes: Alpine troopers, *Alpini* in Italian, were troops specialized in mountain operations. The specialization of the *Bersaglieri* is a type of infantry whose soldiers were in origin trained with rifle shotguns.

Table A.5: Balance checks: Taxpayers Characteristics

Variable	Coefficient	Standard error	Variable	Coefficient	Standard error
Income (log)	0.0066	(0.0085)	Business owner: retail	0.0009	(0.0033)
High education	0.0003	(0.0018)	Business owner: firm	-0.0022	(0.0020)
Liberal arts	0.0001	(0.0018)	Business owner: laboratory	-0.0001	(0.0002)
Notary	0.0002	(0.0007)	Craftsman	-0.0007	(0.0017)
Lawyer	-0.0002	(0.0008)	Landlord	0.0042	(0.0039)
Accountant	-0.0002	(0.0002)	Tenant	-0.0028	(0.0039)
Physician	-0.0003	(0.0010)	Rentier	-0.0004	(0.0006)
Veterinarian	0.0006	(0.0004)	Transport	0.0008	(0.0006)
Pharmacist	-0.0000	(0.0006)	Tailor/Fabric sales	0.0006	(0.0013)
Engineer	0.0001	(0.0004)	Barber shop	-0.0003	(0.0006)
Architect	0.0000	(0.0000)	Timber sales	-0.0004	(0.0011)
Surveyor	0.0001	(0.0002)	Miller	0.0002	(0.0009)
Agronomist	-0.0001	(0.0002)	Butcher	-0.0008	(0.0016)
Actuary	-0.0000	(0.0000)	Bakery	0.0005	(0.0005)
Business owner: hotels	0.0005	(0.0005)	Shepherd	0.0015	(0.0023)
Business owner: restaurant	0.0029	(0.0022)	Bricklayer	-0.0012	(0.0008)

Notes: All the dependent variables except *Income (log)* are dummy variables indicating the profession reported in the headline. The specification of the analysis is as in Table 2 (column 4), and it includes Municipality fixed effects, Surname fixed effects interacted with Province fixed effects and baseline controls which include: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), dummy capturing whether the military victim was young (below the median age), number of taxpayers from the same surname-municipality, terciles of income FE (not in the analyis on *Income (log)*). The sample includes taxpayers whose families suffered a loss due WWI and whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. Standard errors clustered at the municipality level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

Table A.6: Balance checks: Taxpayers Characteristics - Family aggregation

Variable	Coefficient	Standard error	Variable	Coefficient	Standard error
Income (log)	0.0068	(0.0109)	Business owner: retail	0.0049	(0.0039)
High education	0.0001	(0.0023)	Business owner: firm	-0.0006	(0.0020)
Liberal arts	0.0001	(0.0023)	Business owner: laboratory	-0.0002	(0.0003)
Notary	0.0001	(0.0011)	Craftsman	-0.0009	(0.0023)
Lawyer	-0.0009	(0.0010)	Landlord	-0.0005	(0.0039)
Accountant	-0.0002	(0.0002)	Tenant	-0.0036	(0.0040)
Physician	-0.0002	(0.0012)	Rentier	-0.0001	(0.0006)
Veterinarian	0.0006	(0.0004)	Transport	0.0010	(0.0007)
Pharmacist	0.0003	(0.0009)	Tailor/Fabric sales	0.0012	(0.0016)
Engineer	0.0002	(0.0004)	Barber shop	-0.0001	(0.0007)
Architect	0.0000	(0.0000)	Timber sales	0.0011	(0.0013)
Surveyor	0.0001	(0.0002)	Miller	-0.0003	(0.0012)
Agronomist	-0.0003	(0.0003)	Butcher	-0.0011	(0.0019)
Actuary	-		Bakery	0.0007	(0.0006)
Business owner: hotels	0.0004	(0.0007)	Shepherd	0.0009	(0.0029)
Business owner: restaurant	0.0036	(0.0026)	Bricklayer	-0.0016	(0.0013)

Notes: All the dependent variables except *Income (log)* are variables indicating the share of people within a family practicing the profession reported in the headline. The specification of the analysis is as in Table 2 (column 4), and it includes Municipality fixed effects, Surname fixed effects interacted with Province fixed effects and baseline controls which include: dummy variable indicating that at least one military victim (within the family) was an army official, dummy variable indicating that at least one military victim (within the family) was awarded a Medal of Honor (three dummies capturing either a bronze, silver or gold medal), share of young military victims (below the median age), number of taxpayers from the same surname-municipality, share of taxpayers in a certain family in each income tercile (not in the analyis on *Income (log)*). The sample includes the families suffering a loss due WWI and whose relative(s) serve only in artillery, infantry or in one of the corps of *Bersaglieri*. Standard errors clustered at the municipality level in parenthesis. * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$.

B. Appendix - Misclassification

B.1. Misclassification in OLS

From Aigner (1973), we know that the OLS estimates bias, in the presence of misclassification, depends on the level of the false positive rate (FPR) or the probability of being misclassified as treated when not actually treated and the false negative rate (FNR) or the probability of being misclassified as not treated when actually treated.

In our context, we discuss the misclassification problem considering the general case of having three potential treatment statuses (see the last three Columns of Table 1). Consider the outcome variable Y_i, for example a measure of tax evasion, modeled as:

$$Y_i = \beta_0 + \beta_1 T_{1i} + \beta_2 T_{2i} + \epsilon_i$$

where, T_{1i} is the observed indicator for individual i being in T_1 (death of a relative in the battlefield), T_{2i} is the observed indicator for individual i being in T_2 (death of a relative for other reasons), T_3 (no deaths of a relative) is the reference category (not included in the regression), β_0 is the intercept, and ϵ_i is the error term. The observed treatment status can be expressed in terms of the true status and the misclassification probabilities:

$$T_{1i} = T^*_{1i}(1 - \pi_{13}) + T^*_{3i}\pi_{31}$$

$$T_{2i} = T^*_{2i}(1 - \pi_{23}) + T^*_{3i}\pi_{32}$$

Where π_{13} is the probability that an individual who truly belongs to T_1 (battlefield deaths) is misclassified as T_3 (no deaths), π_{23} is the probability that an individual who truly belongs to T_2 (non-battlefield deaths) is misclassified as T_3 (no deaths), π_{31} is the probability that an individual who truly belongs to T_3 (no deaths) is misclassified as T_1 (battlefield deaths), and π_{32} is the probability that an individual who truly belongs to T_3 (no deaths) is misclassified as T_2 (non-battlefield deaths). We exclude the possibility of misclassifying T_1 with T_2 and vice-versa, as they are not suffering from any linkage problem as it could be the case when comparing T_3 with both T_1 and T_2. Substituting the observed treatment variables into the regression model, we get:

$$Y_i = \beta_0 + \beta_1[T^*_{1i}(1 - \pi_{13}) + T^*_{3i}\pi_{31}] + \beta_2[T^*_{2i}(1 - \pi_{23}) + T^*_{3i}\pi_{32}] + \epsilon_i$$

Rearranging and grouping the terms, we have:

$$Y_i = \beta_0 + T^*_{1i}\beta_1(1-\pi_{13}) + T^*_{2i}\beta_2(1-\pi_{23}) + T^*_{3i}(\beta_1\pi_{31} + \beta_2\pi_{32}) + \epsilon_i$$

The coefficients of the true treatment indicators in which we are interested are the ones related to T^*_{1i} and T^*_{2i}.[27] Isolating and rearranging the terms of interest, we find that the estimated $\hat{\beta}_1$ and $\hat{\beta}_2$ are a biased estimate of the true βs.

$$\hat{\beta}_1 = \beta_1(1-\pi_{13}) \qquad \beta_1 = \frac{\hat{\beta}_1}{1-\pi_{13}}$$

$$\hat{\beta}_2 = \beta_2(1-\pi_{23}) \qquad \beta_2 = \frac{\hat{\beta}_2}{1-\pi_{23}}$$

Specifically, they are attenuated proportionally to π_{13} and π_{23}, representing the respective false positive rates. Importantly, the estimated bias in $\hat{\beta}_1$ is equivalent to the bias encountered when running an OLS regression including only observations classified as having a relative who died during the war, with the treatment being the cause of death (see the Panel A: OLS of Table 2).

Instead, the presence of bias for the estimate of a treatment that considers any cause of death vs. non-death of a relative (see the first three Columns of Table 1) can be addressed by expanding on the previous results:

$$\hat{\beta} = \beta(1-\pi_{13}-\pi_{23}-\pi_{31}-\pi_{32}) \qquad \beta = \frac{\hat{\beta}}{(1-\pi_{13}-\pi_{23}-\pi_{31}-\pi_{32})}$$

In this case, the estimate signs may change if the sum of the misclassification probabilities exceeds one. In a scenario where all individuals are assigned to the wrong group, the estimates of β will be flipped.

B.2. Misclassification in IV

From Meyer and Mittag (2017), Bingley and Martinello (2017) and DiTraglia and Garcia-Jimeno (2019), we know that misclassification in the treatment status of the IV is upward bias and is the inverse of the OLS bias. The intuition is that while the misclassification does not affect the reduced form, $\text{Cov}(Y_i, Z_i)$, the first stage, $\text{Cov}(T_{1i}, Z_i)$, is affected with a downward bias.

[27] The misclassification in T^*_{3i} is affecting the intercept.

Due to misclassification, $\text{Cov}(T_{1i}, Z_i)$ is attenuated by the factor $(1 - \pi_{13})$, leading to an upward bias in the IV estimate of β_1. Specifically:

$$\hat{\beta}_1^{IV} = \frac{\text{Cov}(Y_i, Z_i)}{\text{Cov}(T_{1i}, Z_i)}$$

$$\hat{\beta}_1^{IV} = \frac{\beta_1^{IV}}{1 - \pi_{13}} \qquad \beta_1^{IV} = \hat{\beta}_1^{IV}(1 - \pi_{13})$$

Therefore, our estimates for the IV in the constrained sample, where only taxpayers identified as having a relative who died during the war are upward bias.