Supplementary Appendix for "Scars of War: the Legacy of WWI Deaths on Civic Capital and Combat Motivation"

Felipe Carozzi, Edward Pinchbeck, Luca Repetto

Appendix A. Data

Appendix A.1. Data Sources

Data on British service personnel killed in both wars is obtained from the Commonwealth War Graves Commission (CWGC) (Commonwealth War Graves Commission, 2021), an intergovernmental organisation dedicated to marking, recording and maintaining the graves, memorials and memories of the men and women of the Commonwealth forces who died in both World Wars. Open data from this organisation contains information on names, time of death, rank, fighting unit, honours (e.g., gallantry medals), age and a string from which we can extract the location of origin of dead soldiers. Data on locations is augmented using information from Forces War Records (FWR), a military genealogy specialist website (Forces War Records, 2021).

Data on 4,135,092 war records of soldiers mobilised during WWI is obtained from FamilySearch, a not-for-profit organisation which offers on-line access to large genealogical datasets (FamilySearch, 2021). FamilySearch draws its information from the British Army Service Records for 1914 to 1920. These records contain information on enrolled soldiers including names, place of residence, birthplace, age at the time of enlistment, year and unit in which the soldier was enlisted.[39] When cleaning and processing this information, we use as reference the Table of Organisation of each regiment as detailed in James (2012).

Individual-level information on the English and Welsh population before the Great War is obtained from the 1911 Census of population. The data we use originates from Schürer and Higgs (2014) and is distributed by IPUMS (Minnesota Population Center, 2019). We use this data both at the individual level in Section 5.3 and to construct aggregates at the parish level. From this source, we obtain information on the occupational composition of the workforce and several income proxies including the number of servants and the number of rooms per household. We obtain aggregate area-level information for Census years 1901-1931, as well as digital maps for parishes, districts and constituencies from "A Vision of Britain through Time" (VoB), an online library of spatial data created by the Geography Department at the University of Portsmouth (University of Portsmouth, 2011). The parish-level population counts in the VoB data come from the *Census Reports* that were published following each Census. There are

[39] Digitised versions of these records can be consulted at `www.ancestry.co.uk`. The FamilySearch collection, which includes the extracted data, is called "United Kingdom, World War I Service Records, 1914-1920". The original sources of this information are the "Burnt documents" (record code WO 363) and the "Unburnt collection" (record code WO 364), which are kept in the National Archives at Kew in London. The Burnt Documents are roughly 2.5 million records on WWI soldiers which survived the fire resulting from an incendiary bomb hitting the War Office Record Store in 1940. The Unburnt Collection is made of soldier information obtained from pension claims. This collection was stored separately in 1942 and, therefore, did not suffer the fate of many of the Burnt Documents.

known to be discrepancies between the population counts in this source and the more recently published micro-data, for example because not all records have survived or there is ambiguity in the true parish in the individual level records. Consequently, in general we use the counts from the *Census Reports* where available. Further, to minimise discrepancies we also implement the corrections to assigned parishes in the 1911 micro-data using the look-up tables published on the I-CeM website.[40]

We use data from a number of sources to obtain spatially disaggregated proxies for civic capital. Data on war memorials built both before and after the Great War are obtained from the Imperial War Museum memorial registry (Imperial War Museum, 2021*b*), and complemented with information on Listed memorials from Historic England and the Welsh equivalent, Cadw (Cadw, 2023; Historic England, 2023). Information on registered charities and their location is from the Charity Commission for England and Wales (Charity Commission for England and Wales, 2023). Data on mutual societies – a type of enterprise that can be likened to a cooperative – are obtained from the Financial Conduct Authority (Financial Conduct Authority, 2023). Data on branches of the British Legion – a veterans' association set up after WWI – are contained in the Charity Commission data but are incomplete in terms of addresses so we complete these using the Royal British Legion website and internet directories, including the website `www.192.com`, and link these addresses back to parishes using official postcode directories (Office for National Statistics, 2022). Our election data are from the Constituency-Level Elections Archive (Kollman et al., 2019). Data on financial contributions to voluntary hospital schemes come from the Voluntary Hospital Database (see `http://www.hospitalsdatabase.lshtm.ac.uk/the-voluntary-hospitals-database-project.php`). We transcribe information on volunteers for ARP (Air Raid Precaution) duties as at June 1939 that are held in file HO 186/1 of the National Archives in Kew.

Our work also relies on a number of other ancillary sources. Because information from CWGC on medal recipients is limited to soldiers who lost their lives, we also collect and geolocate data on all military recipients of the Victoria Cross (with year of award). These are 1,353, of which we are able to geolocate 666 to England and Wales' parishes. Data on recipients of the George Cross (only those awarded to soldiers and during WW2) are from Wikipedia. These are 79, of which we can geolocate 51 to England and Wales. Complete records on DSO and DCM recipients are not publicly available. By combining data from website www.tracesofwar.com, Forces War Records (collection *UK, WWII, Recipients of the DCM 1939-1945*), the British Newspaper Archive (range Dec. 1938 – Dec. 1947), and the London Gazette we successfully geolocate 926 DCM and 1,127 DSO recipients.

We use the Imperial War Museum's Lives of the First World War database (Imperial War Museum, 2021*a*) to create lists of soldier surnames and to construct counts of WWI conscientious objectors by parish. A Parliamentary return provides counts of mobilised soldiers eligible to vote by constituency in 1945 (H.M Stationary Office, 1945). We compute Pythagorean dis-

[40] Available at https://www.essex.ac.uk/-/media/newparids11.txt?la=en, accessed on May 5, 2023.

tance to the nearest WWI barracks using parish centroids and coordinates of barracks from Historic England (Historic England, 2021). We do likewise for distance to Regimental Headquarters after geolocating them manually. We obtain 1918-1919 influenza deaths by district from the Supplement to the Eighty-First Annual Report of the Registrar-General (Johnson, 2001). Finally, we create a list of Pals Battalions using information in James (2012) and Becke (1938).

Appendix A.2. Spatial Units of Analysis and Reconciliation

Our main analysis is based on a 1911 parish-level dataset covering England and Wales. We take 1911 as our reference year because it was the last Census conducted before the onset of the Great War in 1914. The civil parishes we use in our analysis are administrative units corresponding to the lowest level of local government in the United Kingdom. Civil parishes evolved from ecclesiastical parishes during the 19^{th} century, but by 1880 had no religious or ecclesiastical duties. In 1911, the territories of England and Wales were divided into 14,664 parishes, of which 13,404 in England and 1,260 in Wales. We drop all parishes that had zero population in 1911 – usually parcels of empty land in remote rural areas – and 10 additional parishes that have repeated names within the same county.

Parishes are nested within local government districts, of which there were 1,861 in 1911, and in turn within 52 counties. In some specifications we use data for 505 constituencies, which are electoral units that are distinct from the aforementioned local government areas. Parish boundaries change over time and in some cases variables are only available at other (higher) levels of aggregation. In order to aggregate or re-weight information to common boundaries we use a spatial matching procedure based on the assumption of uniform population distribution within parishes. Because our main spatial units (parishes) are relatively small (10 sq. km on average) and parish boundaries are often quite stable in the 30-year period we study, we expect the measurement error induced by making this assumption to be limited. We selectively group parishes using three criteria. First, using a semi-automated approach we group together around 2,000 parishes that have similar names and lie within the same district. For example, we group High Abbotside with Low Abbotside and we also group Ledbury Urban with Ledbury Rural. Second, we group a similar number of very small parishes with other proximate parishes where they form a grouped parish council or when they together comprise a 1911 Census Enumeration District (ED).[41] Third, we group a small number of suburban parishes with the corresponding city parish. For example, we group the suburb of Sculcoates with Hull. Reflecting that soldiers from the London area often report their location as being London, we also group together London parishes based on the historical conurbation definition available from Vision of Britain. Finally, we exclude 26 parishes, of which 25 are parcels of empty land and the other is unnamed.

[41] Following the Local Government Act 1894, parishes with populations over 300 were required to form councils. Smaller parishes could elect to form grouped parish councils with neighbouring parishes and we use these functional groupings where they exist. We group other small parishes with the largest parish that lies in the same 1911 ED on the basis that EDs were formed such that a Census enumerator could cover the area in a single day on foot.

After applying these restrictions and grouping parishes, our final dataset encompasses 10,807 parishes, of which 9,796 are in England and 1,011 are in Wales.

Our data on 1911 parishes come from two different sources: the 1911 Census micro-data from I-CeM and the *Census Reports* from VoB. These sources use different parish codes and contain a slightly different set of parishes, so we create a mapping file and reconcile the data before conducting analysis.

Appendix A.3. Geolocation Procedure, Measurement Error, and Validation

Our empirical analysis requires adequately geolocating soldiers based on information on their place of birth and residence from the sources described above. Here we provide details of the geolocation procedures used to assign soldiers to their parish of origin. We also produce a series of figures that serve as validation for the resulting parish-level aggregates in WWI mobilisation and soldiers killed in both wars.

The CWGC data on soldiers killed during WWI includes 796,601 records.[42] Given that our analysis will focus on England and Wales only, we remove servicemen born in Scotland, Ireland, and abroad. We then extract information for residence or birthplace (or both) from either the birthplace and residence fields in FWR or the "additional information" string included in the CWGC source.

The CWGC dataset on soldiers killed in WW2 has information on 435,696 deaths (of which 67,591 were civilians) during 1939-1945. For about 344,000 of them (79% of total), some additional information is provided in the form of a short text that very often includes the location of origin.

Geolocation of WWI dead soldiers proceeds by combining a) direct string matches with parish names based on data from FWR on historic county and location of birthplace/residence, b) direct string matching as above but based on the CWGC additional information field, and c) latitudes and longitudes obtained from a batch geolocating service to which we input the FWR locations. For the batch geolocation process, we use a service provided by the company OpenCageGeo, which is based on OpenStreetMap and is available across platforms. In order to validate the geolocation process used by this source, we randomly selected 800 individual servicemen and validated the imputed locations by hand. Only 9 observations in this sample were incorrectly imputed and 6 of these 9 were imputed to nearby areas. Hence, we conclude that the geolocation process based on this method is sufficiently reliable for our purposes, resulting in a limited amount of measurement error.

Geolocation of WW2 soldiers is slightly different because FWR information is of much lower quality, and proceeds as follows: a) direct string matching with parish and historic county names based on the CWGC additional information field, combined with b) extraction of location information from the CWGC additional information field, geolocated using OpenCageGeo, and

[42] This number is in line with the 702,410 born in the British Isles and killed in the war, as reported by the British government (BWO, 1922) because the CWGC data also includes men from British dominions and Commonwealth countries.

c) direct string matching based on data from FWR on historic county and location of residence and birthplace.

The data on parish of origin (birthplace or residence) of mobilised men in WWI – obtained from FamilySearch – has a slightly different structure and, therefore, we use a different procedure from the one used for CWGC/FWR data.[43] To match the FamilySearch records to an individual parish we combine: a) a direct string match with parish names for records that have both an historic county and a location, b) direct string matching with parish names for records that include no county information (only match to parishes with unique names), c) hand matching of a fraction of remaining records carried out by identifying locations via GoogleMaps. We are able to geolocate just over 2.6 million of these records.

When using this data together with the CWGC information on deaths to construct our instrument, we need to restrict to servicemen that could be geolocated and had information on the battalion. Specifically, we start with our initial sample of 4.1 million war records and exclude 2.05 million records for which the battalion is missing. Then, we drop 57,795 entries that are duplicates in terms of all variables, and 72,131 records dated before 1905 or after 1920. Of the remaining sample, we then drop 735,768 individuals that could not be geolocated, as well as 28,578 from regiments with zero or negligible mortality, such as the Hussars. Finally, to ensure we have enough observations to construct the shares serving in each battalion, we drop 38,281 soldiers from battalions with less than 100 servicemen in the data.

Because of the measurement error deriving from the geolocation and the incompleteness of the FamilySearch records, some parishes exhibit values of mobilisation or WWI and WW2 deaths that are unusually large relative to their population. To ensure that these possible outliers are not driving the results, we identify all parishes in which the number of mobilised, WWI deaths and WW2 deaths have per-capita values above the 99^{th} value of the respective distribution. We then replace those figures with the imputed number of dead and of mobilised obtained by multiplying the 1911 parish population by the district-level death or mobilisation rates, as appropriate.[44] In Appendix Appendix C we show that results are robust to not applying this correction.

As an additional step to validate the WWI mobilisation figures derived from FamilySearch – and the associated geolocation process – we first investigate the relationship between mobilisation and 1911 population figures from the Census. The associated binned scatter plot of both variables in log scale is depicted in panel A of Figure A.6. We can observe a clear positive relationship, which is what we would expect given the nature of the mobilisation process. The associated univariate regression yields a fairly high R-squared of 0.63 and a slope coefficient of 0.99.

We can jointly validate the parish-level mobilisation and deaths figures by looking at the

[43] For example, the batch geocoding procedure that we used and validated when using FWR data on locations for killed soldiers yields very poor results when used with the FamilySearch strings.

[44] We use the historical county when district and grouped parish boundary coincide. Note that the outlier adjustment is not applied to the mobilisation totals that enter the calculation of the instrument z_i.

relationship between mobilisation rates and death rates (i.e., the relationship between both mobilisation and deaths divided by population). The associated binned scatter plot (in log scale) is provided in Panel B of Figure A.6. Again we find a positive and almost linear relationship, in line with expectations. The associated univariate regression yields an R-squared of 0.18 and a slope coefficient of 0.37, indicating that there was a clear relationship between mobilisation and deaths – as expected – but that there was substantial unexplained variation in deaths after accounting for differences in mobilisation and population.

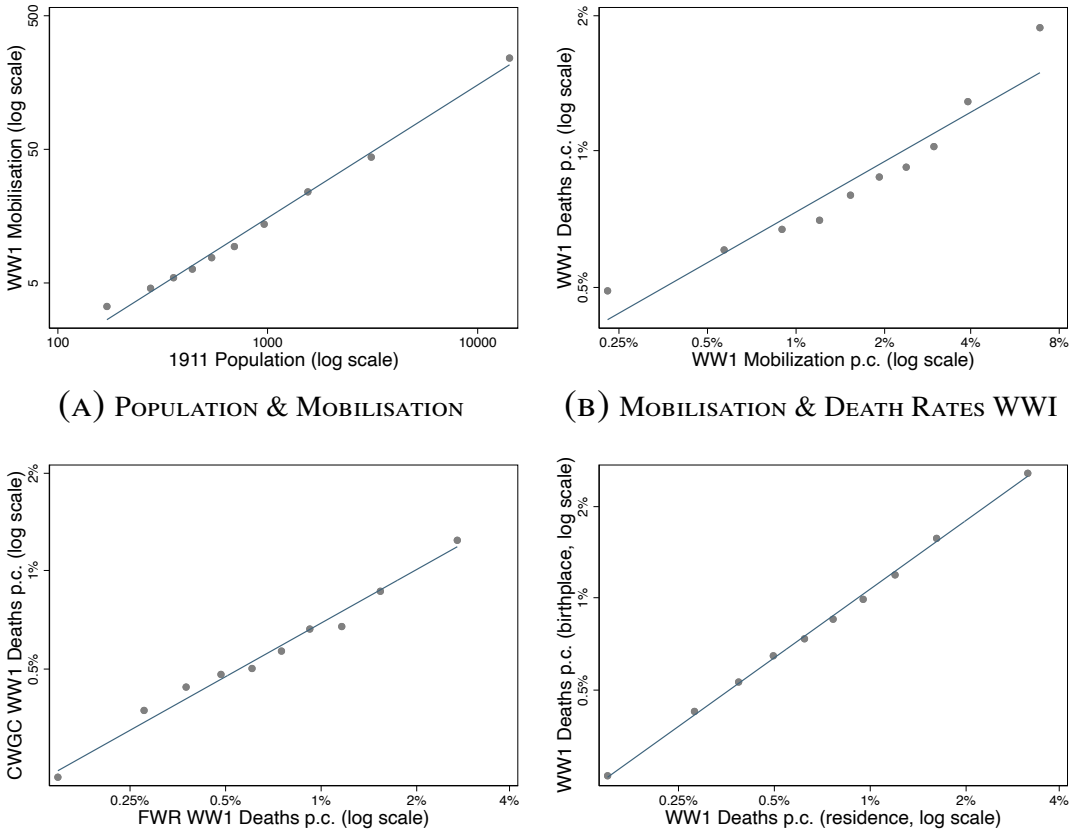
To validate the geolocation procedure for deaths we can use figures for death rates at the parish level constructed using two different sources of the underlying location of origin data: the "additional information" string in the CWGC data and data on birthplace or residence in the FWR records. The corresponding binned scatter plot is shown in Panel C of Figure A.6. There is clearly a positive and close to linear relationship between death rates from both sources. The associated univariate regression yields an R-squared of 0.22 and a slope coefficient of 0.54.

Finally, we can also compare death rates calculated from data on either birthplace or residence from FWR. We expect that both sources would yield very similar figures for deaths because most people reside in the same parish in which they were born. The associated binned scatter plot is provided in Panel D of Figure A.6 and shows a clearly positive and linear relationship. The univariate regression yields an R-squared of 0.53 and a slope coefficient of 0.74.

We can further validate the baseline measure of WWI deaths used in the paper with one constructed using the number of dead commemorated in local memorials. As discussed in Section 2, memorials often include a list of names of the local servicemen who lost their lives in the war. We aggregate these figures at the parish level and investigate the correlation between the parish-level death rate thus constructed and the death rate constructed using our main measure. Results of these comparisons for both WWI and WW2 are illustrated in panels A and B of Figure A.7. The depicted relationships are positive and close to linear. The associated univariate regressions yield elasticities of over 0.2, significant at all conventional levels.

Figure A.6

Validation: Mobilisation and Death Rates



(A) Population & Mobilisation (B) Mobilisation & Death Rates WWI

(C) Death Rates WWI – FWR & CWGC (D) Death Rates WWI – Residence & Birth

Notes: Binned scatter plots of: **Panel A** – parish-level WWI mobilisation and 1911 population, both in logs; **Panel B** – log death rates for WWI calculated from the CWGC source in the horizontal axis and from log mobilisation rates; **Panel C** – log death rates calculated using FWR and CWGC information. **Panel D** – death rates at the parish level calculated using FWR information based on birthplace data and residence data. Fitted line corresponds to OLS estimates using the underlying data.

Appendix A.4. Details of Civic Capital Measures

We use data from a number of sources to create measures of civic capital in the pre-WWI and inter-war periods. Our approach to measure civic capital is to examine results across several distinct measures to provide a comprehensive picture and assess robustness.

Appendix A.4.1. War Memorials

Our first measure of civic capital is the presence of one or more listed war memorials in a parish. We focus on listed, rather than all war memorials, as listed status indicates memorials have historical or architectural significance. The data on war memorials chiefly comes from Historic England and Cadw - the public bodies responsible for caring for and promoting historic and heritage assets in England and Wales, respectively. We classify listed memorials as memorialising different conflicts (e.g., Boer War; WWI; WW2) using additional information contained within the IWM Memorial Register. As we wish to examine civic capital in the inter-war period, we define listed WWI memorials as those that commemorate WWI and were built before the start of WW2. For balancing checks we define pre-WWI memorials as those built before 1914, and Boer War memorials as those that commemorate those conflicts. Since we have location details (postcode or grid reference) for the vast majority of memorials, we are able to assign counts of memorials to 1911 parishes with a high degree of accuracy.

One potential concern about using listed memorials is that the listed building regime only began in earnest following WW2. A large number of listed WWI memorials were not listed until the centenary of WWI in 2014-2018 under a project by Historic England that aimed to add 2,500 memorials to the list. As listing can occur because a structure is "at risk",[45] a concern is that the memorials that were added after 2014 are not listed because of the effort communities made to honour WWI soldiers, but because they were subsequently neglected or else happen to be located in places that were being considered for renewal or redevelopment in the 2010s. We therefore exclude these memorials in some specifications.

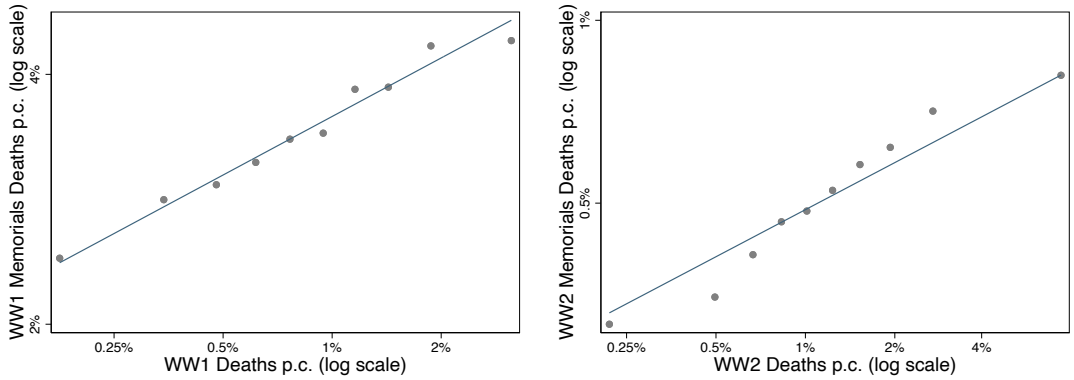
Appendix A.4.2. Charities, Mutuals, and British Legion Branches

We use two further measures of civic capital that are based on the formation of new charities, mutuals, and British Legion branches in the inter-war period. The main data sources for constructing these measures are the Charity Commission's Register of Charities, and the Financial Conduct Authority's Mutuals Public Register. For charities, we first extract the first year recorded in the governing document description data field. We obtain a year for more than 90% of charities in this way. We then restrict attention to the approximately 48,500 entries present before 1939. We geolocate around 12,700 of these from postcodes in the data, and a further 12,000 from string matching the location given in the area of benefit field to a unique parish name in our dataset, before dropping roughly 4,200 relating to the formation of Scouts and

[45] See, for example, https://www.warmemorials.org/listing-england/, last accessed on December 5, 2025.

Figure A.7

Validation: Death rates WWI and WW2



(a) WWI Dead Rate: Memorials & CWGC (b) WW2 Dead Rate: Memorials & CWGC

Notes: **Panel A**: binned scatter plot of the relationship between WWI death rates from memorials (vertical axis) and from the CWGC data (horizontal axis). **Panel B**: binned scatter plot of the relationship between WW2 death rates from memorials (vertical axis) and from the CWGC data (horizontal axis). Fitted line corresponds to OLS estimates using the underlying data.

Guides groups. For mutual societies, the raw data contains registration year. We begin with some 7,600 pre-1939 mutuals of which we are able to geolocate around two thirds. We construct counts of British Legion branches from the charities data. There are 2,074 branches listed in the data. We supplement the geolocation approach used for charities with manual searches of the Royal British Legion website and internet directories to obtain addresses/postcodes and hence parishes. By doing so we assign a 1911 parish to close to two thirds of the branches.

Appendix A.4.3. Additional Measures of Civic Capital

We use three additional measures of civic capital in our analysis. The first corresponds to contributions to voluntary hospital schemes. Voluntary hospitals were independent hospitals that, prior to the advent of the National Health Service (NHS) in 1948, provided general and acute services at low or reduced costs. They largely relied on donations and philanthropic funding and were typically staffed by unpaid doctors. The source of our data is the Voluntary Hospital database. We use information on all voluntary hospitals in the database in the period 1906-1939. The information changes over time as in some years a "voluntary gifts" total is recorded whereas for other years more disaggregated categories are reported instead. We compute a consistent measure of voluntary gifts by summing income from donations, collection boxes, and church collections. We link hospitals to parishes using the spatial coordinates recorded in the original dataset.

A second alternative measure of civic capital used in the analysis corresponds to counts of volunteers to serve in the Air Raid Precautions (ARP) civil defence service on June 1939. The ARP was the largest civil defence service in WW2 and was set up by the Air Raid Precautions Act that came into effect on 1 January 1938. The ARP was largely run by volunteers and various local and national recruitment initiatives ran in 1938 and then throughout the war to encourage participation. By September 1939, some 1.6 million men and women had volunteered. Our

source of data on ARP volunteers is official statistics held in the National Archives that tabulate returns from local authorities in June 1939 and record male and female volunteer counts for 83 large local authorities (County Boroughs) and 62 County remainders (which are aggregations of smaller districts).

The final measure of civic capital that we use is electoral turnout in national elections in the period between December 1910 and November 1935. The source of our data is the Constituency-Level Elections Archive voting data for England and Wales (Kollman et al., 2019). As constituency names and boundaries change throughout the period, we clean constituency names and use spatial re-weighting to harmonise the data to common spatial units. We compute voting turnout as the number of valid votes divided by the number of eligible voters in each constituency-election combination.

Appendix A.5. Linking 1911 Census to Military Records

As described in the main text of the paper, we exploit that we can access the full 1911 Census including names and addresses and unique individual and household identifiers to estimate how WWI deaths within households affect the behaviour of men in WW2. We take all male children in the 1911 Census aged 0-8 (so aged 28 to 36 at the start of WW2), then link these children to WW2 deaths. We separately link WWI deaths to all the men in the 1911 Census that could have fought in WWI. We then combine this second merge with the children dataset to identify which children had fathers and other household members that died in WWI.

We follow the following steps: First, we correct some minor 1911 Census parish errors using a file issued by IPUMS in December 2020. We create two files from the 1911 Census that will be matched to the war dead. The first file, which will be linked to WWI dead, comprises men aged between 10 and 50 in 1911 (and hence between 17 and 57 by the end of WWI). These are potential fathers and cohabiting household members of children in 1911. The second file, which will be linked to WW2 dead, is a file of male children aged between 0 and 8 in 1911 including the forenames of the boys, the forenames of their cohabiting father, and a household identifier.

We then prepare the war dead data for both WWI and WW2 for the ABE merge. There are 796,601 WWI dead in our data, of which some 383,000 are potentially matchable as age, forename, and surname fields are non-missing. There are 435,696 WW2 dead in our data. We only attempt to match around 85,000 that have birth years listed in the range of 1898 to 1912 (we include 1912 to allow for possible mistakes in the records).

We next run merges using the ABE algorithm. For matching WWI soldiers to 1911 Census men we use four matching strategies: (i) surname, forename, birth year and birthplace; (ii) surname, forename, birthyear and parish of residence; (iii) surname, forename, birthplace and birth year; (iv) surname, forename, birth year and district of residence. For matching WW2 soldiers to 1911 Census children we also use four matching strategies: (i) surname, forename, birth year, county of residence, and father's forename initial; (ii) surname, forename, birth year, and county of residence; (iii) surname, forename, birth year, district of residence, and father's forename initial; (iv) surname, forename, birth year, and district of residence.

In each case we use the default ABE parameters, NYSIIS standardised names, and allow the option to use standard nicknames. Note that the ABE matching procedure only considers records to be matched when matches are unique.

In the final step we combine the 1911 Census with the outputs of the ABE merges. We first take all boys aged 0 to 8 in the 1911 Census and we use the ABE WW2 merge to create an indicator variable for those who died in WW2 (we code non-matched children as 0). This provides our dependent variable. We then use the ABE WWI merge to create an indicator for children whose father died in WWI (we code non-matched fathers as 0). Finally, we link in the ABE WWI merge into our dataset for a second time but now merging on the household identifier rather than the person identifier. By doing so we can then create an indicator for whether a household member other than the father died in WWI (we code non-matched households as 0).

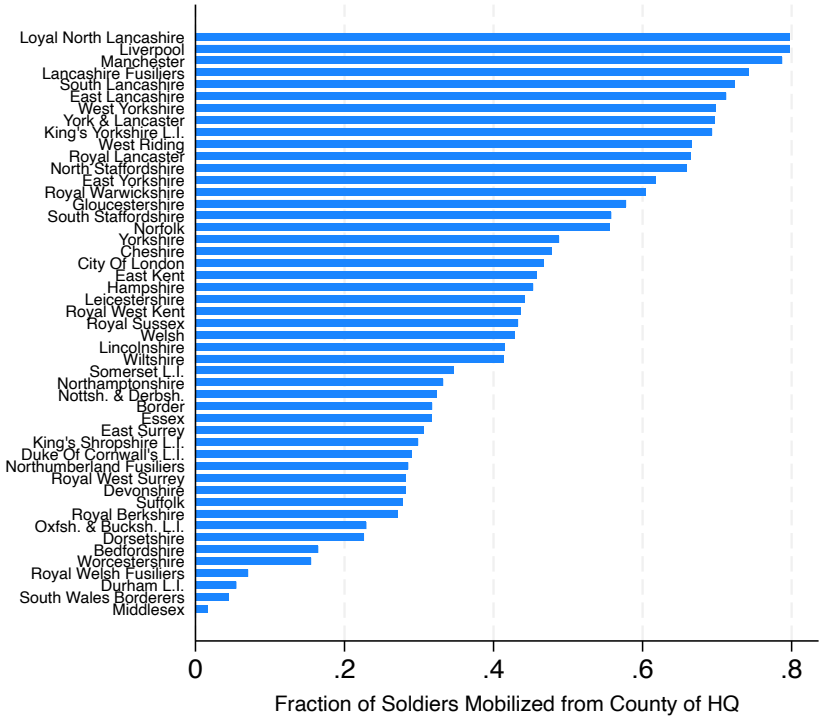
Appendix B. Additional Figures and Tables

FIGURE B.1
BRITISH ARMY REGIMENTS' RECRUITMENT AREAS IN WWI

Notes: Edited extract of a poster originally published by the Parliamentary Recruitment Committee, London, in 1915. Image from the Imperial War Museum archive. © IWM Art.IWM PST 11946. Enlarged section introduced by the authors. Note that not all regiments had a specific recruitment area. Some regiments such as the Royal Field Artillery, Royal Garrison Artillery or the Royal Rifle Corps recruited from all over the United Kingdom.

FIGURE B.2

WWI REGIMENTS AND LOCALISED RECRUITING



Notes: Horizontal axis represents the fraction of soldiers who served in a given regiment whose parish of origin is in the same county as the regiment's headquarters. Regiments organised in the vertical axis correspond to the 45 regiments in the British Army that had pre-specified recruiting areas.

TABLE B.1
FIRST-STAGE RESULTS

	(1)	(2)
A. Log Specification		
	$Log(d^{WW1})$	$Log(d^{WW1})$
z	0.239	0.220
	(0.023)	(0.022)
F-stat	111	104
R2	0.82	0.83
Obs.	6578	6578
B. Rate Specification		
	$D.rate^{WWI}$	$D.rate^{WWI}$
$\tilde{z}$	0.239	0.223
	(0.049)	(0.044)
F-stat	24	25
R2	0.21	0.27
Obs.	8180	8180
Historic County FE	Y	Y
Extra Controls & Reg. mob. shares	N	Y

Notes: First-stage OLS estimates of the effect of the instrument on WWI deaths at the parish level. Panel A corresponds to the first-stage of our baseline IV specification in logarithms. The instrument z is defined as indicated in Section 4.1. All specifications in this panel control for the logarithms of 1911 population and WWI mobilisation, and historic county fixed effects. Columns 2 and 4 include extra demographic and economic controls and regiment mobilisation shares (see text for details). Panel B corresponds to the first-stage of our specification in death rates per capita. The instrument $\tilde{z}$ corresponds to the predicted death rate per capita obtained by exponentiating z and dividing by 1911 population. All specifications in this panel control for WWI mobilisation rates per capita, and historic county fixed effects. Columns 2 and 4 include extra demographic and economic controls and regiment mobilisation shares (see text for details). Standard errors clustered at the historic county level in parentheses.

Table B.2

Effect of WWI deaths on WW2 medals (Soldier-level)

	(1)	(2)	(3)	(4)
A. OLS: Medal WW2				
$Log(d^{WW1})$	0.219	0.244	0.210	0.166
	(0.085)	(0.088)	(0.082)	(0.076)
Mean of dep.var.	2.158	2.165	2.157	2.158
Obs.	214369	203520	214360	214369
B. OLS: N. Medals WW2				
$Log(d^{WW1})$	0.307	0.333	0.291	0.245
	(0.104)	(0.107)	(0.100)	(0.092)
Mean of dep.var.	2.436	2.446	2.435	2.436
Obs.	214369	203520	214360	214369
C. Poisson: N. Medals WW2				
$Log(d^{WW1})$	0.316	0.330	0.299	0.233
	(0.096)	(0.098)	(0.092)	(0.083)
Mean of dep.var.	2.436	2.446	2.458	2.436
Obs.	214369	203520	212338	214369
Parish controls	Y	Y	Y	Y
Age FE	N	Y	N	N
Regiment FE	N	N	Y	N
Rank FE	N	N	N	Y

Notes: Soldier-level analysis of the effect of WWI deaths on the probability of receiving one or more WW2 gallantry medals (Panel A) or the number of medals received (Panels B and C). All outcomes are multiplied by 100. Sample restricted to servicemen who died in WW2 and whose place of origin was successfully geolocated. Estimation is carried out using OLS in Panels A and B and a Poisson model in Panel C. Different sets of controls are used in each column as indicated in the table foot. Parish controls are the full set of controls used in our parish-level analysis. Age fixed effects are dummies for age in 1939 (with catch-all dummies for individuals above 65 and below 10, respectively). Rank fixed effects are dummies for each rank. Regiment fixed effects are dummies for serving in a given regiment. Standard errors clustered at the parish level in parentheses.

Table B.3
Effect of WWI Deaths on Memorials and Civic Capital Measures

	(1) Mutuals/Char.	(2) Memorials	(3) Legions
A. LPM (dummy outcome)			
$Log(d^{WW1})$	0.018	0.019	0.022
	(0.007)	(0.005)	(0.004)
Mean of dep.var.	0.42	0.14	0.10
Obs.	7878	7878	7878
B. Poisson (count outcome)			
$Log(d^{WW1})$	0.147	0.012	0.025
	(0.032)	(0.007)	(0.008)
Mean of dep.var.	1.08	0.17	0.13
Obs.	7878	7878	7788

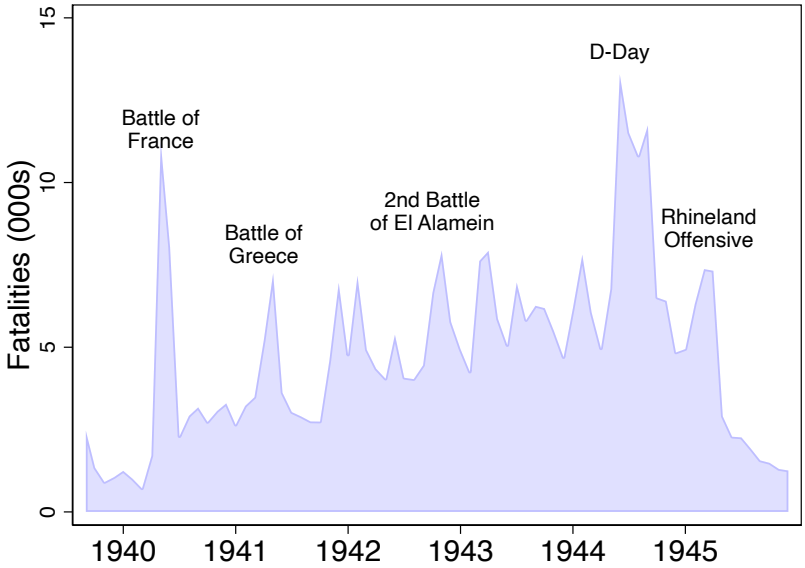
Notes: Effect of WWI deaths on the charities and mutuals established (column 1), listed memorials built (column 2) and British Legion branches (column 3). Panel A presents estimates for linear probability models where the outcomes are dummies taking value 1 if the corresponding institution is present in the parish. Panel B shows average marginal effects from a Poisson model estimated using the corresponding count variables instead. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Table B.4
Effect of WWI Deaths on WW2 Deaths in Campaigns with no German Presence

	OLS		IV	
	$Log(d^{WW2}_{Other})$	$Log(d^{WW2}_{Other})$	$Log(d^{WW2}_{Other})$	$Log(d^{WW2}_{Other})$
$Log(d^{WW1})$	0.184	0.167	0.639	0.417
	(0.021)	(0.020)	(0.117)	(0.110)
First stage F-stat			32.1	32.6
Obs.	3114	3114	3114	3114
Extra Controls & Reg. mob. shares	N	Y	N	Y

Notes: Estimates of the effect of WWI deaths on WW2 deaths taking place in the East Africa and South-East Asian campaigns. Columns 1 and 2 report OLS estimates and columns 3 and 4 report IV estimates. All specifications control for the logarithms of 1911 population and WWI mobilisation, and historic county fixed effects. In columns 2 and 4, we add the full set of controls described in Section 4.1 as well as regiment mobilisation shares. First-stage F-statistics are provided for IV estimates. Standard errors clustered at the historic county level in parentheses.

Figure B.3
Timeline of WW2 Deaths of British Service Personnel



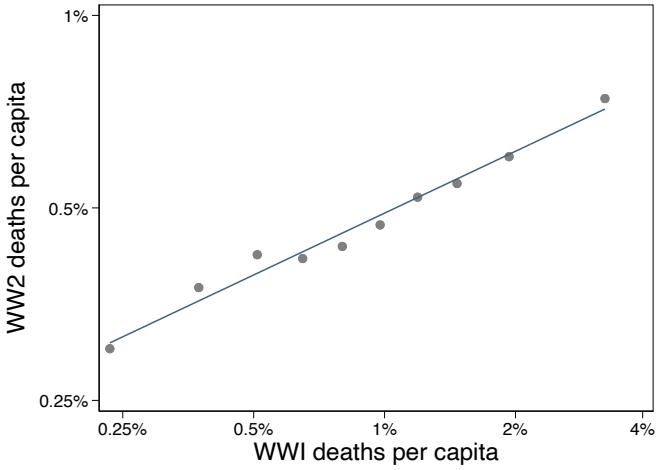
Notes: Number of British Army, Navy and Air Force fatalities in each month during WW2. Overlaid text indicates the name of five key battles: Battle of France (May 1940), Battle of Greece (April 1941), 2nd Battle of El Alamein (October 1942), D-Day (June 1944), and Rhineland Offensive (February 1945). Source: authors' elaboration based on Commonwealth War Graves Commission data.

Table B.5
Reciprocity

	(1) $\Delta houses$	(2) New station	(3) $Log(d^{WW2})$	(4) $Log(d^{WW2})$	(5) $Log(d^{WW2})$
$Log(d^{WW1})$	0.257	0.004	0.255	0.256	0.255
	(1.007)	(0.003)	(0.016)	(0.016)	(0.016)
$\Delta houses1921-31$			0.003		0.003
			(0.000)		(0.000)
$Newstation1918-39$				0.091	0.072
				(0.055)	(0.053)
Mean of dep.var.	21.68	0.04	1.70	1.70	1.70
Obs.	6682	6682	6682	6682	6682

Notes: Parish level OLS estimates to assess the role of reciprocity. Columns 1 and 2 regress percentage changes in housing between 1921 and 1931 (winsorised at the 1st and 99th percentiles) and an indicator for a new rail station between 1918 and 1939 on WWI deaths. Columns 3 to 5 regress WW2 deaths on WWI deaths and these variables. The full set of controls and county fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Figure B.4
Death Rates in WWI and WW2



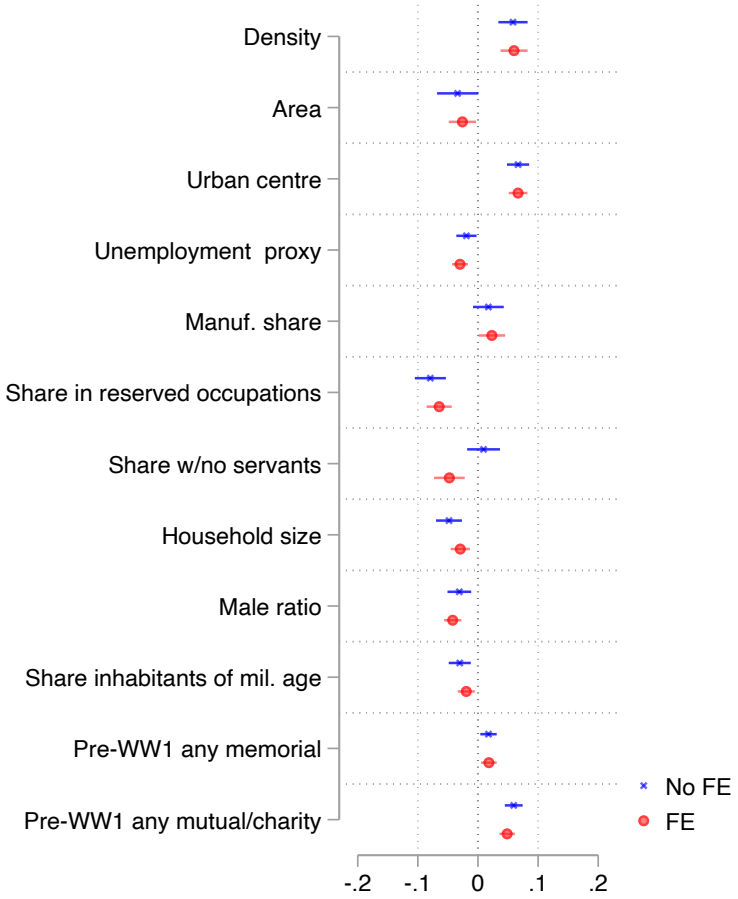
Notes: Binned scatter plot (log scale) of the death rate in WWI, defined as the number of service personnel killed in the war divided by 1911 population at the parish level, and the death rate in WW2, defined as WW2 deaths over 1931 population (last available figure).

Table B.6
WWI Deaths and Election Turnout: DID Estimates

	(1) Turnout (rate)	(2) Log(Turnout)
$Log(d^{WW1}) \times$ Post	0.011	0.016
	(0.006)	(0.009)
Obs.	3806	3806

Notes: Difference-in-difference estimates. Post period defined as the years after WWI. Outcome in column 1 is the turnout rate, defined as the fraction of eligible voters, winsorised at the 1st and 99th percentiles. Outcome in column 2 is the log number of voters who cast their vote in each election. All columns control for interactions between the logarithms of 1911 population, mobilisation and historic county dummies with a post dummy. All specifications include constituency and year fixed effects. Standard errors clustered at the constituency level in parentheses.

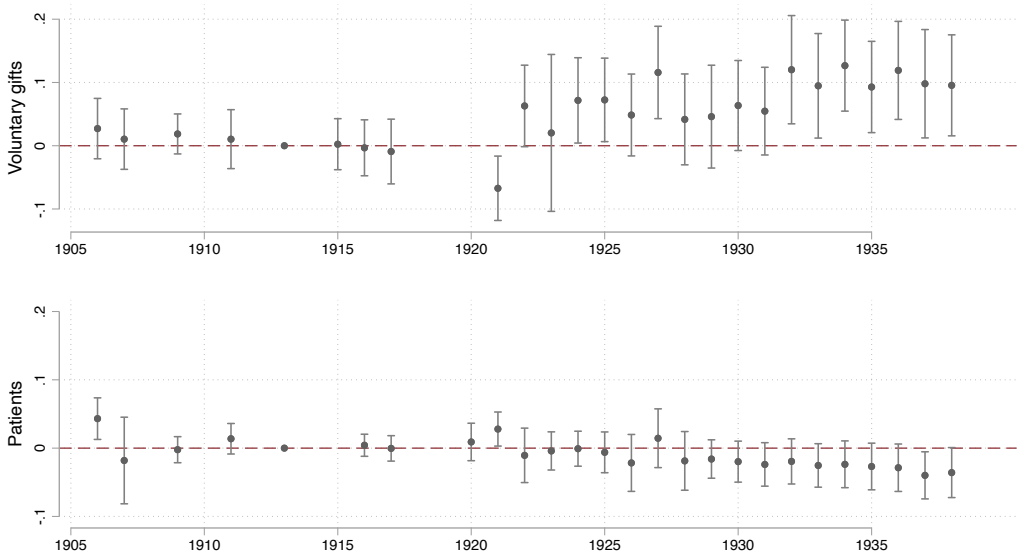
Figure B.5
Correlates of WWI mortality



Notes: Univariate OLS regressions of $Log(d^{WWI})$ on several parish-level pre-WWI characteristics, with and without historic county fixed effects. The log of 1911 population is included in all estimations. All variables are standardised. Confidence intervals constructed using standard errors clustered at the historic county level.

FIGURE B.6

EVENT-STUDY GRAPHS: DONATIONS TO VOLUNTARY HOSPITALS



Notes: Dots represent estimates corresponding to each of the interactions of the (log) number of WWI deaths with year dummies. Top panel estimates obtained using the logarithm of the sum of voluntary gifts received by a hospital in a given year. Bottom panel estimates obtained using the logarithm of the number of patients treated by the hospital in a year. All specifications include hospital fixed effects and year effects. Standard errors clustered at the hospital level in parentheses.

TABLE B.7

EFFECT OF WWI DEATHS ON INTER-WAR OUTCOMES

	(1) Unempl.	(2) Fem.LFP	(3) Pop.Growth	(4) Working Age	(5) Male Ratio
A. 1921					
$Log(d^{WW1})$	0.037	-0.241	-0.976	-0.237	0.074
	(0.076)	(0.206)	(1.308)	(0.131)	(0.069)
Mean dep.var.	7.77	28.74	8.20	65.62	8.50
Observations	1801	1801	1801	1801	1801
B. 1931					
$Log(d^{WW1})$	0.007	-0.354	1.122	-0.255	-0.080
	(0.077)	(0.236)	(3.272)	(0.112)	(0.067)
Mean dep.var.	5.79	29.66	23.01	67.88	7.27
Observations	1802	1802	1802	1802	1802

Notes: OLS estimation results of the effect of WWI deaths on inter-war economic and demographic outcomes (in rates in percentage points, all winsorised at the 1st and 99th percentiles) at the district level. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Table B.8

1918 Flu

	(1) $Log(d^{Flu})$	(2) $Log(d^{WW2})$	(3) $Log(d^{WW2})$	(4) $Log(d^{WW2})$
$Log(d^{WW1})$	0.033	0.453		0.457
	(0.049)	(0.077)		(0.080)
$Log(d^{Flu})$			-0.114	-0.134
			(0.120)	(0.140)
Mean of dep.var.	5.39	5.64	5.64	5.64
Obs.	265	265	265	265
R2	0.87	0.86	0.83	0.86

Notes: District-level OLS regression estimation results on 1918 flu deaths. The full set of controls is included in all specifications. Standard errors clustered at the historic county level in parentheses.

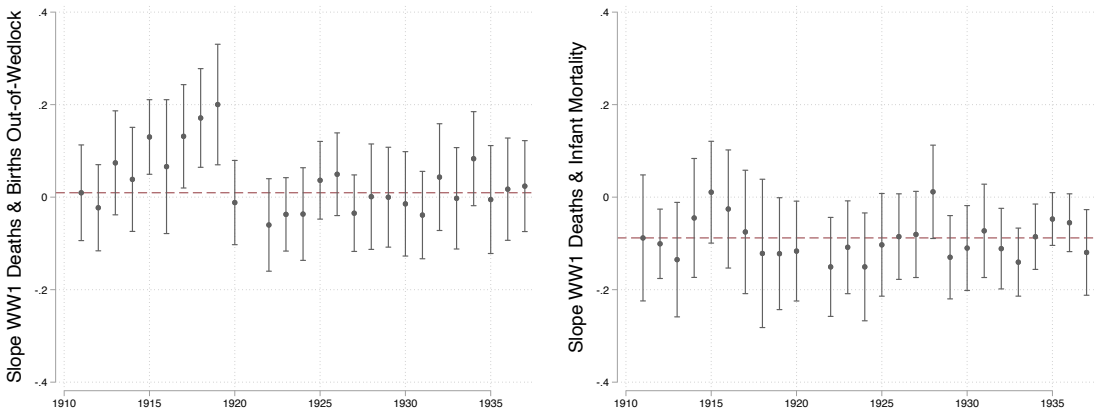
Table B.9

Effect of WWI deaths on male suicides in 1919

	(1) Suic./pop ‰	(2) Suic./death ‰	(3) N Suic. (Poisson)
$Log(d^{WW1})$	-0.000	-0.024	0.008
	(0.000)	(0.028)	(0.062)
Mean of dep.var.	0.01	0.45	3.93
Obs.	1778	1778	1727

Notes: OLS estimation results of the effect of WWI deaths on the number of male suicides per capita and per death multiplied by 100 both winsorised at the 1st and 99th percentiles (columns 1 and 2), and Poisson estimation results using the count of male suicides (column 3). The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Figure B.7
Event-Study Graphs: Pre & Inter-War Economic Outcomes



Notes: Each point is an estimate for yearly interactions of the (log) number of WWI deaths on the outcome, which is winsorised at the 1st and 99th percentiles. All specifications include district fixed effects, year effects and interactions between year dummies and the log of WWI mobilisation (see text). Standard errors clustered at the historic county level. No data available for 1921.

TABLE B.10

ROBUSTNESS: ALTERNATIVE DEFINITIONS OF IV – EFFECT OF WWI DEATHS ON WW2 DEATHS

	(1) $Log(d^{WW2})$	(2) $Log(d^{WW2})$	(3) $Log(d^{WW2})$	(4) $Log(d^{WW2})$
A. IV using regular instrument				
$Log(d^{WW1})$	0.496	0.518	0.502	0.587
	(0.071)	(0.069)	(0.072)	(0.169)
First stage F-stat	52.6	62.8	58.7	30.1
Obs.	5687	5687	5687	5687
R2	0.76	0.77	0.77	0.76
B. IV Using 1917-1918 Deaths				
$Log(d^{WW1})$	0.594	0.623	0.595	0.753
	(0.068)	(0.067)	(0.072)	(0.142)
First stage F-stat	84.8	90.6	85.3	42.7
Obs.	5308	5308	5308	5308
R2	0.76	0.76	0.77	0.74
C. IV Excluding Pals Battalions				
$Log(d^{WW1})$	0.430	0.429	0.398	0.364
	(0.137)	(0.155)	(0.157)	(0.232)
First stage F-stat	36.8	39.9	35.5	23.2
Obs.	5556	5556	5556	5556
R2	0.77	0.78	0.78	0.78
D. IV Using Infantry Battalions				
$Log(d^{WW1})$	0.566	0.601	0.561	0.555
	(0.105)	(0.103)	(0.105)	(0.140)
First stage F-stat	40.9	49.6	46.2	27.6
Obs.	5562	5562	5562	5562
R2	0.76	0.76	0.77	0.77
Controls	Y	Y	Y	Y
County FE	N	Y	Y	Y
Regiment mob. shares	N	N	Y	N
Pred. regiment deaths (z^r)	N	N	N	Y

Notes: IV estimation results of the effect of WWI deaths on WW2 deaths. In Panel A we report baseline results using eq. 1 and different sets of controls for comparison. In Panels B-D we define mobilisation and the instrument differently. Specifically, mobilisation is constructed using all soldiers (Panel B), only soldiers serving in battalions other than Pals (Panel C), or in infantry battalions (Panel D). The number of WWI deaths and the instrument are constructed using only deaths occurring in 1917-18 (Panel B), deaths of soldiers from battalions other than Pals (Panel C), or from infantry battalions (Panel D). Different sets of controls and fixed effects are used in each column (see text for details). Standard errors clustered at the historic county level.

Table B.11

Robustness: Effect of WWI Deaths on WW2 Deaths – Dealing with the Log of Zero

	(1) no const.	(2) c = 0.5	(3) c = 1	(4) c = 2
A. OLS				
$Log(d^{WW1}+c)$	0.256	0.273	0.293	0.318
	(0.016)	(0.017)	(0.017)	(0.016)
Obs.	6682	10807	10807	10807
R2	0.77	0.72	0.76	0.79
B. IV				
$Log(d^{WW1}+c)$	0.504	0.627	0.657	0.699
	(0.067)	(0.060)	(0.045)	(0.035)
Obs.	5778	10807	10807	10807

Notes: OLS (panel A) and IV (panel B) estimation results of the effect of WWI deaths on WW2 deaths. In column 1 we report the baseline estimates from the model in logarithms, where parishes with zero reported WWI or WW2 deaths are dropped. In columns 2-4 we estimate our baseline model adding a constant *c* to the number of dead before taking logarithms for both the outcome (the number of WW2 dead), the variable of interest (the number of WWI dead), and (in panel B) the instrument. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Table B.12

Robustness: Effect of WWI Deaths on WW2 Deaths - iOLS Estimator

	(1) δ = 0.1	(2) δ = 1	(3) δ = 2	(4) δ = 10	(5) δ = 100
$Log(d^{WW1})$	0.218	0.238	0.243	0.252	0.259
	(0.027)	(0.021)	(0.020)	(0.021)	(0.022)
Obs.	10807	10807	10807	10807	10807
λ stat.	1.042	1.039	1.038	1.034	1.030

Notes: iOLS estimation results of the effect of WWI deaths on WW2 deaths using Bellégo, Benatia and Pape (2022)'s iterative OLS estimator. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Table B.13
Robustness: Spatial HAC Standard Errors

	OLS		IV	
	(1) $Log(d^{WW2})$	(2) $Log(d^{WW2})$	(3) $Log(d^{WW2})$	(4) $Log(d^{WW2})$
$Log(d^{WW1})$	0.278	0.255	0.560	0.504
HAC Errors:	(0.018)	(0.018)	(0.076)	(0.080)
Clustered Errors:	(0.017)	(0.017)	(0.066)	(0.067)
Obs.	5778	5778	5778	5778
Extra Controls	N	Y	N	Y
Regiment mob. shares	N	Y	N	Y

Notes: Estimates of the effect of WWI deaths on WW2 deaths. In all columns standard errors are computed incorporating spatial dependence in the error term using a spatial heteroskedasticity and autocorrelation robust standard errors following the tradition of Conley (1999), using a Bartlett kernel with a 50km bandwidth to model dependence. Columns 1 and 2 correspond to OLS estimates obtained using the *reg2hdfespatial* Stata command by Fetzer (2020), which is itself based on the previous implementation by Hsiang (2010). Columns 3 and 4 correspond to IV estimates obtained using the *spatial_hac_iv* Stata command created by Foreman (2020). All specifications include mobilisation and economic controls. Regiment mobilisation shares are included in columns 1 and 3. In columns 2 and 4 we control for our measure of predicted deaths constructed using regiment-level mortality, z^r.

Table B.14
Robustness: Effect of WWI Deaths on WW2 Medals – Dealing with the Log of Zero

	(1) no const.	(2) c = 0.5	(3) c = 1	(4) c = 2
A. OLS				
$Log(d^{WW1}+c)$	0.040	0.031	0.043	0.059
	(0.006)	(0.004)	(0.005)	(0.006)
Obs.	6578	10807	10807	10807
R2	0.31	0.28	0.28	0.28
B. IV				
$Log(d^{WW1}+c)$	0.051	0.118	0.129	0.141
	(0.026)	(0.021)	(0.020)	(0.021)
Obs.	6578	10807	10807	10807

Notes: OLS (panel A) and IV (panel B) estimation results of the effect of WWI deaths on an indicator equal to one if at least one soldier in the parish received a medal for bravery during WW2. In column 1 we report the baseline estimates from the model in logarithms, where parishes with zero reported WWI or WW2 deaths are dropped. In columns 2-4 we estimate our baseline model adding a constant c to the number of dead before taking logarithms. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Table B.15

Robustness: Alternative definitions of IV – Effect of WWI Deaths on WW2 Medals

	(1) $Medal^{WW2}$	(2) $Medal^{WW2}$	(3) $Medal^{WW2}$	(4) $Medal^{WW2}$
A. IV using regular instrument				
$Log(d^{WW1})$	0.114	0.106	0.050	0.141
	(0.029)	(0.029)	(0.031)	(0.052)
First stage F-stat	65.8	80.8	75.7	40.1
Obs.	6433	6433	6433	6433
B. IV Using 1917-1918 Deaths				
$Log(d^{WW1})$	0.143	0.142	0.070	0.212
	(0.030)	(0.031)	(0.034)	(0.052)
First stage F-stat	95.6	105.8	97.9	52.0
Obs.	5958	5958	5958	5958
R2	0.23	0.24	0.31	0.17
C. IV Excluding Pals Battalions				
$Log(d^{WW1})$	0.183	0.163	0.089	0.194
	(0.045)	(0.049)	(0.050)	(0.066)
First stage F-stat	54.3	63.6	57.0	39.6
Obs.	6257	6257	6257	6257
R2	0.18	0.22	0.30	0.19
D. IV Using Infantry Battalions				
$Log(d^{WW1})$	0.147	0.118	0.051	0.131
	(0.037)	(0.038)	(0.040)	(0.050)
First stage F-stat	50.1	64.0	58.3	33.4
Obs.	6282	6282	6282	6282
R2	0.22	0.25	0.31	0.25
Controls	Y	Y	Y	Y
County FE	N	Y	Y	Y
Regiment mob. shares	N	N	Y	N
Pred. regiment deaths (z^r)	N	N	N	Y

Notes: IV estimation results of the effect of WWI deaths on an indicator equal to one if at least one soldier in the parish received a medal for bravery during WW2. In Panel A we report baseline results using eq. 1 and different sets of controls for comparison. In Panels B-D we define mobilisation and the instrument differently. Specifically, mobilisation is constructed using all soldiers (Panel B), only soldiers not serving in Pals (Panel C), or serving in infantry battalions (Panel D). The number of WWI deaths and the instrument are constructed using only deaths occurring in 1917-18 (Panel B), deaths of soldiers not serving in Pals (Panel C), or from soldiers serving in infantry battalions (Panel D). Different sets of controls and fixed effects are used in each column (see text for details). Standard errors clustered at the historic county level.

Table B.16

Robustness: WW2 Honours – Spatial HAC Standard Errors

	OLS		IV	
	(1) Medal WW2	(2) Medal WW2	(3) Medal WW2	(4) Medal WW2
$Log(d^{WW1})$	0.050	0.037	0.127	0.051
HAC Errors:	(0.004)	(0.004)	(0.025)	(0.026)
Clustered Errors:	(0)	(0)	(0)	(0)
Obs.	7878	7878	6578	6578
Extra Controls	N	Y	N	Y
Regiment mob. shares	N	Y	N	Y

Notes: Estimates of the effect of WWI deaths on an indicator equal to one if at least one soldier in the parish received a medal for bravery during WW2. In all columns standard errors are computed incorporating spatial dependence in the error term using a spatial heteroskedasticity and autocorrelation robust standard errors following Conley (1999), using a Bartlett kernel with a 50km bandwidth to model dependence. Columns 1 and 2 show OLS estimates obtained using the *reg2hdfespatial* Stata command by Fetzer (2020), which is itself based on the previous implementation by Hsiang (2010). Columns 3 and 4 show IV estimates obtained using the *spatial_hac_iv* Stata command created by Foreman (2020). All specifications include mobilisation and economic controls, as well as regiment mobilisation shares.

Table B.17

Robustness: Alternative IV Definitions – Effect of WWI Deaths on Civic Capital

	(1) Memorial	(2) Legion	(3) Mutual/char.
A. IV Using 1917-1918 Deaths			
$Log(d^{WW1})$	0.097	0.073	0.078
	(0.031)	(0.025)	(0.050)
First stage F-stat	126.3	126.3	126.3
Obs.	6015	6015	6015
B. IV Excluding Pals Battalions			
$Log(d^{WW1})$	0.106	0.084	0.137
	(0.036)	(0.026)	(0.047)
First stage F-stat	122.8	122.8	122.8
Obs.	6380	6380	6380
C. IV Using Infantry Battalions			
$Log(d^{WW1})$	0.110	0.069	0.119
	(0.029)	(0.022)	(0.045)
First stage F-stat	95.4	95.4	95.4
Obs.	6416	6416	6416

Notes: IV estimation results of the effect of WWI deaths on proxies for civic capital measures in the inter-war period. Estimates obtained using modified versions of the instrument described in the main text. In Panel A, we build our instrument by calculating death rates using only deaths taking place in 1917 and 1918. In Panel B, the instrument is built excluding Pals' Battalions (see Section 2). In Panel C, the instrument is built using only infantry regiments. Outcomes are a dummy indicating whether a WWI listed memorial (col. 1), a branch of the British Legion (col. 2) or a mutual or charity (col. 3) was established in the parish in the inter-war period. Associated first-stage F-statistics indicated in each panel foot. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

TABLE B.18

ROBUSTNESS: EFFECT OF WWI DEATHS ON CIVIC CAPITAL – OLS ESTIMATES

	(1) no const.	(2) c = 0.5	(3) c = 1	(4) c = 2
A. Outcome: Listed Memorial Dummy				
$Log(d^{WW1} + c)$	0.019	0.022	0.030	0.041
	(0.005)	(0.004)	(0.005)	(0.007)
Obs.	7878	10807	10807	10807
R2	0.21	0.19	0.20	0.20
B. Outcome: Legion Branch Dummy				
$Log(d^{WW1} + c)$	0.022	0.019	0.026	0.036
	(0.004)	(0.003)	(0.004)	(0.005)
Obs.	7878	10807	10807	10807
R2	0.23	0.22	0.22	0.22
C. Outcome: Charity/Mutual Dummy				
$Log(d^{WW1} + c)$	0.018	0.024	0.029	0.034
	(0.007)	(0.006)	(0.007)	(0.008)
Obs.	7878	10807	10807	10807
R2	0.23	0.24	0.24	0.24

Notes: OLS estimates of the effect of WWI deaths on proxies for civic capital measures in the inter-war period. In Panel A, the outcome is a dummy taking value 1 if the parish has a WWI listed memorial. In Panel B, the outcome is a dummy taking value 1 if the British Legion created a branch in the parish in the inter-war period. In Panel C, the outcome is a dummy taking value 1 if a mutual or charity was recorded as created in the parish during the inter-war period. Baseline estimates from the model in logarithms, where parishes with zero reported WWI deaths and/or zero mobilisation are dropped, are reported in column 1. In columns 2-4, we estimate our baseline model adding a constant *c* to the number of dead before taking logarithms. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Table B.19

Robustness: Effect of WWI Deaths on Civic Capital – IV Estimates

	(1) no const.	(2) c = 0.5	(3) c = 1	(4) c = 2
A. Outcome: Listed Memorial Dummy				
$Log(d^{WW1} + c)$	0.095	0.077	0.095	0.112
	(0.031)	(0.026)	(0.024)	(0.023)
Obs.	6578	10807	10807	10807
R2	0.20	0.18	0.18	0.18
B. Outcome: Legion Branch Dummy				
$Log(d^{WW1} + c)$	0.069	0.108	0.109	0.105
	(0.022)	(0.026)	(0.023)	(0.021)
Obs.	6578	10807	10807	10807
R2	0.22	0.15	0.18	0.20
C. Outcome: Charity/Mutual Dummy				
$Log(d^{WW1} + c)$	0.146	0.155	0.142	0.123
	(0.049)	(0.046)	(0.038)	(0.033)
Obs.	6578	10807	10807	10807
R2	0.20	0.20	0.22	0.23

Notes: Instrumental variable estimates of the effect of WWI deaths on our parish-level measures of civic capital. In Panel A, the outcome is a dummy taking value 1 if the parish has a WWI listed memorial. In Panel B, the outcome is a dummy taking value 1 if the British Legion created a branch in the parish in the inter-war period. In Panel C, the outcome is a dummy taking value 1 if a mutual or charity was recorded as created in the parish during the inter-war period. Baseline estimates from the model in logarithms, where parishes with zero reported WWI deaths and/or zero mobilisation are dropped, are reported in column 1. In columns 2-4, we estimate our baseline model adding a constant *c* to the number of dead before taking logarithms for both the variable of interest (the number of WWI dead) and its instrument. The full set of controls and fixed effects is included in all specifications. Standard errors clustered at the historic county level in parentheses.

Appendix C. Additional Robustness Checks

In this Appendix we present a series of robustness checks which supplement those described in Section 6 in the text. We first provide estimates of the effect of WWI deaths on the presence of listed memorials by including in our set of memorials also those listed after the centenary anniversary of the beginning of WWI in 2014. OLS and IV estimates including or excluding memorials listed post-2014 are reported in Table C.1 together with the baseline ones and show that the inclusion of these additional memorials has little effect qualitatively in the OLS and IV estimates.

TABLE C.1
ROBUSTNESS: EXCLUDING MEMORIALS LISTED AFTER THE WWI CENTENARY COMMEMORATIONS

	Excluding Listed post-2014		Including Listed post-2014	
	Memorial	Memorial	Memorial	Memorial
$Log(d^{WW1})$	0.019	0.095	0.021	0.114
	(0.005)	(0.031)	(0.008)	(0.049)
Obs.	7878	6578	7878	6578
Estimator	OLS	2SLS	OLS	2SLS

Notes: OLS and IV estimates of the effect of WWI deaths on various indicators for the presence of war memorials in the parish. In all columns the outcome is an indicator taking value one if there is a listed memorial in the parish. In columns 1 and 2, this definition excludes memorials listed after the centenary of the beginning of the war in 2014. In columns 3 and 4, these are included in the set of memorials when building the outcome. All specifications include the full set of controls, fixed effects for historic county and mobilisation shares. Standard errors clustered at the historic county level in parentheses.

We next conduct a second robustness check for memorials in which we regress the count of *listed* war memorials on WWI deaths, controlling for the total number of public WWI memorials in the parish. This is designed to probe whether our estimates of the response of listed memorials are driven by a mechanical relationship between WWI deaths and memorials in general (listed and non-listed). To define the number of *public* WWI memorials we count all WWI memorials in the parish excluding corporate memorials and those memorials that commemorate only a single death (typically, these are gravestones). OLS and IV estimates in Table C.2 show that results are broadly robust to controlling for the number of public memorials, suggesting our findings are not wholly driven by a mechanical relationship between deaths and memorials.

Our results are also robust to dropping the grouped parish comprising London from the estimation sample. This unit comprises several parishes roughly corresponding to the London conurbation in 1911. As shown in the summary results provided in Table C.3, the exclusion

Table C.2

Robustness: Controlling for the number of all WWI memorials

	Excluding Listed post-2014		Including Listed post-2014	
	Memorials	Memorials	Memorials	Memorials
$Log(d^{WW1})$	0.012	0.102	0.018	0.207
	(0.007)	(0.034)	(0.010)	(0.051)
Obs.	6578	6578	6578	6578
Estimator	OLS	2SLS	OLS	2SLS

Notes: OLS and IV estimates of the effect of WWI deaths on the count of Listed WWI memorials in the parish, controlling for the count of all public group memorials. In all columns the outcome is the count of listed memorials in the parish. All specifications include the count of all public group war memorials as a control, the full set of controls, fixed effects for historic county and mobilisation shares. In columns 1 and 2, this definition excludes memorials listed after the centenary of the beginning of the war in 2014. In columns 3 and 4, these are included in the set of memorials when building the outcome. Standard errors clustered at the historic county level in parentheses.

of London has virtually no effect on our point estimates and, as a result, has no impact on the qualitative conclusion of the analysis, either for deaths or civic capital.[46]

Table C.3

Robustness: dropping London

	(1) $Log(d^{WW2})$	(2) Medal	(3) Mutual/char.	(4) Memorial	(5) Legion
A. OLS Estimates					
$Log(d^{WW1})$	0.256	0.037	0.018	0.019	0.022
	(0.016)	(0.005)	(0.007)	(0.005)	(0.004)
Obs.	6681	7877	7877	7877	7877
B. IV Estimates					
$Log(d^{WW1})$	0.504	0.051	0.147	0.095	0.069
	(0.067)	(0.027)	(0.049)	(0.031)	(0.022)
Obs.	5777	6577	6577	6577	6577

Notes: OLS and IV estimation results of the effect of WWI deaths on different outcomes as indicated in each column header when dropping the grouped parish of London. Panel A corresponds to OLS estimates. Panel B corresponds to 2SLS estimates. All specifications include the full set of controls, fixed effects for historic county and mobilisation shares. Standard errors clustered at the historic county level.

We conduct an additional robustness check by using an alternative measure of mobilisation. The measure of WWI mobilisation used in parish-level analyses throughout the paper is based

[46] Similarly, the exclusion of London has no impact on the soldier-level results from Section 5. Results available upon request.

on servicemen for which we observe not only the location of origin but also the battalion of mobilisation. This choice reflects that we wish to use the same set of soldiers to create our mobilisation control and to build our instrument. In that way, we ensure that we appropriately account for m_i in the expression for the instrument derived in Section 4.1. However, we may be concerned that this variable measures mobilisation with error. To provide reassurance on this point, we repeat the analysis on our main outcome variables but now using the log of total mobilisation to measure mobilisation. Results are reported in Tables C.4 and C.5 and yield estimates in line with those reported in Section 4.2.

TABLE C.4
ROBUSTNESS: CONTROL FOR ALTERNATIVE MEASURE OF MOBILISATION

	(1) $Log(d^{WW2})$	(2) $Log(d^{WW2})$	(3) $Log(d^{WW2})$
A. OLS Estimates			
$Log(d^{WW1})$	0.298	0.278	0.265
	(0.016)	(0.015)	(0.015)
Obs.	7753	7753	7753
B. IV Estimates			
$Log(d^{WW1})$	0.638	0.584	0.567
	(0.085)	(0.073)	(0.078)
Obs.	5778	5778	5778
Controls	Y	Y	Y
County FE	N	Y	Y
Regimental shares	N	N	Y

Notes: OLS and IV estimates of the effect of WWI deaths on WW2 deaths at the parish level. Mobilisation variable built using all available geolocated soldiers in the FamilySearch data source. Panel A corresponds to OLS estimates. Panel B corresponds to 2SLS estimates. Sets of controls and fixed effects in each specification as indicated in the table foot. Standard errors clustered at the historic county level in parentheses.

We conduct a final robustness check in which we apply as few restrictions and manipulation to the data as possible. Specifically, we undo the imputation of extreme outliers that affects some observations in the sample used in our parish-level analyses (see Appendix Appendix A.3). A summary of results for the main estimates in the paper obtained when not imputing these variables is reported in Table C.6. Once again, point estimates are very similar to the ones reported in the paper in Sections 4.2 and 4.4.

Appendix D. Analysis of compliers

The different magnitude of our IV and OLS estimates might be due to the fact that the treatment effect of WWI deaths is larger in parishes affected by the instrument (the "compliers")

Table C.5
Robustness: Control for Alternative Measure of Mobilisation

	(1) Medal	(2) Medal	(3) Medal
A. OLS Estimates			
$Log(d^{WW1})$	0.044	0.042	0.034
	(0.005)	(0.004)	(0.005)
Obs.	9553	9553	9553
B. IV Estimates			
$Log(d^{WW1})$	0.085	0.082	0.028
	(0.036)	(0.034)	(0.036)
Obs.	6578	6578	6578
Controls	Y	Y	Y
County FE	N	Y	Y
Regimental shares	N	N	Y

Notes: OLS and IV estimates of the effect of WWI deaths on a WW2 medal indicator at the parish level. Mobilisation variable built using all available geolocated soldiers in the FamilySearch data source. Panel A corresponds to OLS estimates. Panel B corresponds to 2SLS estimates. Sets of controls and fixed effects in each specification as indicated in the table foot. Standard errors clustered at the historic county level in parentheses.

Table C.6
Robustness: No Imputation of WWI Mobilisation or Deaths

	(1) $Log(d^{WW2})$	(2) Medal	(3) Mutual/Charity	(4) Memorial	(5) Legion
A. OLS Estimates					
$Log(d^{WW1})$	0.275	0.036	0.018	0.018	0.021
	(0.017)	(0.006)	(0.007)	(0.005)	(0.004)
Obs.	6687	7879	7879	7879	7879
	$Log(d^{WW2})$	Medal	Mutual/Charity	Memorial	Legion
B. IV Estimates					
$Log(d^{WW1})$	0.522	0.062	0.147	0.115	0.078
	(0.086)	(0.032)	(0.056)	(0.036)	(0.026)
Obs.	5781	6579	6579	6579	6579

Notes: OLS and IV estimation results of the effect of WWI deaths on different outcomes as indicated in each column header. Data on WWI deaths and mobilisation obtained without any imputation. Panel A corresponds to OLS estimates. Panel B corresponds to 2SLS estimates. All specifications include the full set of controls, as well as fixed effects for historic county and regiments of mobilisation. Standard errors clustered at the historic county level in parentheses.

than in other parishes. IV estimates will then be larger because IV in general estimates the average treatment effect only for compliers (Imbens and Angrist, 1994). To investigate this

possibility, we characterise the set of compliers in our setting, following Imbens and Rubin (1997).

The original method assumes there is a binary "treatment" variable and a binary instrument. In our setting, both our treatment variable $log(d_i^{WWI})$ and the instrument z_i are continuous, so we discretise them by creating indicators for each variable being above the median, denoted $Z_i = 1(z_i \geq Med(z))$ and $D_i = 1(log(d_i^{WWI}) \geq Med(log(d_i^{WWI})))$. Also denote $D_i(0)$ and $D_i(1)$ the values of the treatment for individual i that would be obtained given the instrument $Z_i = 0$ and $Z_i = 1$, respectively.

The population can then be partitioned in four groups: the *never-takers*, units that are never "treated" irrespective of the value of the instrument: $D_i(0) = 0, D_i(1) = 0$; *always-takers*, units with $D_i(0) = 1, D_i(1) = 1$, *compliers*, for which $D_i(0) = 0, D_i(1) = 1$. The last group of *defiers*, for which $D_i(0) = 1, D_i(1) = 0$ is ruled out by the usual monotonicity assumption. Let ϕ_n, ϕ_a, ϕ_c be the population frequencies of the three types of individuals. Under the standard assumptions of the LATE theorem Imbens and Angrist (1994), we can only learn about the causal effect of D on Y for the sub-population of parishes that are affected by the instrument.

As Imbens and Rubin (1997) discuss, while we cannot in general identify compliers from the data, we can identify some of the non-compliers. For instance, parishes that have $Z_i = 0$ and $D_i = 1$ must be always-takers. Similarly, parishes that have $Z_i = 1$ but $D_i = 0$ must be never-takers (since defiers are ruled out). If one is willing to assume that the instrument is fully independent of the potential outcomes $Y_i(0), Y_i(1)$, one could thus fully characterise the distribution of $Y_i(1)$ for always takers, denoted $g_a(y)$. Analogously, in large samples, we can know the distribution of $Y_i(0)$ for never takers. Because by assumption the instrument is also independent of the type $C_i = a, n, c$, in large samples we can also know the population proportions of each type: $\phi_n = Pr(D_i = 0|Z_i = 1), \phi_a = Pr(D_i = 1|Z_i = 0)$, and hence we can obtain $\phi_c = 1 - \phi_n - \phi_a$.

With a similar procedure, we can calculate averages of several covariates for each group. For instance, the average for some variable W for always-takers can be obtained by the sample equivalent of $E(W_i|C_i = a)$. By the law of iterated expectations, the equivalent for compliers can be found as $E(W_i|C_i = c) = \frac{1}{\phi_c}(E(W_i) - E(W_i|C_i = a)\phi_a - E(W_i|C_i = n)\phi_n)$.

Table D.1 below reports sample averages for several parish-level characteristics for each of the three groups, as well as for the full sample. Compliers – parishes that have high WWI mortality when the instrument predicts them to – have much higher population and density than always- and never-takers. They also have slightly lower mobilisation per capita than the average parish, and lower mortality than always-takers.

These results suggest that our IV estimates predominantly use variation from relatively large and densely populated parishes. Individuals living in large villages and cities may be more exposed to the commemoration of the War. The visibility of memorials in densely populated areas will be higher. Also, ceremonies, parades, and other forms of celebration and remembrance may be easier to organise and be better attended in urban centres than in more dispersed, rural communities. In turn, this means that a given WWI mortality shock could plausibly lead to greater

accumulation of civic capital in denser communities. This hypothesis is consistent with finding IV estimates that are larger than those obtained using OLS.

TABLE D.1
DESCRIPTIVE STATISTICS BY COMPLIER GROUP

	Full sample	**Always T.**	**Never T.**	**Compliers**
Population 1911	3332.30	1217.31	702.60	5935.16
Density 1911	252.87	121.79	115.31	405.72
Share in reserved occupations	0.38	0.37	0.40	0.37
Male ratio	0.50	0.50	0.50	0.50
Mobilisation WW1	277.02	42.25	48.77	544.76
Mobilisation Rate WW1 (%)	5.29	5.24	8.33	4.32
Number WW1 Dead	53.27	11.44	2.65	104.28
Death Rate WW1 (%)	0.97	1.42	0.54	0.74
Death Rate WW2 (%)	0.47	0.52	0.48	0.43
Listed WW1 Memorial indicator	0.30	0.33	0.28	0.27
British Legion indicator	0.08	0.05	0.04	0.11
Mutual/charity indicator	0.36	0.38	0.33	0.35
Share ϕ	1.00	0.38	0.15	0.47

Notes: Sample averages of several parish-level characteristics for different groups of parishes by complier status. The last row presents estimates for the share of parishes in each group.

References for Appendix

Becke, Archibald Frank. 1938. *Order of Battle of Divisions Part 3A. New Army Divisions.* Vol. 1, HM Stationery Office.

Bellégo, Christophe, David Benatia, and Louis Pape. 2022. "Dealing with logs and zeros in regression models." *arXiv preprint arXiv:2203.11820*.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-experimental shift-share research designs." *The Review of Economic Studies*, 89(1): 181–213.

BWO. 1922. "Statistics of the Military Effort of the British Empire During the Great War, 1914-1920."

Cadw. 2023. "Listed Buildings in Wales." [Online; accessed 1-May-2022].

Charity Commission for England and Wales. 2023. "The Register of Charities." [Online; accessed 26-Aug-2021].

Commonwealth War Graves Commission. 2021. "Online Casualty Database." [Online; accessed 1-Jan-2021].

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1–45.

FamilySearch. 2021. "United Kingdom World War I Service Records." [Online; accessed 1-Jan-2021].

Fetzer, Thiemo. 2020. "Can workfare programs moderate conflict? Evidence from India." *Journal of the European Economic Association*, 18(6): 3337–3375.

Financial Conduct Authority. 2023. "Mutuals Public Register." [Online; accessed 1-Feb-2022].

Forces War Records. 2021. "Online Database." [Online; accessed 1-Jan-2021].

Foreman, Timothy. 2020. "The Effects of Dust Storms on Economic Development." *Working Paper*.

Historic England. 2021. "Listed Buildings (points) [dataset]." [Online; accessed 1-Feb-2021].

Historic England. 2023. "Listed War memorials in England [dataset]."

H.M Stationary Office. 1945. "Parliamentary Constituencies (Electors) in England and Wales." Return to the House of Commons, dated 13th June 1945.

Hsiang, Solomon M. 2010. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." *Proceedings of the National Academy of sciences*, 107(35): 15367–15372.

Imbens, Guido W, and Donald B Rubin. 1997. "Estimating outcome distributions for compliers in instrumental variables models." *The Review of Economic Studies*, 64(4): 555–574.

Imbens, Guido W, and Joshua D Angrist. 1994. "Identification and estimation of local average treatment effects." *Econometrica: journal of the Econometric Society*, 467–475.

Imperial War Museum. 2021*a*. "Lives of the First World War." [Online; accessed 1-Mar-2021].

Imperial War Museum. 2021*b*. "War Memorials Register." [Online; accessed 1-Feb-2021].

James, Brigadier EA. 2012. *British Regiments 1914-1918.* Andrews UK Limited.

Johnson, Niall. 2001. "1918-1919 Influenza Pandemic Mortality in England and Wales. [data collection]." UK Data Service. SN: 4350.

Kollman, K, A Hicken, D Caramani, D Backer, and D Lublin. 2019. "Constituency-Level Elections Archive [data file and codebook]." Produced and distributed by Ann Arbor, MI: Center for Political Studies, University of Michigan.

Minnesota Population Center. 2019. "Integrated Public Use Microdata Series, International: Version 7.2 [dataset]."

Office for National Statistics. 2022. "Postcode Directories, 2006-2022." *UK Data Service Census Support*, Accessed August 10, 2022 https://borders.ukdataservice.ac.uk/

pcluts.html.

Schürer, Kevin, and Edward Higgs. 2014. "Integrated Census Microdata (I-CeM), 1851-1911 [data collection]." *UK Data Service. SN*, 7481.

University of Portsmouth. 2011. "Great Britain Historical GIS Project." [Online; accessed 1-Feb-2022].