

Appendix for online publication

The intergenerational transmission of conflict

Gaia Narciso and Battista Severgnini

A. Matching Techniques

In order to link the individuals to the different types of administrative records available, we consider two distinct techniques:

1. *Manual matching.*

We match the individuals manually looking at the different information contained both in administrative records from the Irish Military Archives and the 1911 Irish Census. In addition, we complement and counter-check the matching by making use of the information provided by Connell (2015) and Foster (2015) and by two websites, namely www.irishmedals.ie and www.irishvolunteers.org. Manual matching is based on first and last names and age, allowing for a ± 2 years discrepancy in age as reported in the 1911 Census and the dataset on insurgents. The manual matching involved the translation of Irish names in the Census into their English version or the matching of shortened names. Residence at the time of the 1911 Census and at the time of the pension or medal application is taken into account, but it is not a necessary requirement for the matching.

We construct two indicators on the basis of the manual matching, a conservative indicator (*rebel_con*) and a less conservative one (*rebel_lib*). These two methodologies allow us to identify 508 and 823 individuals as insurgents respectively. To simplify the manual matching, we only consider individuals who were 16 or older at the time of the 1911 Census.

2. *Matching following the methodology suggested by Abramitzky, Mill, and Perez (2021).*

We proceed in 4 different steps:

- (a) We make uniform first and last names by removing incorrect characters (*e.g.*, question and exclamation marks) and by making uniform letters when they are not common in the Irish alphabet (*e.g.*, \AA or \O) and substituting shortened first names with more standardized names using the code written by Abramitzky et al. (2012) and

integrated by a list of Irish names and family names provided by Woulfe (1923). In addition, we translate our information into names which are phonetically equivalent using the phonetic algorithm NYSIIS (Atack et al., 1992).

- (b) We divide both the Military Archives and the Census datasets by blocking the first letter of names and family names and consider age differences lower than 5 years in absolute value.
- (c) We translate both names and family names into phonetic equivalents using the NYSIIS algorithm. Within each block, we compute the phonetic differences using the Jaro-Winkler distances as an indicator (Jaro, 1989; Winkler, 2006). Furthermore, we use an Expectation Maximization algorithm (Dempster et al. 1977) for assigning the probabilities that each two records are a true match.
- (d) We construct two indicators based on efficiency and accuracy. The algorithm chooses hyperparameters for maximizing a weighted average of the true positive rate (TPR) and the positive predictive value (PPV) and of the efficiency of the indicator related to the true positive rate, $TPR = TP/(TP+FN)$, for the accuracy the positive predictive value, $PPV = TP/(TP+FP)$, where TP , FN , and FP are the number of true positives, false negatives, and false positives, respectively. Having a set of different matched individuals, we consider a conservative sample (with $TPR=80$ and $PPV=30$) and a more lenient sample (with $TPR=60$ and $PPV=10$).

The automated matching is based on the entire 1911 Irish Census, with no restrictions in terms of age. Overall, the conservative measure (*rebel_8030*) identifies 169 insurgents, while the less conservative measure (*rebel_6010*) identifies 759 rebels.

B. Tables and Figures

Table B1: Individual characteristics

	Mean	Min	Max	SD
<i>Individual characteristics</i>				
<i>Age</i>	31.93	10	65	15.06
<i>Female</i>	50.21%	0	1	
<i>Literate</i>	93.60%	0	1	
<i>Catholic</i>	72.86%	0	1	
<i>Married</i>	34.63%	0	1	
<i>Irish</i>	14.25%	0	1	
<i>Household size</i>	5.96	1	20	2.73
<i>Occupations</i>				
<i>Professional</i>	15.64%	0	1	
<i>Clerical</i>	1.68%	0	1	
<i>Sales</i>	2.60%	0	1	
<i>Service</i>	5.14%	0	1	
<i>Agriculture</i>	21.88%	0	1	
<i>Production</i>	15.31%	0	1	
<i>Rebel</i>	0.0465%	0	1	
<i>Rebel_lib</i>	0.0268%	0	1	
<i>Rebel_con</i>	0.0168%	0	1	
<i>Rebel_8030</i>	0.0052%	0	1	
<i>Rebel_6010</i>	0.0290%	0	1	
<i>Rebel_pen</i>	0.021%	0	1	

Total number of observations: 2,886,364

Source: Authors' calculations using the administrative historical data, as described in Section 2. The sample only includes individuals who were over 10 years old and under 65 at the time of the 1911 Census.

Table B2: Population distribution by county of residence

<i>County of residence</i>	<i>Freq.</i>	<i>Percent</i>
Antrim	328,519	11.38
Armagh	80,833	2.80
Carlow	24,196	0.84
Cavan	61,095	2.12
Clare	68,166	2.36
Cork	258,580	8.96
Derry	92,460	3.20
Donegal	108,109	3.75
Down	208,728	7.23
Dublin	310,924	10.77
Fermanagh	39,777	1.38
Galway	114,350	3.96
Kerry	100,007	3.46
Kildare	39,274	1.36
Kilkenny	50,868	1.76
Laois	36,081	1.25
Leitrim	40,788	1.41
Limerick	93,799	3.25
Longford	28,967	1.00
Louth	41,629	1.44
Mayo	122,460	4.24
Meath	43,849	1.52
Monaghan	43,551	1.51
Offaly	38,585	1.34
Roscommon	62,090	2.15
Sligo	50,318	1.74
Tipperary	100,931	3.50
Tyrone	95,246	3.30
Waterford	53,244	1.84
Westmeath	38,245	1.33
Wexford	69,635	2.41
Wicklow	41,060	1.42
Total	2,886,364	100.00

Source: Authors' calculations using data using the 1911 Census, as described in Section 2.

Table B3: Population distribution by province of residence

<i>Province of residence</i>	<i>Freq.</i>	<i>Percent</i>
Connacht	390,006	13.51
Leinster	763,313	26.45
Munster	674,727	23.38
Ulster	1,058,318	36.67
Total	2,886,364	100.00

Source: Authors' calculations using data using the 1911 Census, as described in Section 2.

Table B4: Balance test based on the list of rebels

<i>Variable</i>	<i>Mean of matched</i>	<i>Difference Matched- Non Matched</i>	<i>N. of observations</i>
Male	0.89	-0.04	4,662
Year of birth	1890.06	2.40***	3,297
Year of birth (individuals born before 1891)	1883.80	0.55	1,312

Source: Authors' calculations using data using the list of rebels and the 1911 Census, as described in Section 2. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B5: Place of birth characteristics

	Mean	Min	Max	SD
<i>County of birth level</i>				
<i>Excess Mortality Rate (per '000s)</i>	0.09	0.04	0.17	0.03
<i>Crop Failure rate</i>	0.24	-0.40	0.68	0.20
<i>Literacy share</i>	0.79	0.70	0.84	0.09
<i>Claimants 1798</i>				
<i>Out-migration rate</i>	0.04	0.02	0.09	0.02
<i>Barony</i>				
<i>Blight indicator</i>	0.59	0.00	1.00	0.49
<i>IV - Blight infection rate</i>	0.07	-1.93	2.13	0.30
<i>Townland</i>				
<i>GAA Clubs</i>	0.43	0	1	0.49

Source: Authors' calculations using data from different geographical and historical sources, as described in Section 2. Missing data for the variable *Claimants* have been replaced with zeros for 4 counties.

Table B6: Main Specification

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dependent Variable: Rebel</i>						
<i>Crop failure rate</i> <i>(County of birth)</i>	0.00033 (0.000)*** [0.0024]	0.00033 (0.000)*** [0.0014]				
<i>Excess Mortality</i> <i>(County of birth)</i>			0.00329 (0.001)** [0.0425]	0.00335 (0.001)** [0.0503]		
<i>Blight indicator</i> <i>(Barony of birth)</i>					0.00012 (0.000)*	0.00016 (0.000)**
<i>Individual characteristics</i>						
<i>Age</i>	-0.00002 (0.000)*** [0.0000]	-0.00002 (0.000)*** [0.0000]	-0.00002 (0.000)*** [0.0000]	-0.00002 (0.000)*** [0.0000]	-0.00001 (0.000)***	-0.00001 (0.000)***
<i>Female</i>	-0.00085 (0.000)** [0.0000]	-0.00084 (0.000)** [0.0000]	-0.00085 (0.000)** [0.0000]	-0.00084 (0.000)** [0.0000]	-0.00046 (0.000)***	-0.00046 (0.000)***
<i>Literate</i>	0.00010 (0.000) [0.3691]	0.00011 (0.000) [0.3251]	0.00010 (0.000) [0.3640]	0.00011 (0.000) [0.3107]	0.00006 (0.000)	0.00006 (0.000)
<i>Catholic</i>	0.00036 (0.000)** [0.0000]	0.00036 (0.000)** [0.0000]	0.00036 (0.000)** [0.0000]	0.00036 (0.000)** [0.0000]	0.00018 (0.000)***	0.00019 (0.000)***
<i>Married</i>	-0.00013 (0.000)** [0.0013]	-0.00015 (0.000)** [0.0007]	-0.00013 (0.000)** [0.0021]	-0.00015 (0.000)** [0.0013]	-0.00006 (0.000)***	-0.00006 (0.000)***
<i>Household size</i>	0.00002 (0.000)** [0.0104]	0.00002 (0.000)* [0.0271]	0.00002 (0.000)** [0.0086]	0.00002 (0.000)* [0.0290]	0.00001 (0.000)	0.00000 (0.000)
<i>Occupation and Location characteristics</i>						
<i>Professional</i>	-0.00042 (0.000)** [0.0056]	-0.00041 (0.000)** [0.0063]	-0.00042 (0.000)** [0.0050]	-0.00041 (0.000)** [0.0067]	-0.00017 (0.000)**	-0.00018 (0.000)**
<i>Clerical</i>	0.00009 (0.000) [0.3444]	0.00001 (0.000) [0.8678]	0.00009 (0.000) [0.3402]	0.00001 (0.000) [0.8772]	-0.00021 (0.000)	-0.00022 (0.000)
<i>Sales</i>	-0.00033	-0.00040	-0.00033	-0.00040	-0.00009	-0.00013

	(0.000)*	(0.000)*	(0.000)*	(0.000)*	(0.000)	(0.000)
	[0.0977]	[0.0331]	[0.1044]	[0.0365]		
<i>Service</i>	-0.00022	-0.00024	-0.00021	-0.00024	-0.00020	-0.00021
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)***	(0.000)***
	[0.0038]	[0.0014]	[0.0034]	[0.0015]		
<i>Agriculture</i>	-0.00038	-0.00033	-0.00038	-0.00033	-0.00027	-0.00025
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)***	(0.000)**
	[0.0754]	[0.1114]	[0.0831]	[0.1159]		
<i>Production</i>	0.00013	0.00008	0.00013	0.00008	-0.00001	-0.00002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.1346]	[0.3616]	[0.1255]	[0.3342]		
<i>Other country of birth variables</i>						
<i>Emigration rate -</i>	-0.00690	-0.00604	-0.01026	-0.00925	0.01115	0.01276
	(0.002)***	(0.003)**	(0.003)***	(0.003)***	(0.002)***	(0.002)***
	[0.0134]	[0.0395]	[0.0067]	[0.0231]		
<i>Literacy rate</i>	-0.00065	-0.00117	-0.00084	-0.00127	0.00061	0.00027
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
	[0.3302]	[0.1012]	[0.3439]	[0.1667]		
Observations	2,751,083	2,751,083	2,751,083	2,751,083	759,618	759,450
R-squared	0.001	0.005	0.001	0.005	0.001	0.004
Location	Yes	No	Yes	No	Yes	No
Characteristics						
County of residence FE	Yes	No	Yes	No	Yes	No
District of residence FE	No	Yes	No	Yes	No	Yes

***p < 0.01, **p < 0.05, *p < 0.1. *Crop failure rate*, *Excess Mortality*, and *Blight Indicator* are the three measures of the Famine severity, as described in Section 2. Location characteristics include: Share of Catholics, Share of Men, Share of aged 25-40 and Literacy rate in the electoral district of residence, as described in Section 2. Columns 1-4: Robust standard errors clustered by county of birth in parentheses. P-values for t-test of parameter significance using wild bootstrapped standard errors presented in brackets (Cameron and Miller 2015). Columns 5-6: Robust standard errors clustered by barony of birth in parentheses.

Table B7: Concurring factors

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Dependent variable: Rebel</i>					
<i>Crop failure rate</i> (<i>County of birth</i>)	0.00033 (0.000)*** [0.0038]	0.00034 (0.000)*** [0.0010]				
<i>Excess mortality</i> (<i>County of birth</i>)			0.00367 (0.001)*** [0.0179]	0.00405 (0.001)*** [0.0171]		
<i>Blight indicator</i> (<i>Barony of birth</i>)					0.00015 (0.000)**	0.00019 (0.000)***
<i>GAA Clubs</i>	0.00012 (0.000) [0.2155]	-0.00019 (0.000) [0.8540]	0.00012 (0.000) [0.2091]	-0.00017 (0.000) [0.8769]	-0.00000 (0.000)	0.00018 (0.000)
<i>Claimants 1798</i>	0.00000 (0.000) [0.8944]	0.00002 (0.000) [0.6000]	0.00001 (0.000) [0.6728]	0.00002 (0.000) [0.4026]	0.00004 (0.000)	0.00004 (0.000)
Observations	2,751,083	2,751,083	2,751,083	2,751,083	759,618	759,450
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Occupations	Yes	Yes	Yes	Yes	Yes	Yes
Location Characteristics	Yes	No	Yes	No	Yes	No
County of residence FE	Yes	No	Yes	No	Yes	No
District of residence FE	No	Yes	No	Yes	No	Yes

***p < 0.01, **p < 0.05, *p < 0.1. The variable *GAA Clubs* is the number of Gaelic Athletic Associations (GAA) at townland level. The variable *Claimants 1798* is the number of rebellion episodes per county per 1,000 during the 1798 Revolution (Catwell, 2011). Columns 1-4: Robust standard errors clustered by county of birth in parentheses. P-values for t-test of parameter significance using wild bootstrapped standard errors presented in brackets (Cameron and Miller 2015). Column 5-6: Robust standard errors clustered by barony of birth in parentheses.

We gather data on social unrest episodes per 1,000 individuals during the 18th century at a county of birth level. Data on the 1798 rebellion episodes (*Claimants 1798*) are collected from Cantwell (2011).

Social capital may be an important confounding factor in our analysis. To this end, we consider the density of the Gaelic Athletic Association (GAA) branches as an indicator of the presence of social capital. Numerous studies have highlighted the connection between GAA associations and the Irish independence movement. According to Rouse (2015) “[t]his organization came to symbolize the idea of Irish independence and was used to draw a line between the oppressed and their oppressors”. We measure GAA density as the number of associations divided by population at town level in Ireland.

In the analysis conducted in the paper, we have focused on the role played by the Famine in determining the probability of joining the independence movement. But did the Famine have a direct effect on the probability of rebellion, or can we identify other potential alternative mechanisms? Could previous acts of rebellion, rather than the Great Irish Famine, explain the probability of participating in the independence movement in the 20th century? What is the role of social capital in affecting the probability of taking part in the rebellion? We tackle these issues in Table B6, which presents the results of a specification that includes the entire set of controls (as in column 2 of Table 1) and investigates the role of a variable capturing the extent of the Revolution of 1798 (Claimants 1798). We do not observe a relationship between the probability of becoming a rebel and previous insurgency activity in the county of birth, while the sign and statistical significance of the Famine measure is unchanged. The estimation presented in Table B6 also controls for the diffusion of social capital in affecting the decision to rebel. To this end, we consider the density of the Gaelic Athletic Association (GAA) branches as an indicator of the presence of social capital, described in Section 2. The results of the estimation including this measure of social capital confirm our main findings, while we do not find any statistically significant effect of local GAA clubs on the probability of rebelling. We can conclude that, even when controlling for potential concurring factors, there is evidence in support of the Famine’s intergenerational legacy of rebellion.

Table B8: Disaggregated indicators

	(1)	(2)	(3)	(4)	(5)
	<i>Manual matching</i>		<i>Automated Matching</i>		<i>Pension recipients</i>
Dependent variable:	<i>Rebel_lib</i>	<i>Rebel_con</i>	<i>Rebel_6010</i>	<i>Rebel_8030</i>	<i>Rebel_pen</i>
Panel A					
<i>Potato Crop Failure</i>	0.00023 (0.000)*** [0.0003]	0.00016 (0.000)*** [0.0027]	0.00014 (0.000)*** [0.0075]	0.00004 (0.000)*** [0.0205]	0.00022 (0.000)*** [0.0002]
Individual Characteristics	Yes	Yes	Yes	Yes	Yes
Occupations	Yes	Yes	Yes	Yes	Yes
Location Characteristics	Yes	Yes	Yes	Yes	Yes
County of residence FE	Yes	Yes	Yes	Yes	Yes
Observations	2,751,083	2,751,083	2,751,083	2,751,083	2,751,083
Panel B					
<i>Excess Mortality Rate</i>	0.00182 (0.001)* [0.0999]	0.00179 (0.001)* [0.0814]	0.00158 (0.001)** [0.0727]	0.00042 (0.000) [0.1427]	0.00183 (0.001)* [0.1253]
Individual Characteristics	Yes	Yes	Yes	Yes	Yes
Occupations	Yes	Yes	Yes	Yes	Yes
Location Characteristics	Yes	Yes	Yes	Yes	Yes
County of residence FE	Yes	Yes	Yes	Yes	Yes
Observations	2,751,083	2,751,083	2,751,083	2,751,083	2,751,083
Panel C					
<i>Blight indicator</i>	0.00015 (0.000)**	0.00007 (0.000)***	0.00004 (0.000)	-0.00004 (0.000)	0.00013 (0.000)***
Individual characteristics	Yes	Yes	Yes	Yes	Yes
Occupations	Yes	Yes	Yes	Yes	Yes
Location characteristics	Yes	Yes	Yes	Yes	Yes
Province Indicators	Yes	Yes	Yes	Yes	Yes
Observations	759,618	759,618	759,618	759,618	759,618

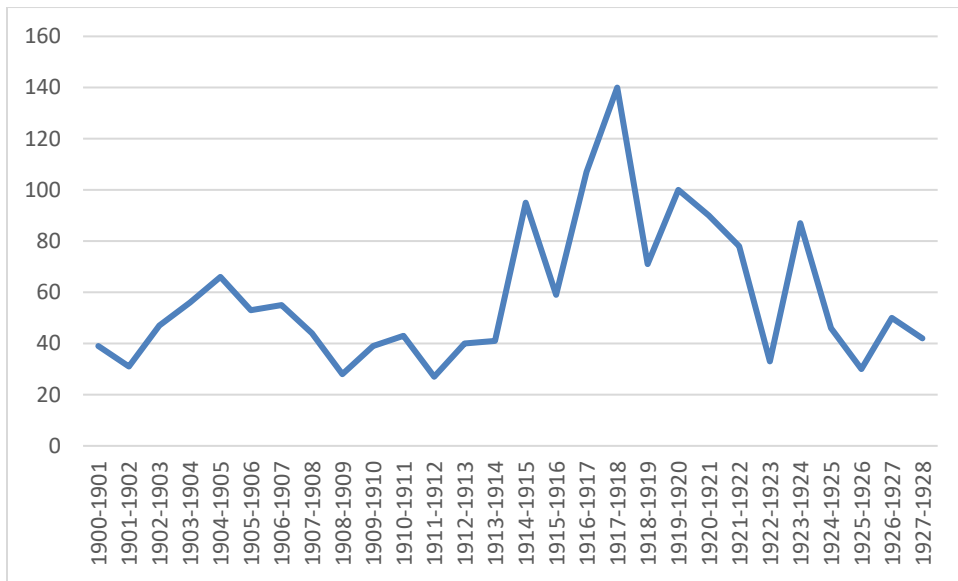
***p < 0.01, **p < 0.05, *p < 0.1. Panel A and B: Robust standard errors clustered by county of birth in parentheses. P-values for t-test of parameter significance using wild bootstrapped standard errors presented in brackets (Cameron and Miller 2015). Panel C: Robust standard errors clustered by barony of birth in parentheses. Columns 1 and 2 present the specification using the two indicators of rebellion based on the manual matching. The first indicator, *Rebel_lib*, is the less conservative measure, while *Rebel_con* is the more stringent one. Columns 3 and 4 present the evidence using the two indicators arising from the automated matching. We distinguish between the less conservative measure (*Rebel_6010*) and the more conservative one (*Rebel_8030*). Finally, in Column 5 the dependent variable is the indicator *Rebel_pen*, which takes the value 1 if individual *i* is identified as a rebel according to at least one of the 4 measures and if the individual is granted a pension and 0 otherwise.

Table B9: Placebo IV analysis

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
FIRST STAGE						
<i>Blight Indicator</i>						
Placebo instrument	0.46199 (1.750)	0.35186 (1.747)	-0.48372 (1.638)	-0.48084 (1.581)	-2.19229 (1.819)	-2.24977 (1.725)
SECOND STAGE						
<i>Rebel</i>						
Blight Indicator	-0.00171 (0.008)	0.00010 (0.002)	0.00416 (0.013)	0.00294 (0.008)	0.00118 (0.001)	0.00109 (0.001)*
Observations	364,906	364,906	364,906	364,487	364,906	364,487
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Occupations	Yes	Yes	Yes	Yes	Yes	Yes
Location Characteristics	No	No	Yes	Yes	Yes	Yes
Concurrent factors	No	No	No	No	Yes	Yes
County of residence FE	Yes	Yes	Yes	No	Yes	No
District of residence FE	No	No	No	Yes	No	Yes

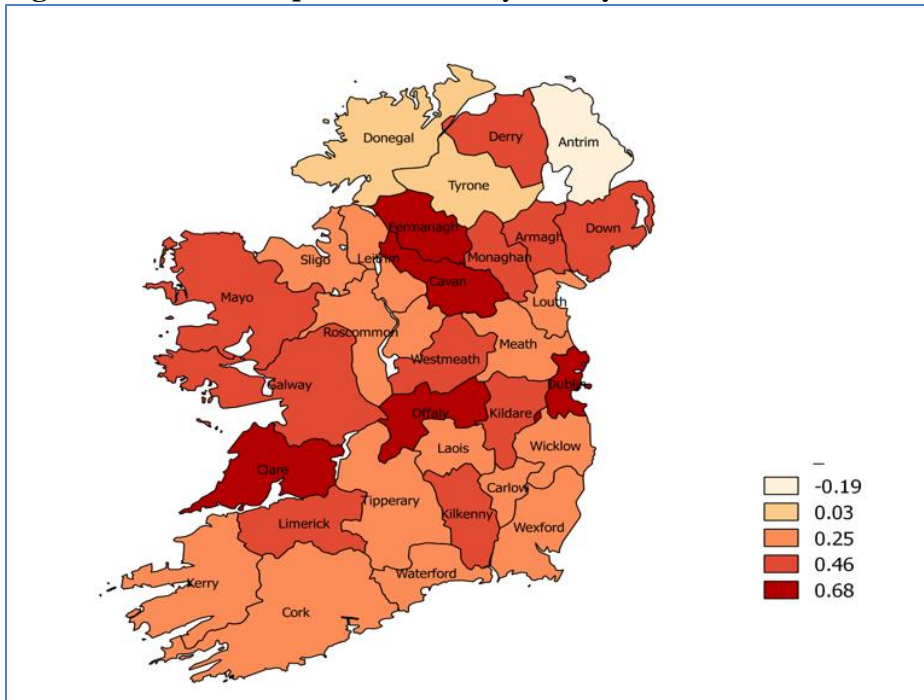
Note: First and second stage of the placebo Instrumental Variable analysis presented in Section 6. Robust standard errors clustered at barony level. *** p<0.01, ** p<0.05, * p<0.1

Figure B1: Number of articles mentioning the word "Famine" on Irish newspapers



Source: Irish Newspaper Archives (www.irishnewsarchive.com)

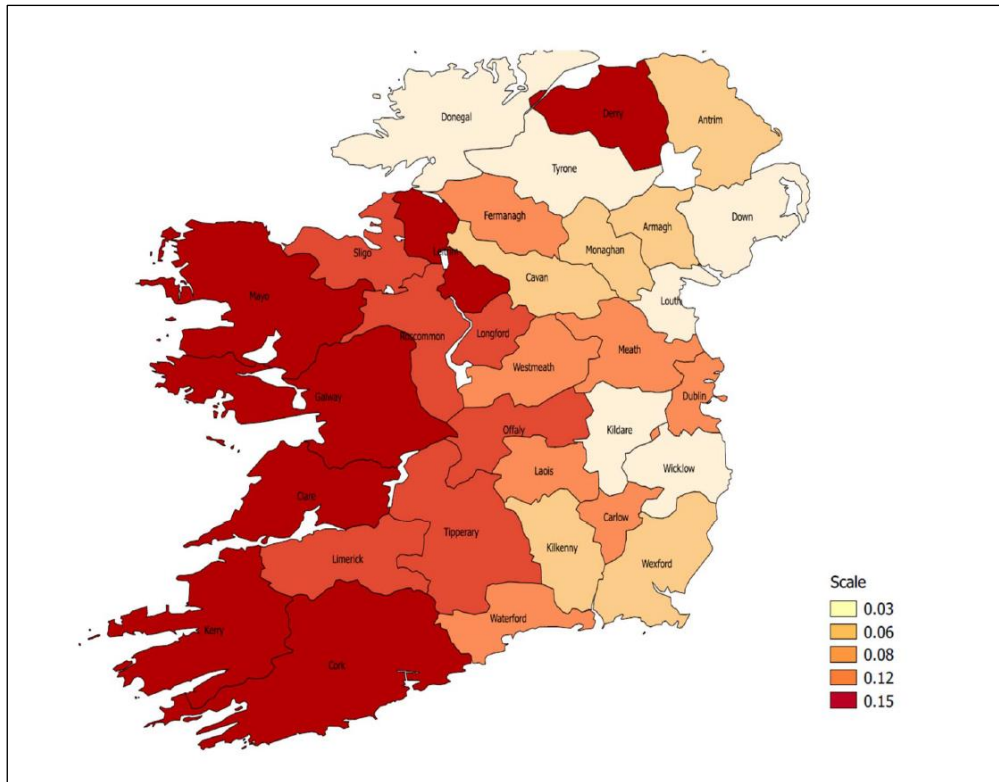
Figure B2: Potato crop failure rates by county.



Note: Potato crop failure rates 1846-1845. Extent of the Famine represented according to a scale ramp. Darker shades represent higher incidence.

Source: Authors' calculations using data described in Section 2. The geographical coordinates of cities and the borders of Irish counties and related Geographical Information System (GIS) data during the 19th and 20th centuries are extracted from the EURATLAS files (Nuessli, 2011).

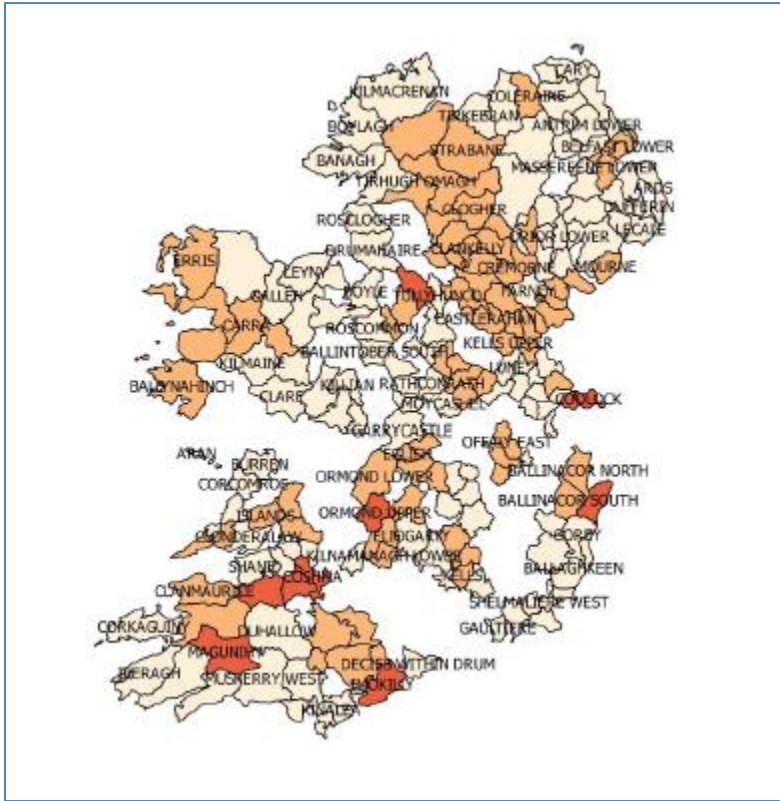
Figure B3: Excess mortality rates by county



Note: Excess Mortality Rate (in %) per thousand individuals by county. Extent of the Great Irish Famine represented according to a five scale interval ramp. Dark shades represent higher incidence.

Source: Authors' calculations using data described in Section 2. The geographical coordinates of cities and the borders of Irish counties and related Geographical Information System (GIS) data during the 19th and 20th centuries are extracted from the EURATLAS files (Nuessli, 2011).

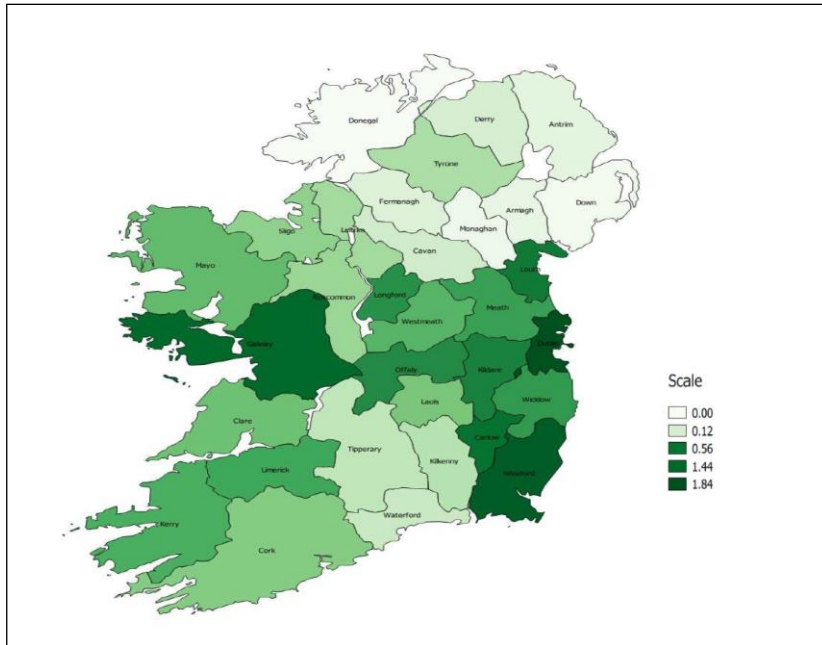
Figure B4: Goodspeed (2016)'s blight indicators by barony



Note: Goodspeed (2016)'s map of blight indicators. Darker shades represent more severe blight.

Source: See Goodspeed (2016) for a full description of the indicator. The geographical coordinates of cities and the borders of Irish counties and related Geographical Information System (GIS) data during the 19th and 20th centuries are extracted from the EURATLAS files (Nuessli, 2011).

Figure B5: Geographical distribution of rebels, by county of birth



Note: Geographical distribution of the ratio of the total number of rebels over total population (per 10,000s) by county of birth, as reported by the 1911 Irish Census. Darker shades represent higher incidence

Source: Authors' calculations using data described in Section 3.

Figure B6: Cumulative distribution of t-statistics, based on Placebo 2

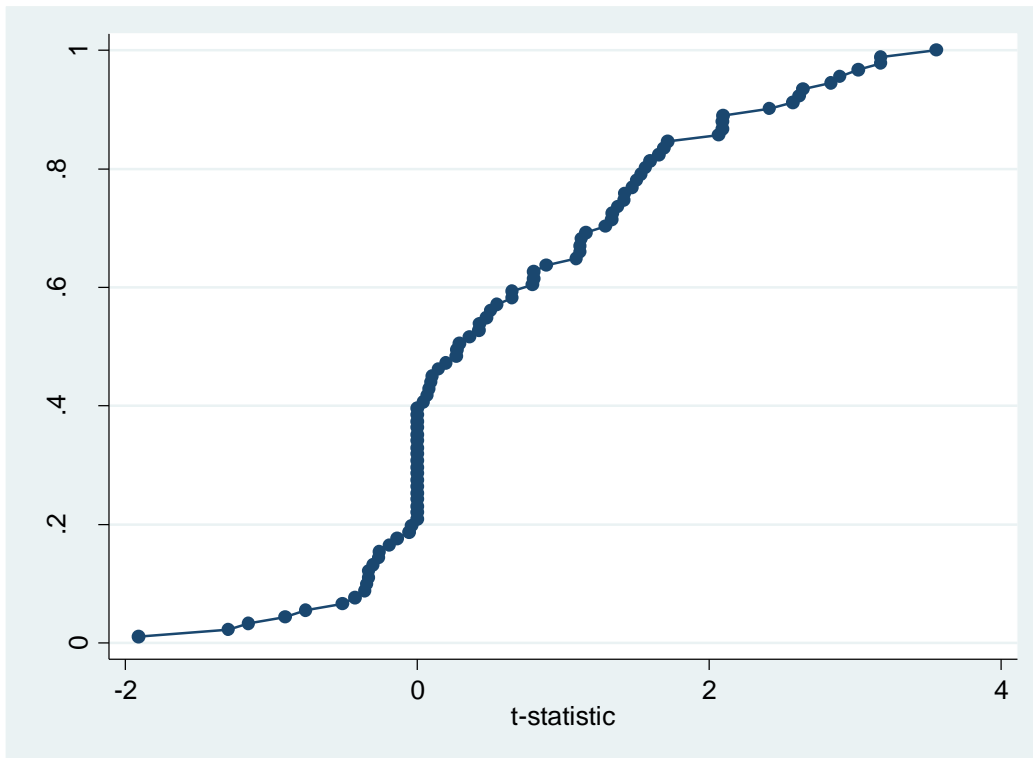
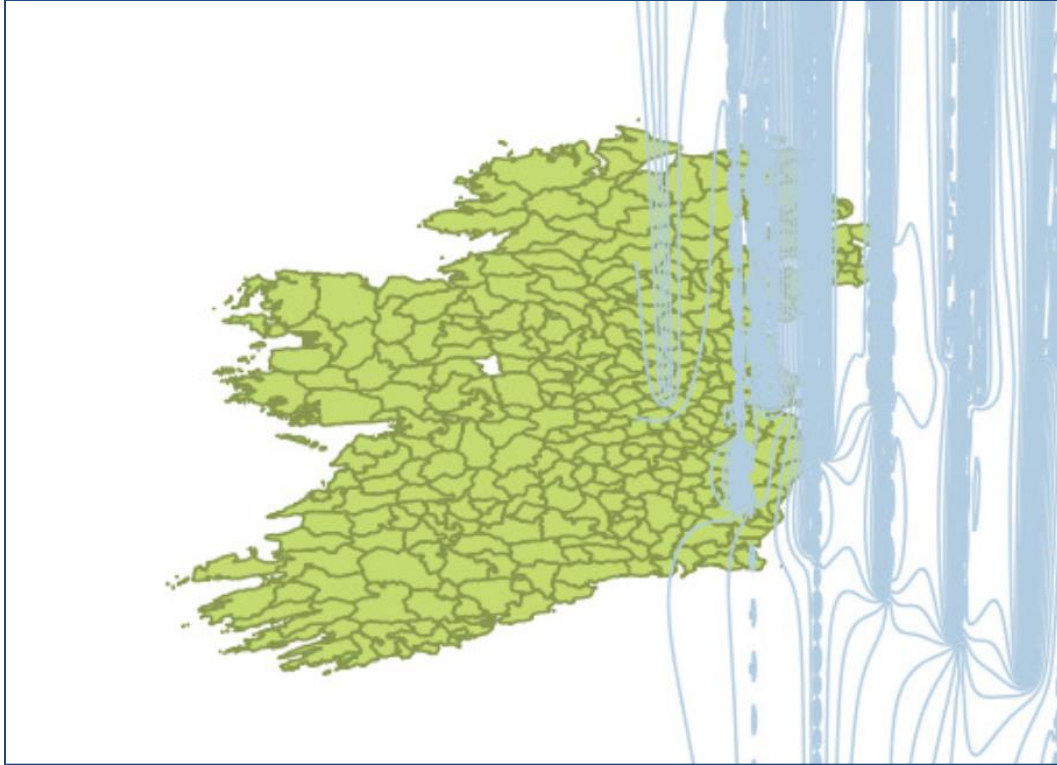


Figure B7: Placebo exercise 2 - Simulation of the potato blight on April 23rd, 1916



Note: The figure displays the results of the simulation of the potato blight dispersion using the wind direction (represented by the blue lines) on Irish island during on April 23rd, 1916, as described in Section 7.

Online Appendix

References

- Abramitzky, R., L. P. Boustan, and K. Eriksson (2012): “Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration,” *American Economic Review*, 102(5), 1832-1856.
- Abramitzky, R., R. Mill, and S. Perez (2021): “Linking Individuals Across Historical Sources: a Fully Automated Approach,” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*.
- Atack, J., F. Bateman, and M. E. Gregson (1992): “Matchmaker, Matchmaker, Make Me a Match A General Personal Computer-Based Matching Program for Historical Research,” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 25(2), 53-65.
- Connell, J. (2015): *Dublin in Rebellion: A Directory 1913-1923*. Lilliput Pr Ltd, Dublin, 2nd edn.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977): “Maximum Likelihood from Incomplete Data via the EM Algorithm.,” *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), 1-38.
- Foster, R. (2015): *Vivid Faces: The Revolutionary Generation in Ireland, 1890-1923*. Penguin, London.
- Jaro, M. A. (1989): “Advances in Record-Linkage Methodology as Applied to Matching the 1995 Census of Tampa, Florida,” *Journal of the American Statistical Association*, 84(406), 414-420.
- Nuessli, C. (2011): *Euratlas*. Euratlas-Nuessli, Yverdon-les-Bains.
- Winkler, W. (2006): “Overview of Record Linkage and Current Research Directions,” Division Research Report Series 2, U.S. Bureau Statistical Research.
- Woulfe, P. (1923): *Sloinnte Gaedheal is Gall. Irish names and surnames*. M. H. Gill, Dublin.