Regular Article

The deep roots of rebellion

Gaia Narciso [a,*], Battista Severgnini [b]

[a] *Department of Economics, Trinity College Dublin, Address: Arts Building, College Green, Dublin 2, Ireland*
[b] *Department of Economics, Copenhagen Business School, Address: Porcelænshaven 16A, DK, 2000, Frederiksberg, Denmark*

A B S T R A C T

This paper analyzes the triggers of rebellion and documents the historical roots of conflict using a unique dataset at the individual level. Drawing on evidence from the Famine (1845–1850) and its effect on the Irish Revolution (1916–1921), we show how negative shocks can explain social unrest in the long run. These findings are confirmed by the analysis of surnames, which enables the study of socio-cultural persistence over time. The instrumental variable analysis based on the wind direction that determined the spread of the potato blight that caused the Famine provides further evidence in support of the legacy of rebellion.

1. Introduction

How do large adverse shocks shape the decision to rebel? Social unrest and civil conflicts are usually studied *ex-post,* making it hard to disentangle the short- and long-run factors that trigger the rebellion decision at the individual level.[1] This paper contributes to the literature by shedding light on the individual determinants of joining a rebellion and on the historical roots of conflict.

We analyze whether past adverse shocks can explain social unrest by drawing on evidence from the Great Irish Famine (1845–1850) and its effect on the Irish Revolution against British rule (1916–1921). We show that individuals whose families had been most affected by the Irish Famine were more likely to participate in the rebellion against British rule during the revolutionary period, over 70 years after the Famine. We rely on two key sources of variation to identify the causal effects of the Famine on participation in the rebellion. Firstly, we provide evidence of the impact of the Famine on insurgency at the individual level by exploiting the geographical and temporal distribution of surnames and surnames' differential exposure to the Famine. Secondly, we implement an instrumental variable analysis based on the wind direction that determined the spread of the potato blight that caused the Famine.

We construct a unique dataset based on several historical sources and proceed in four steps. We start by considering individuals from the totality of the 1911 Irish Census. This dataset is an exceptional source of information at the individual and household level shortly *before* the start of the Irish revolutionary era. Furthermore, we use the lists of rebels, primarily provided by the Irish Military Archives, and match them with the 1911 Irish Census using manual techniques and automated statistical methods. This allows us to investigate the individual characteristics of those who joined the independence movement. Next, we gather a set of measures of the severity of the Famine at the local level. We provide strong evidence of the long-run relationship between the extent of the Famine and the probability of participating in the uprising two generations after the Famine. Inspired by studies on the importance of surnames for studying socio-cultural persistence across generations (*e.g.*, Güell et al., 2014; Clark and Cummins, 2015; and Bleakley and Ferrie, 2016), we exploit variation in the distribution of surnames over space and time to track how rebellion animosity lingered for more than one generation. Finally, due to potential endogeneity issues, we conduct an instrumental variable analysis based on the evidence related to the smooth and isotropic spread of the potato blight in Ireland (Zadoks and Kampmeijer, 1977; and Cavalli-Sforza and Feldman, 1981). Guided by phytopathology studies, we exploit the exogenous spreading of the potato blight and construct a "donut" instrumental variable based on the wind direction that favored the spread of the blight. The instrumental variable analysis and placebo exercises confirm our main findings. To the best of our knowledge, this is the first study that highlights the role of large adverse shocks in explaining social unrest in the long run.

---

* Corresponding author.
E-mail addresses: narcisog@tcd.ie (G. Narciso), bs.eco@cbs.dk (B. Severgnini).
[1] Collier and Hoeffler (2004), Miguel et al. (2004), Acemoglu and Robinson (2006), Gleditsch (2009), Blattman and Miguel (2010), Caselli et al. (2015), Kuran (1991), Cantoni et al. (2019), Costa and Kahn (2003), Krueger (2007, 2015) and Humphreys and Weinstein (2008).

2. Historical background

2.1. The Great Irish Famine (1845–1850)

The Irish Famine, caused by the diffusion of a potato blight, was one of the biggest tragedies of modern history. Over the period 1845–1850, about 1 million people out of a population of 8.5 million died due to starvation and related diseases, while around 1 million emigrated, mainly to North America (Ó Gráda, 1989).

The potato played an essential role in setting living standards in Ireland: over the previous centuries, it became the main staple due to its nutritional content and relative ease of cultivation in the Irish climate (Ó Gráda, 1993; Feehanan, 2012).

A fungus, the *Phytophthora infestans*, caused the potato blight that led to the Great Irish Famine. Originating in Mexico (Goss, 2014), it struck much of Europe (Ó Gráda, 1989 and 1994; Kenealy, 2002). The epidemic was most severe in Ireland due to the widespread planting of potatoes and a favorable and *unusual* combination of weather conditions. Infected potato tubers produce zoospores, which can move through the potato plant transmitting the *Phytophthora infestans* to their foliage (Johnson, 2010). Indeed, the blight can be highly contagious, with estimated 300,000 spores per day produced by each lesion on a potato leaf (Agrios, 2005). Spores can further spread to other plants by water or carried by the wind.

The blight broke out in Ireland in 1845, when the *Phytophthora* destroyed about one-third of the potato crop (Ó Gráda, 2006). The following year was characterized by an almost complete potato crop failure due to an unusually warm winter and autumn, followed by damp summers. In 1847, the extent of the blight was minimal, but due to the limited availability of seed potatoes from the previous year, the total yields were low. The *Phytophthora* hit severely again in 1848 when crops failed almost completely.

Westminster initially provided famine relief under the *Poor Relief (Ireland) Act* of 1838, which had established workhouses for the poor. By 1847, workhouses had reached full capacity, while the number of individuals working in public works swelled from 27,000 in September 1846 to 700,000 in March 1847 (Ó Gráda, 1994; 1999). Eventually, relief in the form of public works provided by Westminster was replaced by soup kitchens: by the summer of 1847, 3 million people received food rations.

Sen (1999) identifies the role of cultural alienation and hostility in his analysis of the Great Irish Famine. Indeed, the consensus among modern historians and critics at the time was that these efforts were insufficient, with "*relief being too little, too slow, too conditional and cut off too soon*" (Ó Gráda, 2009).

2.2. The Irish Revolution (1916–1921)

In 1916, the year of the Easter Rising, Ireland was part of the United Kingdom of Great Britain and Ireland, established with the Act of Union in 1801. On Easter Monday, April 24th, 1916, about 150 armed men gathered in front of the General Post Office in Dublin and seized it. The fighting between rebels and British troops lasted for five days and ended with the insurgents' surrender. The Easter Rising was the first act of what later became the War of Independence against British rule. On three occasions, the Irish Members of Parliament had tried to achieve independence via legal ways in Westminster to guarantee Home Rule for Ireland, *i.e.*, the set-up of a Parliament in Dublin. The first two Home Rule Bills were defeated in Westminster, while the third Home Rule Bill eventually passed in 1914. However, the implementation of the third Home Rule Bill stalled due to the breakout of World War I. The rebellion escalated in 1919, with the Irish rebels conducting ambushes and attacks against British barracks throughout the country (Killeen, 2007). The uprising continued until 1921, when the Anglo-Irish Treaty was signed, splitting the Irish counties between Northern Ireland and the Irish Free State.

Although the impact of the Great Irish Famine was immediate, historical evidence suggests that politically motivated rebellion smoldered under the surface for several years after the Famine. Historians identify two reasons for this delay. First, while riots and increased thefts characterized the initial period of the Famine, prolonged undernutrition and disease caused agitations to cease and resignation to emerge as the Famine loomed on (Kenealy, 2002). Second, changes in law enforcement and institutional settings introduced by the British government and the outbreak of the Great War fueled the spread of rebellion in Ireland (Kenealy, 2002). Interestingly, we find that the number of times the word "*Famine*" appeared in Irish newspapers spiked during the revolutionary period (1916–1921).[2]

3. Data and matching techniques

3.1. Data from the Irish Census, Military Archives

The first data source is the 1911 Irish Census, recently digitized by the Irish National Archives. The Census consists of over 3.9 million observations across the thirty-two Irish counties. Tables B1, B2, and B3 in the Online Appendix display the summary statistics at the individual level and the relative geographical distributions.

The Irish Military Archives provide the second source of data. In 1923, all veterans or widows and children of deceased veterans of the Easter Rising and the War of Independence were granted a pension. We identify the rebels based on the list of pension applicants. In total, 4662 pension applications are available. About 82% of the applications were confirmed; overall, 3816 rebels (or their next of kin) were granted a pension. In addition, we complement veterans' names with secondary sources of information on Irish rebels (Connell, 2015; Foster, 2015).

3.2. Matching

We match the overall list of veterans with the 1911 Irish Census. Our matching relies on two linking strategies: manual and automated matching. Section A in the Online Appendix describes the two methodologies implemented.

3.2.1. Manual matching
We manually link the list of rebels by exploiting historical sources available from the Irish Military Archives and on the Internet.[3] Inspired by Ferrie (1996), manual linking is based on two main principles: complete name (first and last name) and age. Given the evidence on age rounding on census forms, we allow for up to a 2-year discrepancy around the age reported in the Census. The manual matching is conducted on individuals who were 16 or older at the time of the 1911 Census.

3.2.2. Automated matching
The second matching strategy is based on the automated statistical method proposed by Abramitzky et al. (2018) and Abramitzky et al. (2021).

3.2.3. Rebel definition
Each matching technique produces two indicators, two of which are more conservative while the other two are less conservative. We pool the four indicators and create a dummy variable *Rebel*, which takes the value 1 if individual *i* in the 1911 Census is labeled as rebel according to at least one of the 4 indicators and 0 otherwise. We identify 1491 rebels, achieving a matching rate of about 24% with the Military Archives list.[4]

---

[2] See Figure B1 in the Online Appendix.
[3] http://www.irishmedals.ie and http://www.irishvolunteers.ie.
[4] The matching rate is in line with the literature. See for example, Long and Ferrie (2013).

The two matching techniques identify two different sets of insurgents. On average, the manual technique is better at identifying the more historically known rebels or those with less common last names. The manual method also allows us to identify women by searching for the maiden name of female insurgents in the 1911 Census. Due to the frequent recurrence of particular surnames and first names, manual matching is less effective at matching more common names with a similar birth year. In these cases, an overall conservative approach was adopted, excluding these observations and thus leading to a relatively smaller number of veterans with common last names to be manually matched with the 1911 Census. The automated technique performs better in matching the more common names since it allows matching to a greater section of the Irish Census. We consider the two matching methods complements rather than substitutes in categorizing individuals in the Census. However, as a robustness check, we also present the results for the four different rebellion indicators separately.

Table B4 in the Online Appendix provides the balance tests comparing the rebels matched to the 1911 Census with respect to those not linked with the Census. There is no statistically significant difference in terms of gender between the matched and unmatched list of rebels. Average matched rebels are older than unmatched rebels. However, this difference may arise from the restrictions imposed on the manual matching, which focuses on individuals above the age of 16 in 1911.[5]

3.3. Other historical sources

We consider three additional administrative sources to study Irish family names' distribution over time. The first source is the *Down Survey*, the first historical detailed land survey on a national scale, which contains the distribution of family names in all townlands from 1656 to 1658. The second source is a survey for determining poverty status in Ireland, known as *Griffith's Valuation*,[6] which is the only genealogical information available in Ireland before the 20th century. Carried out between 1848 and 1867, it provides a representative picture of Irish society. The third source is the 1901 Irish Census, which is available in its entirety.

3.4. Measuring the famine and other covariates

We construct three main measures of the Famine.

3.4.1. Potato crop failure

We measure the extent of the Famine in terms of potato crop failure rate between 1845 and 1846 at the county level. Information on potato production at the county level relies on the statistical data provided by Bourke (1959).

3.4.2. Excess mortality

We measure the extent of the Famine in terms of the excess mortality rate per thousand inhabitants at the county level between 1846 and 1850, based on the data provided by Cousens (1960) and calculated as the ratio of excess deaths at the county level during the Famine years and county population in 1841.

3.4.3. Blight indicator

We use the information on the severity of the blight infection at the barony level, based on the work by Goodspeed (2016). The *blight indicator* is based on textual analysis of the reports of the Parliamentary Relief Commission at the time of the Famine. We construct a dummy

variable that takes the value of 1 if the Famine is reported to be severe and 0 otherwise.[7]

Many authors have investigated the impact of the Famine on Irish emigration, particularly to the United States (Ó Grada and O'Rourke, 1997; Narciso et al., 2020). A complementary study (Narciso et al., 2020) shows that the Irish Famine was a significant long-run driver of individuals' migration choices. For the purpose of this analysis, we collect data on the emigration ratio at the county of birth, measured as the change in population between 1852 and 1851 over population in 1841, as reported by the Irish Emigration Database (*Emigration rate COB*).

Finally, data on daily wind direction used for constructing the instrumental variable in both the IV analysis and placebo exercises are taken from Delaygue et al. (2019), based on the original work by Bourke and Hubert (1993).[8]

4. Estimation strategy

We investigate the determinants of participating in the rebellion and the role of the Famine on the probability of becoming a rebel by estimating the following equation:

$$Rebel_{icd} = \alpha + \beta Famine_c + \gamma X_{icd} + \vartheta C_c + \delta Z_d + \varphi_d + \varepsilon_{icd} \tag{1}$$

The dependent variable, $Rebel_{icd}$ is an indicator variable, which takes the value 1 if individual i, born in location c, living in district electoral division d, takes part in rebellion activities and 0 otherwise. We control for a set of individual characteristics, X_{icd}, such as age, gender, literacy, occupational dummies, being Catholic, marital status, and household size. The variable $Famine_c$ measures the extent of the Famine in the location of birth c. We use three different measures to capture the extent of the Famine: (1) potato crop failure rate between 1845 and 1846 measured at the county of birth level; (2) excess mortality rate per thousand individuals at the county of birth level; (3) a blight indicatory at the barony of birth level. We control for a set of variables at the location of origin level, C_c, *i.e.*, emigration rates between 1851 and 1852 and literacy rate in 1911 at the county of birth level. We include a set of residence characteristics as of 1911, Z_d, namely the share of men, the share of Catholics, the share of individuals aged between 25 and 40 years old, and adult literacy rate at the electoral district level.[9] Some specifications include the district of residence fixed effects (φ_d). Standard errors are clustered at the location of birth level, *i.e.*, at the county or the barony of birth, depending on the measure of the Famine used. Given the relatively small number of clusters at county level, we also include clustered standard errors using the wild cluster bootstrap technique (Cameron and Miller, 2015) based on 10,000 replications.[10]

5. Baseline estimation results

Table 1 presents the linear probability model estimation results of the baseline specification. We restrict our analysis to the sample of individuals over 10 years old (at least 16 during the Irish Revolution) and

---

[5] Due to data limitations, age and gender are the only variables available for the comparison. Figure B2 in the Online Appendix presents the geographical distributions of rebels, by county of birth.

[6] See http://downsurvey.tcd.ie/ and http://www.askaboutireland.ie/griffith-valuation/.

[7] See Goodspeed (2016) for more information on the construction of this variable. Figures B3, B4 and B5 in the Online Appendix present the geographical distribution of the Famine according to these three measures.

[8] See Table B5 in the Online Appendix.

[9] Adult literacy rate is defined as the percentage of individuals over 15 who are literate, *i.e.*, who can read or read and write.

[10] We use the Stata command *bootest* from Roodman et al. (2019).

Table 1
Baseline specification.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	*Dependent Variable: Rebel*					
Crop failure rate	0.00033	0.00033				
(County of birth)	(0.000)[a]	(0.000)[a]				
	[0.0024]	[0.0014]				
Excess Mortality			0.00329	0.00335		
(County of birth)			(0.001)[b]	(0.001)[b]		
			[0.0425]	[0.0503]		
Blight indicator					0.00012	0.00016
(Barony of birth)					(0.000)[c]	(0.000)[b]
Individual characteristics						
Age	−0.00002	−0.00002	−0.00002	−0.00002	−0.00001	−0.00001
	(0.000)[a]	(0.000)[a]	(0.000)[a]	(0.000)[a]	(0.000)[a]	(0.000)[a]
	[0.0000]	[0.0000]	[0.0000]	[0.0000]		
Female	−0.00085	−0.00084	−0.00085	−0.00084	−0.00046	−0.00046
	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[a]	(0.000)[a]
	[0.0000]	[0.0000]	[0.0000]	[0.0000]		
Literate	0.00010	0.00011	0.00010	0.00011	0.00006	0.00006
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	[0.3691]	[0.3251]	[0.3640]	[0.3107]		
Catholic	0.00036	0.00036	0.00036	0.00036	0.00018	0.00019
	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[a]	(0.000)[a]
	[0.0000]	[0.0000]	[0.0000]	[0.0000]		
Married	−0.00013	−0.00015	−0.00013	−0.00015	−0.00006	−0.00006
	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[a]	(0.000)[a]
	[0.0013]	[0.0007]	[0.0021]	[0.0013]		
Household size	0.00002	0.00002	0.00002	0.00002	0.00001	0.00000
	(0.000)[b]	(0.000)[c]	(0.000)[b]	(0.000)[c]	(0.000)	(0.000)
	[0.0104]	[0.0271]	[0.0086]	[0.0290]		
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
Occupations	Yes	Yes	Yes	Yes	Yes	Yes
County of birth controls	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

Note: 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. *Crop failure rate, Excess Mortality,* and *Blight Indicator* are the three measures of the Famine severity, as described in Section 3. County of birth controls include: emigration rates between 1851 and 1852 and literacy rate in 1911. Location characteristics include: share of men, share of Catholics, share of individuals aged between 25 and 40 years old, adult literacy rate at the electoral district level of residence.

[a] $p < 0.01$.
[b] $p < 0.05$.
[c] $p < 0.1$.

under 65 at the time of the 1911 Census.[11]

As detailed in Section 4, we consider the county of birth of each individual in the 1911 Census and match it with our measure of the potato crop failure rate as a way of measuring the severity of the Great Famine (Table 1, columns (1) and (2)). The estimated coefficient of crop failure is positive and statistically significant at the 1% level. A one percent increase in the potato crop failure rate during the Famine in the county of birth raises the probability of becoming a rebel by 0.033 percentage points. The results are robust to the inclusion of individual, occupation, and location characteristics or district of residence fixed effects (column (2)).[12]

Columns (3) and (4) present the estimation results using the second measure of the Famine, namely excess mortality at the county of birth level. In line with the previous results, we find that the higher the excess mortality, the higher the probability of becoming a rebel. A one percent

increase in excess mortality is associated with about a 0.33 percentage point increase in the likelihood of becoming an insurgent.

Our results are also robust when considering a measure of the Famine at a more disaggregated level, *i.e.*, barony level. The analysis at the barony of origin allows for a greater geographical variation in the analysis of the effect of the Famine on the probability of rebelling. Two issues arise in relation to the use of more geographically disaggregated measures of the Famine. First, the townland of origin is available in the 1911 Census for 40% of the census respondents in our sample, as the 1911 Census questionnaire asks about the county of birth or city of birth. Second, in some instances, it may be difficult to distinguish between the county name and the name of the main city of that specific county (*e.g.*, Galway city and county Galway). The independent variable, described in Section 3, is the blight indicator introduced by Goodspeed (2016). Standard errors are clustered at the barony level. Given the number of baronies (365 in total), it is unnecessary to report the bootstrapped standard errors in columns (5) and (6). The blight indicator is positive and statistically significant at the 5% level when the electoral district of residence fixed effects are included (column (6)).

Regarding rebels' characteristics, we find that insurgents are more

---

[11] We check whether spatial autocorrelation might lead to larger t-statistics in our regression The z-score of the Moran index based on OLS residuals of column (3) of Table 1 and aggregated at the county level is 1.57, suggesting that the standard errors are not spatially autocorrelated. Furthermore, the results of the regression analysis allowing for arbitrary spatial clustering (Colella et al., 2019) fully confirm the findings presented in Table 1. Results are available upon request.

[12] The mean crop failure rate is 24% in the sample.

likely to be younger, Catholic, and male. Similarly, household size is positively related to the probability of being a rebel.[13] We also find a statistically significant relationship between marital status and being an insurgent. The specification also controls for occupations and the role of peer effects in influencing the decision to participate in the rebellion. We include a set of variables at the district electoral division level of residence according to the 1911 Census (*i.e.*, the share of men in the district, the share of Catholics, the share of individuals aged 25–40, and adult literacy rate) and the level of development in the county of birth, as measured by the literacy rate in 1911 and the extent of out-migration. The specification includes county of residence indicators (columns (1), (3), (5)) or district of residence fixed effects (columns (2), (4), (6)).[14]

The construction of the *Rebel* indicator involves two different matching strategies (manual and automated linking). Table B8 in the Online Appendix presents the estimation results for each of the four indicators arising from the manual and automated matching. We also propose an alternative dependent variable, which captures any pension recipient (rather than pension applicant) that we are able to identify. The new dependent variable takes the value 1 if individual *i* is identified as a rebel according to at least one of the four measures and if the individual is granted a pension and 0 otherwise. The estimation results provide evidence in support of a strong, positive association between the Famine and the probability of rebelling, across the three measures of the Famine and the four disaggregated rebel indicators.

6. Family exposure to the famine

This Section investigates the extent of family exposure to the Famine. Due to data availability, it is not possible to construct a complete and representative map of genealogical trees of Irish families. As an alternative, we exploit the distribution of individuals' surnames across different years.[15] We expect that family names more exposed to the Famine would be positively related to the likelihood of rebelling.

First, we explore the extent of internal migration within Ireland at the turn of the 20th century. We use the entirety of the 1901 and 1911 Censuses and collect the family names reported in each of these Census rounds. To have homogeneity in family names across different data sources, we translate Irish surnames into English and consider the phonetic version of each last name.[16] Next, we construct an index measuring the geographical distribution of each surname across the two Census rounds. The surname index is similar to the one suggested by Fryer and Levitt (2004) and is measured as:

$$Surname\ Index_{ict} = \frac{\Pr(Surname_i | c, t)}{\Pr(Surname_i | c, t) + \Pr(Surname_i | Ireland - c, t)} \quad (2)$$

For each surname *i* in county *c* at time *t* (t = 1901, 1911), we calculate the ratio between the probability that surname *i* appears in county *c* at time *t*, over the sum of the probabilities that it appears in county *c* and in the rest of the island at time *t*.

We calculate the correlation coefficient between the surname indices over time to measure geographical mobility across the two Censuses. We find that surnames across the two Censuses have a correlation coefficient of 0.96, which suggests low geographical migration within Ireland at the turn of the 20th century and attenuates the possibility of endogenous location choices.

As a second step, we exploit the distribution of surnames considering

two complementary strategies: the first one takes into consideration the geographical distribution of family names since the 17th century; the second strategy is based on a two-stage OLS procedure that exploits the predictive power of surnames (Bleakley and Ferrie, 2016).

6.1. Distribution of last names

The wealth of historical data allows us to measure the exposure of each family name to the Famine and exploit the geographical variation of surnames over time, using a methodology similar to Fryer and Levitt (2004). We compute two Surname Indices (as in (2)) based on two historical data sources, the *Down Survey* (1656–1658) and the *Griffith's Valuation*. For each surname available, we compute the difference between these two indices. This difference can be interpreted as a proxy of the extent of mortality *at surname level* during the Famine period. Next, we match each surname available in the 1911 Census with this proxy of Famine exposure. We estimate a new specification, including the extent of mortality during the Famine at the surname level as the primary independent variable, and test whether individuals whose surname was more exposed to the Famine are more likely to become rebels. Table 2 presents the results of this estimation. The estimated coefficient of surname exposure to the Famine is positive and statistically significant at the 5% level when the entire sample is considered (columns (1) and (2)). Given the patrilineal transmission of last names, columns (3) and (4) present the estimation results for the sample of men. The estimated coefficient of surname exposure to the Famine is positive and statistically significant at the 1% level when the district of residence fixed effects are included in the analysis (column (4)). Based on this exercise, we can conclude that individuals whose family name was more exposed to the Famine two generations earlier are more likely to have participated in the insurgence.

6.2. Predicative power of last names

As a second step in the analysis of family exposure to the Famine, we implement the estimation framework suggested by Bleakley and Ferrie (2016). In the first step, we run an OLS regression of an equation where the dependent variable is our measure of the extent of the Famine (*i.e.*, potato crop failure rate) and the regressors are last name fixed effects. The second stage is a modified version of equation (1), where we include the predictive power of surnames ($\widehat{Famine}_c^{surname}$) and the estimated errors ($\widehat{Famine}_c^{residual}$) from the first stage, as shown in equation (3) below[17]

$$Rebel_{icd} = \alpha + \beta_1 \widehat{Famine}_c^{surname} + \beta_2 \widehat{Famine}_c^{residual} + \gamma X_i + \vartheta C_c + \delta Z_d + \varepsilon_{icd} \quad (3)$$

The results of the estimation of equation (3) are reported in columns (5) to (8) of Table 2. We would expect the estimated coefficient of the predicted value of the Famine at surname level ($\widehat{\beta_1}$) to be positive and statistically significant. This is indeed what we find, both in the full sample (columns (5) and (6)) and in the sample of men only (columns (7) and (8)).

The scope of the two exercises is to explore the extent of family exposure to the Famine and its effect on the probability of rebelling. Both the proxy of family exposure to the Famine (columns (1)–(4)) and the predicted value of the Famine based on last names (columns (5)–(8)) indicate a positive and statistically significant relationship with the probability of taking part in the rebellion, thus providing further

---

[13] Household size is measured as the number of individuals above the age of 10 and below the age of 65 living at the same address.

[14] There are 4 provinces and 3329 district electoral divisions in the 1911 Irish Census. Table B6 in the Online Appendix presents the full set of estimates, Table B7 includes further controls.

[15] Güell et al. (2014); Clark and Cummins (2015); and Bleakley and Ferrie (2016).

[16] We use the Stata command *soundex*.

[17] The estimated errors are included to control for the residual effects.

Table 2
Family exposure to the Famine.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable: Rebel							
Exposure to the Famine	0.00019	0.00021	0.00049	0.00051				
at surname level	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[a]				
	[0.0798]	[0.0452]	[0.0349]	[0.0337]				
$\widehat{Famine}_c^{surname}$					0.00404	0.00507	0.00567	0.00765
					(0.002)[c]	(0.002)[b]	(0.003)[c]	(0.004)[c]
					[0.1173]	[0.0677]	[0.1247]	[0.0833]
$\widehat{Famine}_c^{residual}$					0.00178	0.00328	0.00169	0.00505
					(0.003)	(0.003)	(0.005)	(0.005)
					[0.5985]	[0.3215]	[0.6303]	[0.3823]
Sample	All	All	Men	Men	All	All	Men	Men
Observations	872,056	872,056	441,435	441,435	872,056	872,056	441,435	441,435
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupations	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County of birth controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Characteristics	Yes	No	Yes	No	Yes	No	Yes	No
County of residence FE	Yes	No	Yes	No	Yes	No	Yes	No
District of residence FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: 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). The variable "*Exposure to the Famine at surname level*" measures the geographical variation of last names over time, as described in Section 6.1. The variable $\widehat{Famine}_c^{surname}$ measures the predicted power of surnames, as described in Section 6.2. All regressions include a constant. Individual characteristics include: age, gender, literacy, occupational dummies, being Catholic, marital status, and household size. County of birth controls include: emigration rates between 1851 and 1852 and literacy rate in 1911. Location characteristics include: share of men, share of Catholics, share of individuals aged between 25 and 40 years old, adult literacy rate at the electoral district level.

[a] $p < 0.01$.
[b] $p < 0.05$.
[c] $p < 0.1$.

evidence in support of the long-run impact of the Famine.[18]

7. Instrumental variable analysis and placebo tests

7.1. Instrumental variable analysis

A concern may arise in relation to a potential bias in our econometric estimation. First, some of the regressors are based on historical data, which might suffer from measurement error. Second, the results could be biased by potential confounding factors positively related to the extent of the Famine and the likelihood of joining the Irish rebellion movement.

To mitigate these potential issues, we construct an instrumental variable based on the specific characteristics of the diffusion of the potato blight that caused the Famine. The spread of the blight was favored by the relatively flat terrain on the island and driven by factors independent from human action (Zadoks and Kampmeijer, 1977; Cavalli-Sforza and Feldman, 1981).

We rely on the work conducted in the field of phytopathology (Van der Plank, 1963 and Kushalappa and Ludwig, 1982), which quantifies the potato blight infection rate as a gradient of the number of infected plants in a particular area compared to the infection rate in the nearest surroundings. We construct a measure of the severity of the potato blight, R_{iJ}, for barony *i* as the difference between a measure of barony *i*'s soil suitability for potato cultivation and the average soil suitability for potato cultivation of a set of selected nearby baronies *J, i.e.*

$$R_{iJ} = ln\frac{FAOi}{1 - FAOi} - \frac{1}{J}\sum_{j \in J} ln\frac{FAOj}{1 - FAOj} \qquad (4)$$

where FAO$_i$ is an index of soil suitability for potato cultivation for barony *i*, according to the FAO database on Global Agro-Ecological Zones. The index measures potential potato cultivation based on information on both climatic and land characteristics at $0.5 \times 0.5°$ cells (about 56×56 km), whereby a higher index represents greater potato suitability. The index FAO$_j$ measures land suitability of nearby barony *j*, selected on the basis of an "*eaten donut*" strategy.[19] For each barony *i*, the set *J* of nearby baronies in equation (4) does not include the immediate neighbors of barony *i*, but only comprises the neighbors of neighbors, taking into account *wind direction* at the time the blight arrived in Ireland in August–September 1845.

Fig. 1 provides an example of the construction of this instrument for a representative barony, Athlone. First, we identify the baronies sharing a common border with the barony of interest (Ballintober South, Ratheline, Killkenny West, Brawny, Clononan, Garrycastle, Moycanm Clionomacon, and Killinie). Second, we select only the neighboring baronies that have had a direct effect on the transmission of the potato blight following the wind direction in August–September 1845 (Killian, Ballintober South, Rathcline, and Kilkenny West). Finally, the value of the severity of the potato blight for Athlone is computed as the unweighted average of the infection rate of the baronies, which had a direct impact on Killian, Ballintober South, Ratheline, and Kilkenny West (in yellow, Ballymoe, Kilconnell, Moydow, Rathconrath, Roscommon, and Shrule). According to the "*eaten donut*" procedure, the instrument is computed as the average corresponding infection rate of the relevant neighbors of neighbors, where relevance is assessed based on *wind direction*. The rationale for constructing the instrumental variable using a donut procedure rather than adjacent baronies is to reduce the potential bias that may arise if a barony experiencing more severe blight is different from a barony with milder potato blight.

The instrument's construction relies on barometric data from Delaygue et al. (2019), based on the seminal work by Bourke and Hubert (1993). The dataset provides information on *wind direction* per day for

---

[18] Our results should be interpreted as the propensity to participate in the rebellion of individuals and their descendants, who did neither migrate abroad nor died because of the Famine. Due to data availability, it is not possible to quantify a potential bias in sample selection.

[19] This estimation strategy has similarities with the leave-one-out-jacknife IV approach (Angrist et al., 1999) and to the examiner design by Dobbie et al. (2018).
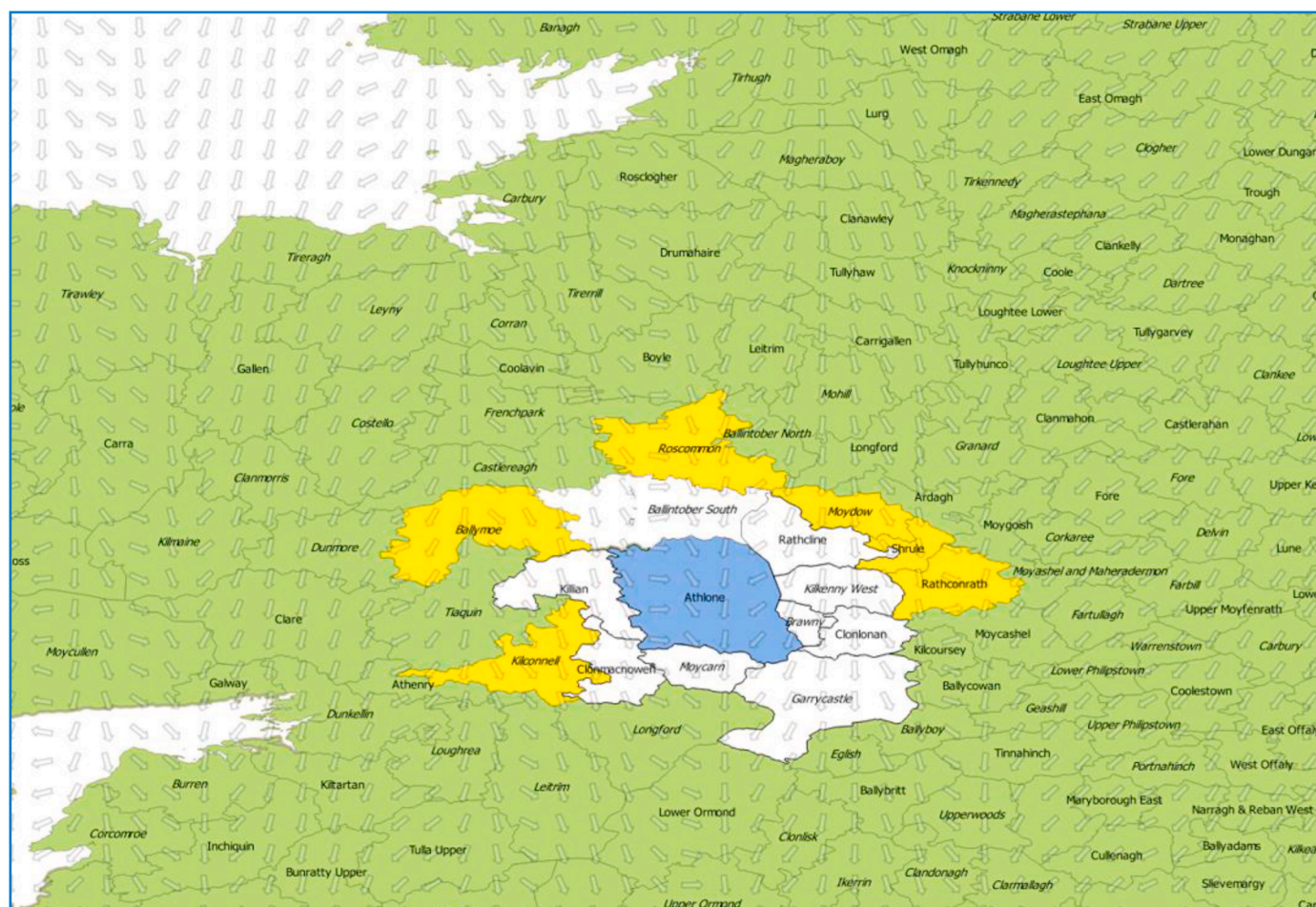
Fig. 1. Construction of the 'donut' for the barony of Athlone. *Note*: The figure provides a visual example of the construction of the "eaten donut" for the barony of Athlone.

the entire Irish island. We use this information on wind bearing to simulate the daily dynamics dispersion of the blight, starting from the date of the first reporting of the arrival of the potato blight (September 6th, 1845, Bourke (1964)) and with an estimated travelling speed of about 50 km per day. [20] [21]

Table 3 reports the IV estimation results. The dependent variable is the rebel indicator at the individual level. Columns (1) to (6) report all specifications presented in the paper. The estimated coefficient of the blight indicator is positive and statistically significant in the basic specification (Panel B, columns (1) and (2)), when controlling for location characteristics (columns (3) and (4)) and when controlling for concurrent factors (columns (5) and (6)). The findings from the instrumental variable analysis are in line with the ones presented in the previous tables, providing further support to the long-run legacy of the Famine on the probability of rebelling.

7.2. Placebo tests

According to the "*eaten donut*" strategy, our instrumental variable consists of two components: land suitability of nearby baronies and wind direction at the time of the potato blight arrival in Ireland

(August–September 1845). To test the validity of our instrument, we conduct three placebo exercises: 1) we consider a set of placebo instrumental variables based on the wind direction calculated in August–September of each year between 1851 and 1950, *i.e.* the century *following* the Famine; 2) we use the wind direction the day *before* the first act of rebellion in 1916 and look at how the blight spores would have spread according to the wind direction recorded on April 23rd, 1916 (the day before the Easter Rising); 3) we construct an alternative placebo instrument, which considers land suitability of the baronies excluded from the "*eaten donut*". We expect that the blight indicator, instrumented for either the ex-post wind direction (placebos 1 and 2) or the less relevant neighboring counties (placebo 3) should not have a statistically significant impact on the probability of becoming a rebel.

7.2.1. Placebo 1: wind direction 1851–1950

Using the same methodological sources and the software code employed to construct our instrument, we construct a set of indicators R_{iJt} based on wind direction in August–September, the same months we consider for our instrument, in each of the years between 1851 and 1950, *i.e.* the 100 years *after* the Famine.

$$R_{iJt} = ln \frac{FAOi}{1 - FAOi} - \frac{1}{J}\sum_{j \in J} ln \frac{FAOjt}{1 - FAOjt} \quad \text{with } t = 1851, .., 1950 \quad (5)$$

where Jt indicates the set of relevant counties based on wind direction measured in August–September of year t. Figure B6 in the Online Appendix shows the empirical cumulative distribution function of the t-statistics obtained by the IV estimates using each available placebo

---

[20] Summary statistics are reported in Table B5 in the Online Appendix.
[21] As additional robustness check, we collect the wind directions reported on Irish newspapers available at the British Newspaper Archive, over the same time period, http://www.britishnewspaperarchive.co.uk. The observations match with our simulation.

Table 3

IV analysis.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: First Stage						
			Blight Indicator			
Potato Blight	1.28617	1.27094	1.59673	1.58093	1.60826	1.56844
Infection rate	(0.289)[a]	(0.306)[a]	(0.478)[a]	(0.495)[a]	(0.551)[a]	(0.561)[a]
Panel B: Second stage						
			Rebel			
Blight Indicator	0.00057	0.00058	0.00039	0.00035	0.00038	0.00036
	(0.000)[a]	(0.000)[a]	(0.000)[b]	(0.000)[b]	(0.000)[b]	(0.000)[b]
Observations	364,906	364,487	364,906	364,487	364,906	364,487
Individual Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Occupations	Yes	Yes	Yes	Yes	Yes	Yes
County of birth controls	No	No	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	No	Yes	No	Yes	No
District of residence FE	No	Yes	No	Yes	No	Yes

Note: Excluded instrument: Potato Blight Infection rate, as described in (4). The first stage is presented in Panel A. Standard errors are clustered at barony level. Individual characteristics include: age, gender, literacy, occupational dummies, being Catholic, marital status, and household size. County of birth controls include: emigration rates between 1851 and 1852 and literacy rate in 1911. Location characteristics include: share of men, share of Catholics, share of individuals aged between 25 and 40 years old, adult literacy rate at the electoral district level. Concurrent factors include: number of GAA clubs and the number of rebellion episodes per county per 1000 during the 1798 Revolution (see the Online Appendix for a discussion of the concurrent factors).

[a] $p < 0.01$.
[b] $p < 0.05$.

indicator, one for each available year. Only in 8.8% of cases, the blight indicator is statistically significant at the 1% level, with a t-statistics higher than 2.58.

7.2.2. Placebo 2: wind direction on April 23rd, 1916

The second placebo exercise employs a specific date to measure wind direction, *i.e.* the day before the first act of rebellion (Easter Rising). We construct our instrument based on how the blight spores would have spread according to the wind direction recorded on April 23rd, 2016.[22] In this case, the blight would not have covered a large part of the Irish territory, confirming the precision of our simulation using wind direction in August–September 1845 and the external validity of our instrument.

7.2.3. Placebo 3: excluded baronies

Finally, we construct an alternative placebo instrument, which considers land suitability of the baronies excluded from the "*eaten donut*". For example, for the barony of Athlone (Fig. 1), the baronies to be considered in this placebo instrument are Castlereagh, Taquim, Longford, Lower Ormond, Eglish, Ballyboy, Ballycowen, Kilcoursey and Moycashel. We would not expect a) this placebo instrument to be correlated with the blight indicator and b) the placebo instrumental variable analysis to produce statistically significant estimated coefficients. The results reported in Table B9 in the Online Appendix confirm our priors.

Overall, the three different placebo exercises provide support of our instrumental variable analysis.

8. Conclusions

This paper studies the triggers of insurgency at the individual level and explores the long-run effect of large adverse shocks on the probability of rebelling. Contributions in social sciences have remarked on the impact of conflict on the long-run growth of countries and on the need to understand its causes. Our original contribution exploits the information from a unique Irish historical dataset during the first two decades of the 20th century. We test whether radical historical events affect the

decision to participate in rebellions, drawing evidence from the effect of the Famine on the Irish independence movement during the Irish Revolutionary period (1916–1921). Robustness checks related to the distribution of family names and instrumental variable regressions, based on specific barometric conditions that favored the dispersion of the blight, confirm our results. Our analysis provides evidence in support of the legacy of uprising and shows how adverse shocks can affect the probability of joining an insurgency in the long-run.

---

[22] Figure B7 in the Online Appendix shows the geographical representation of our simulation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jdeveco.2022.102952.

References

Abramitzky, Ran, Mill, Roy, Perez, Santiago, 2018. Linking individuals across historical sources: A fully automated approach. Historical Methods: A Journal of Quantitative and Interdisciplinary History 53 (2), 94–111. https://doi.org/10.1080/01615440.2018.1543034.

Abramitzky, Ran, Boustan, Leah, Eriksson, Katherine, Feigenbaum, James, Pérez, Santiago, 2021. Automated Linking of Historical Data. Journal of Economic Literature 59 (3), 865–918.

Acemoglu, D., Robinson, J.A., 2006. Economic Origins of Dictatorship and Democracy. Cambridge University Press, Cambridge.

Agrios, G., 2005. Plant Pathology, 5 edn. Academic Pres, San Diego, CA.

Angrist, J.D., Imbens, G.W., Krueger, A.B., 1999. Jackknife instrumental variables estimation. J. Appl. Econom. 14 (1), 57–67.

Blattman, C., Miguel, E., 2010. Civil war. J. Econ. Lit. 48 (1), 3–57.

Bleakley, H., Ferrie, J., 2016. Shocking behavior: random wealth in Antebellum Georgia and human capital across generations. Q. J. Econ. 131, 1455–1495, 1(3).

Bourke, P.M.A., 1959. The extent of the potato crop in Ireland at the time of the Famine. J. Stat. Soc. Inq. Soc. Ireland XX, 1–26.

Bourke, P.M.A., 1964. Emergence of potato blight, 1843-46. Nature (4947), 805–808.

Bourke, P.M.A., Hubert, L., 1993. The Spread of Potato Blight in Europe in 1845-6 and the Accompanying Wind and Weather Patterns. Irish Meteorological Service, Dublin.

Cameron, A.C., Miller, D.L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50 (2), 317–372.

Cantoni, D., Yang, D.Y., Yutchman, N., Zhang, Y.J., 2019. Are protests games of strategic complements or substitutes? Experimental evidence from Hong Kong's democracy movement. Q. J. Econ. 134 (2), 1021–1077. May 2019.

Caselli, F., Morelli, M., Rohner, D., 2015. The geography of interstate resource wars. Q. J. Econ. 130 (1), 267–315.

Cavalli-Sforza, L., Feldman, M., 1981. Cultural Transmission and Evolution: a Quantitative Approach. Princeton University Press, Princeton, NJ.

Clark, G., Cummins, N., 2015. Intergenerational wealth mobility in England, 1858–2012: surnames and social mobility. Econ. J. 125 (582), 61–85.

Colella, F., Lalive, R., Sakalli, S.O., Thoenig, M., 2019. Inference with Arbitrary Clustering," IZA Discussion Papers 12584. Institute of Labor Economics (IZA).

Collier, P., Hoeffler, A., 2004. Greed and grievance in civil wars. Oxf. Econ. Pap. 56, 663-595.

Connell, J., 2015. Dublin in Rebellion: A Directory 1913-1923, second ed. Lilliput Pr Ltd, Dublin.

Costa, D.L., Kahn, M.E., 2003. Cowards and heroes: group loyalty in the American civil war. Q. J. Econ. 118 (2), 519–548.

Cousens, S.H., 1960. Regional death rates in Ireland during the Great famine, from 1846 to 1851. Popul. Stud. 14, 55–74.

Delaygue, G., Brönnimann, S., Jones, P.D., Blanchet, J., Schwander, M., 2019. Reconstruction of Lamb weather type series back to the eighteenth century. Clim. Dynam. 52 (9–10), 6131–6148.

Dobbie, W., Goldin, J., Yang, C.S., 2018. The effects of pretrial detention on conviction, future Crime, and employment: evidence from randomly assigned judges. Am. Econ. Rev. 108 (2), 201–240.

Feehanan, J., 2012. The potato: root of the Famine. In: Crowley, John, Smyth, William J., Murphy, Mike (Eds.), Atlas of the Great Irish Famine, 1842-1852. Cork University Press.

Ferrie, J.P., 1996. A new sample of males linked from the public use microdata sample of the 1850 US federal census of population to the 1860 US federal census manuscript schedules. Hist. Methods 29 (4), 141–156.

Foster, R., 2015. Vivid Faces: the Revolutionary Generation in Ireland, 1890-1923. Penguin, London.

Fryer, R., Levitt, S., 2004. The causes and consequences of distinctively black names. Q. J. Econ. 119 (3), 767–805.

Gleditsch, K.S., 2009. All International Politics Is Local: the Diffusion of Conflict, Integration, and Democratization. University of Michigan Press, Ann Arbor, MI.

Goodspeed, T.B., 2016. Microcredit and adjustment to environmental shock: evidence from the Great famine in Ireland. J. Dev. Econ. 121 (C), 258–277.

Goss, E.M., 2014. The Irish potato Famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. Proceedings of the National Academy of Sciences of the United States 111 (24), 8791–8796.

Güell, M., Rodríguez Mora, J.V., Telmer, C.I., 2014. The informational content of surnames, the evolution of intergenerational mobility, and assortative mating. Rev. Econ. Stud. 82, 693–735, 2(1).

Humphreys, M., Weinstein, J.M., 2008. Who fights? The determinants of participation in Civil War. Am. J. Polit. Sci. 52 (2), 436–455.

Johnson, D., 2010. Transmission of *Phytophthora infestans* from infected potato seed tubers to emerged shoots. Plant Dis. 94, 18–23.

Kenealy, C., 2002. The Great Irish Famine: Impact, Ideology, and Rebellion, British History in Perspective. Palgrave, London.

Killeen, R., 2007. A Short History of the Irish Revolution 1912 to 1927. Gill Books, Dublin.

Krueger, A.B., 2007. What Makes a Terrorist: Economics and Roots of Terrorism. Princeton University Press, Princeton, N.J.

Krueger, A.B., 2015. Attitudes and action: public opinion and the occurrence of international terrorism. Science 325 (5947), 1534–1536.

Kuran, T., 1991. The East European revolution of 1989: is it surprising that we were surprised? Am. Econ. Rev. 81 (2), 121–125. May.

Kushalappa, A.C., Ludwig, A., 1982. Calculation of apparent infection rate in plant disease of a method to correct for host growth. Phytopatology 72, 1373.

Long, J., Ferrie, J., 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. Am. Econ. Rev. 103 (4), 1109–1137.

Miguel, E., Satyanath, S., Sergenti, E., 2004. Economic shocks and civil conflict: an instrumental variable approach. J. Polit. Econ. 112 (4), 725–753.

Narciso, G., Severgnini, B., Vardanyan, G., 2020. The Long-Run Impact of Historical Shocks on the Decision to Migrate: Evidence from the Irish Migration. CReAM Discussion Paper Series.

Ó Grada, C., O'Rourke, K., 1997. Migration as disaster relief: lessons from the Great Irish famine. Eur. Rev. Econ. Hist. 1 (1), 3–25.

Ó Grada, C., 1989. The Great Irish Famine. Princeton University Press, London.

Ó Grada, C., 1993. Ireland before and after the Famine. Explorations in Economic History, 1800-1925. Manchester University Press, Manchester and York.

Ó Grada, C., 1994. Ireland: A New Economic History, 1780-1939. Oxford University Press, New York, NY.

Ó Grada, C., 1999. Black '47 and beyond: the Great Irish Famine in History, Economy and Memory. Princeton University Press, Princeton, NJ.

Ó Grada, C., 2006. Ireland's Great Famine: Interdisciplinary Perspectives. UCD Press, Dublin.

Ó Grada, C., 2009. Famine: A Short History. Princeton University Press, Princeton, NJ.

Roodman, D., Nielsen, M.Ø., MacKinnon, J.G., Webb, M.D., 2019. Fast and wild: bootstap inference in Stata using boottest. STATA J. 19 (1), 40–60.

Sen, A., 1999. Development as Freedom. Anchor Books, New York.

Van der Plank, J.E., 1963. Plant Diseases. Epidemics and Control. Academic Press, New York.

Zadoks, J., Kampmeijer, P., 1977. The role of crop population and their deployment, Illustrated by Means of Simulator, EPIMUL76. Annals of the New Academy of Sciences 287 (1), 164–190.