

DISCUSSION PAPER SERIES

DP17439
(v. 2)

THE IRISH IN ENGLAND

Neil Cummins and Cormac Ó Gráda

**ECONOMIC HISTORY AND PUBLIC
ECONOMICS**

CEPR

THE IRISH IN ENGLAND

Neil Cummins and Cormac Ó Gráda

Discussion Paper DP17439
First Published 05 July 2022
This Revision 25 October 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Economic History
- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Neil Cummins and Cormac Ó Gráda

THE IRISH IN ENGLAND

Abstract

The successful assimilation of ethnic minorities into Western economies is one of the biggest challenges facing the Modern World. The substantial flows of Irish, to England, provide an historical example of this process. However, this has received surprisingly little scholarly attention. We use the universe of probate and vital registers of births, marriages and deaths, from England, 1838 to 2018, to document the status of the Irish in England. We identify the 'Irish' in the records as those individuals with distinctively Irish surnames. From at least the mid 19th century to 2018, the Irish in England have persisted as an underclass, 30-50% poorer than the English. Infant mortality is about 25% higher for the Irish 1838-1950 but has subsequently equalized. Sorting, both into urban areas, and to the North, are important elements in the Irish experience. We discuss the potential roles of selective migration, social mobility, and discrimination in this, and signpost directions for future research.

JEL Classification: N00, N33, N34

Keywords: Inequality, Economic history, Big data

Neil Cummins - n.j.cummins@lse.ac.uk
London School of Economics and Political Science and CEPR

Cormac Ó Gráda - cormac.ograda@ucd.ie
University College Dublin

The Irish in England

Neil J Cummins and Cormac Ó Gráda*

October 25, 2022

Abstract

The successful assimilation of ethnic minorities into Western economies is one of the biggest challenges facing the Modern World. The substantial flows of Irish, to England, provide an historical example of this process. However, this has received surprisingly little scholarly attention. We use the universe of probate and vital registers of births, marriages and deaths, from England, 1838 to 2018, to document the status of the Irish in England. We identify the ‘Irish’ in the records as those individuals with distinctively Irish surnames. From at least the mid 19th century to 2018, the Irish in England have persisted as an underclass, 30-50% poorer than the English. Infant mortality is about 25% higher for the Irish 1838-1950 but has subsequently equalized. Sorting, both into urban areas, and to the North, are important elements in the Irish experience. We discuss the potential roles of selective migration, social mobility, and discrimination in this, and signpost directions for future research.

JEL: N00, N33, N34.

Keywords: inequality; economic history; big data.

1 Introduction

The successful assimilation of ethnic minorities into Western economies is one of the biggest challenges facing the Modern World. The history of the Irish in England provides an important case study of this process. Even before the Great Famine of the 1840s, the Industrial Revolution had attracted hundreds of thousands of Irish immigrants to Britain. They were to be found mainly as unskilled labourers in the slums of cities such as Liverpool, Manchester, Glasgow, Birmingham, and London, and their satellite towns, where they were resented by most of the local population. During the Famine, the pressure they placed on labour markets and on public health made them even less welcome (Williamson (1986); Neal (1997); Darwen et al. (2019)).

As Figure 1.1a illustrates, the annual exodus from Ireland peaked at about 100,000 during and after the Great Famine and remained positive thereafter, with small breaks during the world wars and in the 1970s until the early 1990s. Mainly as a result, the population of the island of Ireland still today (at just under 7m) falls short of its pre-Great Famine level of about 8.5m. Further, if Ireland had followed England’s population growth rates, Ireland should today be an island not of 7m, but of over 40m (Figure 1.1b). These missing 35 million ‘Irish’ are instead to be found elsewhere, primarily in North America and Britain.

Within Britain, the Irish have long been the ‘other’ ethnic group. In the post-Famine decades, mass migration to North America overshadowed migration across the Irish Sea, but the latter continued to be substantial (Figure 1.1c). It is reflected in the number of Irish-born in Britain,

*Cummins: Economic history, LSE and CEPR. Email: n.j.cummins@lse.ac.uk, neilcummins.com. Ó Gráda: Economics, University College Dublin. Thanks to Guido Alfani, John Fitzgerald, Don MacRaild, and to Kevin O’Rourke, Peter Solar and other participants at the Irish Quantitative History Conference at Trinity College Dublin, and seminar participants in Brussels, Bergen and York for comments.

which rose from 419,256 in 1841 to 727,326 in 1851, before falling thereafter from 806,000 in 1861 to 550,040 in Britain in 1911. After World War 1, in the wake of increased US immigration restrictions, Great Britain once again became the main destination of Irish emigrants and continues to be so. In 1971 the Irish-born in Britain reached an all-time high of nearly a million (957,830), but then fell back to 837,464 in 1991 and 681,952 in 2011 (Figure 1.1d). As the immigrants married and had children the numerical gap between the Irish-born and “the Irish in Britain” rose over time, although due to assimilation most of the latter identified as British rather than Irish in the UK census (Hickman (2011)).

While the literature on the Irish in Britain is voluminous and interdisciplinary, studies of how they fared in material terms are rather few. How long did it take them to converge with the rest of the population in terms of economic wellbeing and health? Or did they converge? There is a pervasive sense in the literature that, unlike their cousins who opted for emigration further afield, the stories of the Irish in Britain was not, by and large, ones of success. Even in the new millennium, several studies stress Irish disadvantage in terms of health and life expectancy, with that disadvantage persisting to the third generation (Harding and Balaraman (2001); Delaney et al. (2013); Das-Munshi et al. (2013)).

Accounts of Irish underachievement and marginalisation dominate the historiography. Referring to men of working age in 1972, Heath and Ridge (1983) found that comparing the Irish to the native English, far fewer of the former had achieved white-collar jobs and many more remained unskilled labourers. As might have been expected, transitions from farming to labouring were common, but there were significant flows too from other backgrounds into unskilled labour. However, Li and Heath (2008) find that while the social mobility of Irish males in Britain, as measured by progressing to white-collar salaried employment, lagged behind British males up to the early 1990s, they were surpassing them by the mid-2000s. More recently, Li and Heath (2020) invoke the first six waves (2009-2014) of the Household Longitudinal Study to study relative status by ethnic group. They find that Irishwomen in Britain matched white British in terms of employment and earnings, while the male Irish disadvantage in terms of unemployment can be accounted for by demographic factors.

None of these studies, however, covers a period of more than a few decades, and measures that would straddle longer periods are scarce. Despite nearly two centuries of substantial flows from Ireland to England, and despite this being a central feature of the cultural identity and history of both nations, the socio-economic position of those of Irish heritage within Britain, is poorly understood. There are few empirical studies that assess the social position of the Irish in England, on a consistent basis, over time. Our analysis presents the most extensive documentation of the Irish in England to date.

This paper uses the universe of probate and vital registers of births, marriages and deaths, from England, 1838 to 2018, to document the status of the Irish in England. We identify the ‘Irish’ in the records as those individuals with distinctively Irish surnames. We assign ethnicity to a surname based on the distribution of birthplaces of individuals holding a given surname in the 1911 census of England and Wales. For robustness, we also calculate ethnicity in this way using Onomap, a contemporary classification system based on billions of global records.

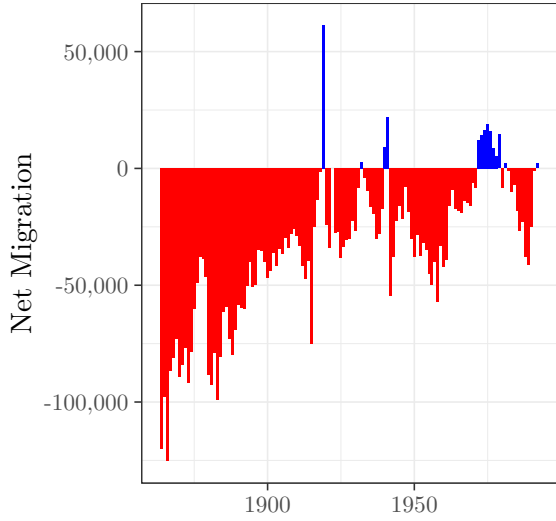
We measure status in two ways; wealth at death, and infant mortality. Thus we capture ethnic inequality both at the start and end of life. The results are stark. From at least the mid 19th century the ‘Irish’ in England have persisted as an underclass. We document the lower wealth, and higher infant mortality, of those with Irish surnames. Using linked data we show that this Irish effect is robust to age controls. Thus this lower wealth is not an artifact of the return migration of richer, older Irish to Ireland. The ‘Irish’ are always poorer than the English, and this pattern is persistent throughout 1858-2018. ¹ We show that the Irish wealth penalty is driven by the experience of the Irish in the North of England. Recently, however, there is evidence that the Irish infant mortality rate has converged with that of the native English. Half of the Irish infant mortality effect is sorting into higher mortality districts.

¹Earlier work by one of us indicates that a significant proportion of probated wealth is ‘hidden’ after 1920 (Cummins (2022b)). We assume here that, conditional on wealth, the Irish are just as likely to hide wealth as the English.

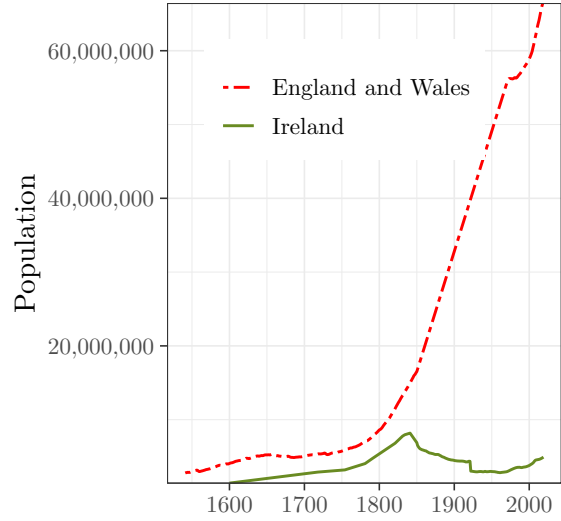
Could the Irish simply have been sending their wealth home? Although emigrant remittances, mainly from the United States, were an important feature of Irish life for a century or more after the Great Famine, hard data on them are lacking. Official data on Irish emigrant remittances are available for 1940 and 1970, when they were considered important enough to be recorded in the national accounts as income. These data are necessarily approximations, but it is reckoned that annual remittances from the United Kingdom to the Republic of Ireland averaged £5.7 million during that period. That implies that such remittances added about 1.5% to Irish GDP in mid-century and 0.5% in the 1960s. The contribution per Irish-born resident of the UK averaged £10-£12 over this period. It may be supposed that as the number of Irish-born declined, the average sum remitted rose as incomes rose. However, the Irish born were a minority of all those with Irish surnames in England, throughout. Thus remittances can only potentially explain a small proportion of the Irish wealth gap.²

The data are presented in section 2, the methodology in section 3 which describes in detail the process and accuracy of the surname ethnicity assignment, and the construction of the three wealth measures, and the infant mortality rate. Section 4 presents the results and section 5 describes the results using an alternative ethnic classification, Onomap. What explains these patterns? We discuss the role of social mobility, migrant selection, the marriage market in the assimilation of the Irish into the English in section 7. Section 8 concludes.

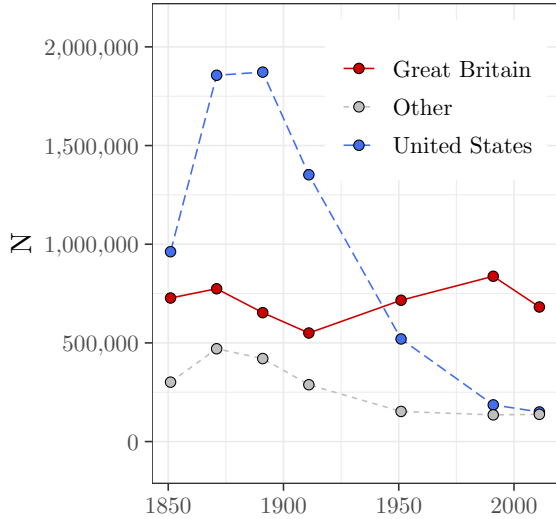
²For estimates of remittances from the United States to the United Kingdom see Schrier (1958, p.167–8). Central Statistics Office (Dublin), Statistical Abstract, various years; Office for National Statistics (2013).



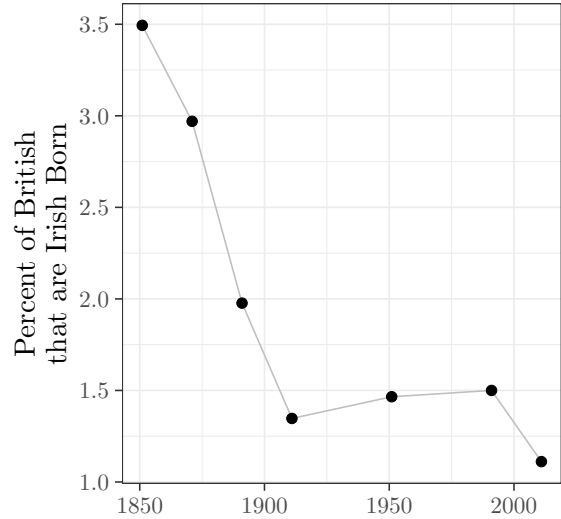
(a) Net Migration, Ireland, 1850-1992



(b) Population Growth, Ireland compared with England & Wales, 1600-2019



(c) Numbers of Irish Born Abroad, 1850-2000



(d) Irish Born as % of British Population

Figure 1.1: The Context of Irish Migration

Sources: Ireland's population, 1600-1850 Ó Gráda (1979), 1850-1951 Rothenbacher (2005), 1951-2019: cso.ie. England and Wales, 1541-1851: Wrigley and Schofield (1981), 1851-1871Rothenbacher (2005), 1971-2019: ONS.

2 Data

2.1 Wealth

We use estimates of wealth-at-death from a complete transcription of the *Principal Probate Registry (PPR) Calendar* entries, 1858-1992. This source records all those who die with wealth in England above the probate threshold.³ Cummins (2021) investigates in depth the quality of the transcription and assesses the credibility of the wealth estimates. The top percentile wealth-share estimates match closely existing estimates from different sources (Inland Revenue) Atkinson and Harrison (1978); Atkinson et al. (1989); Atkinson (2013) and Alvaredo et al. (2018). The PPR wealth data matches well to estimates of wealth reported by Blake and Orszag (1999).⁴

The PPR Calendar data was supplemented by a database of the number of deaths and the number of probates, by surname, 1996-2018. Every probate over this period is listed, by name, on <https://probatesearch.service.gov.uk/#calendar>. Note that the interpretation of probate changes after 2010 where banks had discretion on whether they required an act of probate for estates below £50, 000.⁵

2.2 Registers of Births, Marriages and Deaths, 1837-2007

On the 1st of July 1837 a National Civil Registration system was established in England and Wales. Recently these records have been digitized by various groups interested in family history. We compiled a database of 125,005,217 births 47,082,406 marriages, and 85,932,666 deaths, from 1837 to 2007, for England and Wales by downloading the individual index entries from two such websites: freebmd.com (1837-1980) and familysearch.org (1980-2007). Figure A.2, reported in the appendix, by year for each vital series a comparison of the numbers collected versus that recorded by the official records (from Office for National Statistics (2021b).)

In all cases the harvested counts closely match that expected from official statistics for the vast majority of years between 1837 and 2007. The exceptions are the sharp drops in numbers harvested in the 1970s for births and marriages; this is because the underlying website (freebmd.com) was incomplete for those years when the data was collected.

3 Methodology

3.1 Surnames and Ethnicity

Surnames are hereditary cultural labels typically transmitted along the paternal line of inheritance. Thus surnames can track clusters of genetic related individuals (primarily men). We use surnames as a marker of ethnicity.⁶ We define a surname as “Irish” if the proportion of surname holders born in Ireland of a given surname is above a threshold level in 1911.

We use the 36 million de-anonymized individual records from the special access version of the 1911 census, to examine the distribution of place of birth for the over 500,000 surnames (Schurer and Higgs (2021)). Table 3.1 reports the top 25 most numerous countries of birth listed in 1911. Nearly 90% of those enumerated were born in England, 6.5% were born in Wales, 1.5% in Scotland, 1.2% in Ireland. All other countries each represent far less than 1%.

Based on table 3.1 we pick 11 countries of birth to attribute an ethnicity to surnames. Note that this method requires us to proxy ethnicity by the relative frequency of surname holders

³The PPR Calendars will therefore include wealth-holders dying outside England. The probate threshold during the period 1858-1900 was £10, 1901-1931: £50, 1932-1964: £100, 1965-1974: £500, 1975-1984: £1,500, 1984 onwards: £5,000 Cummins (2021, table 1).

⁴ Appendix figure A.1 for a reproduction of some of these comparisons over time, from Cummins, 2021.

⁵See appendix section B for more detail on this.

⁶Overviews of the use of surnames to infer ethnicity, in the social sciences and genetics, are given in (Mateos, 2007; Mateos et al., 2011).

Table 3.1: Distribution of 1911 Census Population by Country of Birth

Country	N	%
England	28,052,691	89.19
Wales	2,052,922	6.53
Scotland	458,153	1.46
Ireland	372,708	1.18
Russia	72,533	0.23
India and Pakistan	58,598	0.19
Germany	55,237	0.18
France	40,242	0.13
United States	36,025	0.11
Isle of Man	35,111	0.11
Australia	21,410	0.07
Italy	18,412	0.06
Poland	18,253	0.06
Canada	17,493	0.06
South Africa	16,650	0.05
Austria	12,196	0.04
Switzerland	9,877	0.03
Netherlands	7,859	0.02
At Sea	6,082	0.02
Malta	5,863	0.02
Sweden	5,724	0.02
Norway	5,493	0.02
Belgium	5,397	0.02
New Zealand	5,282	0.02
Gibraltar	4,758	0.02

Source: 1911 Census

birth countries. Thus we cannot use this method to categorize Jewish surnames, nor ethnicities that do match distinct countries in 1911.

The countries we use are England, Wales, Scotland, Ireland, Russia, India (which includes contemporary Pakistan as it's before 1947), Germany, France, Italy, Poland and the Netherlands. How to know whether a given surname corresponds to a origin country? Table 3.2 presents the matrix of the proportions born in each of the 11 countries for a set of well known surnames.

As here we do not observe the global distribution of surnames in 1911 but the distribution *within* England, we cannot simply assign the most frequent country-of-birth to a surname. This would classify many names incorrectly. For example, Stewart (Scottish), Cohen (Russian), Murphy (Irish), Durand (French), Van Gelder (Dutch), Becker (German), Posner (Russian and Polish) would all incorrectly be classified as English.⁷

To more correctly attribute Surnames to ethnic origins we therefore cross reference the complete matrix of surnames by country of birth (as represented by the example Surnames in table 3.2), with the average proportions born in England from table 3.1. We first attribute to each surname an ethnic origin based upon the most frequent country of birth. Where there is a country other than England or Wales which accounts for 5% or over of the births of that surname, we update the ethnic origin to that country. This procedure works to attribute correctly all of the example surnames in table 3.2).⁸

⁷An example of how this happens is to imagine a migrant couple, from Ireland, with a unique surname, moving to London in 1900, and having 5 children. By 1911, 5/7, or 71%, of the holders of the name, in England, would be born in England, even if this is arguably an "Irish" family.

⁸Upon inspection, it was apparent that this method incorrectly assigned many Welsh surnames as "English" (e.g. Jones, Edwards and Hughes). This is because of the very unequal population sizes of the two neighboring countries. 44% of Jones, 37% of Hughes and 25% of Edwards, are born in Wales. We therefore updated a surname to "Welsh" if more than 20% of the holders of a surname were born in Wales. As 6.5% of the population of England and Wales was born in Wales (table 3.1), the 20% cutoff here implies that the holders of a "Welsh"

Table 3.2: Example Surnames for Attributing Ethnicity from the 1911 Census

Surname	N	Country of Birth										
		England	Wales	Scotland	Ireland	Russia	Germany	India	France	Italy	Poland	Netherlands
Churchill	4,957	.789	.049	.003	.009	.000	.000	.002	.001	.000	.000	.000
Davies	215,938	.348	.559	.004	.002	.000	.000	.001	.000	.000	.000	.000
Stewart	19,144	.702	.020	.104	.025	.000	.001	.005	.000	.000	.000	.000
Cohen	14,816	.568	.011	.003	.003	.194	.010	.000	.001	.000	.050	.003
Murphy	24,697	.638	.055	.015	.144	.000	.000	.003	.000	.000	.000	.000
Ali	120	.358	.008	.008	.008	.000	.000	.167	.008	.000	.000	.000
Van Gelder	98	.653	.000	.000	.000	.010	.000	.000	.000	.000	.000	.122
Durand	258	.581	.004	.004	.004	.000	.000	.035	.198	.000	.000	.000
Singh	186	.032	.000	.000	.000	.000	.000	.790	.000	.000	.000	.000
Becker	861	.584	.014	.006	.010	.030	.138	.002	.006	.000	.003	.001
Ferrari	252	.425	.020	.000	.000	.000	.000	.000	.000	.226	.000	.000
Posner	328	.579	.003	.000	.003	.155	.015	.003	.000	.000	.113	.000

Note: Calculated from the 1911 census

3.2 Wealth Calculations

We first analyze the relative wealth of the Irish through three measures: 1. the probate rate, 2. average wealth, and 3. the representation of a group in the top 1% of wealth-holders.⁹ For these calculations we combine the PPR wealth data with the death data thus constructing an individual level dataset of *all* deaths, and all wealth at death estimates, 1858 to 1992. From 1996 to 2018, we observe all deaths by ethnicity, and the number of probates, by ethnicity. So for this most recent period we can calculate a probate rate by ethnicity.

The number of adults who die with no wealth, or wealth below the probate threshold, is calculated for ethnicity e as $N_{np}^e = N_{20}^e - N_p^e$ where N_{np} is the number not probated, N_p is the number probated (from the PPR calendars), and N_{20} is the number of adult deaths where age at death is greater, or equal, to 20 years, as is reported in the death registers.¹⁰ For every non-probated adult death (N_{np}^e), we generate one observation that is appended to the PPR database. We assign to these non-probated observations an inferred wealth equal to half the level of wealth observed in the PPR Calendars for the year of death, that was below the probate threshold. This follows the the standard method used by HM Revenue and Customs (Turner (2010, p.628-9)).

The probate rate (pr) is then simply calculated as the simple mean of a probated categorical dummy (D_p):

$$pr^e = \frac{N_p^e}{N_{20}^e} = \bar{D}_p^e \quad (1)$$

We can calculate the probate rate by ethnicity and year from 1858 to 1992, and from 1996 to 2018. As we only observe the number of deaths by ethnicity until 2007, we use the 2006 value of N_{20}^e for every year 2007 to 2018. We justify this based upon the flat trend in the national number of deaths as reported by Office for National Statistics (2021a). It must be recognized that this may be wrong for a specific ethnicity. But in the absence of observed data it is a reasonable approximation.

Average wealth (\bar{w}^e), 1858-1992, is calculated as

surname are at least 300% more likely to have been born in Wales than the average English.

⁹We do not analyze median wealth as the median wealth of adults dying in England is actually below the probate threshold, a point underlined in Cummins (2021). Cummins (2022a) presents estimates for these three measures for all sizable ethnicities dying in England and Wales, 1858-2018.

¹⁰As age at death is only recorded in the death registers from 1866. Therefore to calculate N_{20} for each ethnicity we calculated $\frac{N_{20}^e}{N}$ for all deaths 1866-76, then used this ratio to infer N_{20}^e for ethnicity e by calculating $N_{20}^e = N * \frac{N_{20}^{1866-76}}{N^{1866-76}}$.

$$\bar{w}^e = \frac{\sum w_p^e + \sum w_{np}^e}{N_{20}^e} \quad (2)$$

where w_p and w_{np} represent probated and non-probated wealth. Due to the construction of the synthetic individual level dataset, it is straightforward to calculate average wealth grouped by ethnicity and year.

Finally, representation within the top 1% is calculated as the mean of a dummy variable for having wealth above or equal to the 99th percentile, calculated across all adult deaths, for a given year.

The final sample size for the synthetic PPR-death register data is 71,668,665, 1858-1992, and 12,486,026, for whether an individual is probated, 1996 to 2018.

3.3 Linked Wealth-Death Sample

A concern with the interpretation of average wealth differences by ethnicity is that we could be comparing populations with different demographics. For example, the Irish dying in England, could be a unrepresentative subset of all Irish living in England. A richer, healthier majority may live in England, not die, but later return home to Ireland and die rich, and old.¹¹ Thus we would like to control for age at death, as a check against this, in our analysis.

The PPR Calendar data do not report age at death. But the death registers do, from 1866 to 2007. There are nearly 75 million deaths in England and Wales over this period. Whilst many of these death records have ‘common’ names, in that the first-forename and surname combination appears more than once in a year¹², a large number of these records are ‘unique’. About half, 38 million records, correspond to a first-forename and surname combination that is the only occurrence in a given year. As we want to maximize accuracy, we use only these ‘unique’ names to link the two databases.

The records were linked therefore where there was an exact concordance of first-forename, surname and year of death between the PPR Calendar data and the Death registers. Examples of these links are Mary Crutch (d. 2004), Rollings Watson (d. 1990), Selina Broadhurst (d. 1885), Emily Brand (d. 1937) and Cedric Fielding (d. 1931). As stated above, we only attempt to linked unambiguous matches where a decedent was one of these unique individuals who die in a given year. In other words, any person who held a name that did not uniquely identify a death in a year was dropped from the attempted link.

Table 3.3 reports some details of this process. Of the 52 million adult deaths (deaths of people 20 and above), 22 million are ‘unique’, as defined above. We are able to find 6 million of these adult deaths via linking on name and death year. For those not linked, 18 million, we can infer wealth.¹³

Figure 3.1 reports the average age at death for the linked PPR-Death data, and that for the general population, by gender. Before 1945, probated men and women are significantly older than the general population. This probably reflects the well known social status gradient in mortality. After 1950, females are exactly representative of females in the general population. However from 1950 to about 1975, linked men are *younger*. We speculate that this unexpected pattern is a result of younger men being either richer than older men in this period (and this more likely to make probate), or have a greater tendency to arrange probate, or both.

3.4 Infant Mortality

Infant mortality rates, by ethnicity e , are calculated for 1866 to 2007, from the birth and death registers.

¹¹As noted in section 2, the PPR Calendars record wealth held in England and Wales for decedents. Thus some rich Irish, residing in Ireland, with assets in England will be reported. This will result in a marginal upward bias in our estimates of the wealth of the Irish in England.

¹²For example there are 285 “Elizabeth Jones” dying in 1905.

¹³Note that our ‘unique’ sample is more heavily female (54% versus 50% compared with all adult deaths). This is because there is a greater variety of female forenames.

Table 3.3: Linked Data Characteristics, Unique Adult Deaths to PPR Calendar

	All Adult Deaths	Unique Adult Deaths
N Adult Deaths	52,115,209	22,274,610
N linked to probate		3,758,636
Age	65.72	65.75
sd	17.07	17.33
Female Dummy	0.50	0.54
sd	0.50	0.50
Birth Year	1,872.60	1,876.50
sd	33.01	32.58
Death Year	1,938.32	1,942.64
sd	35.57	34.53
Real Wealth		21,463.35
sd		333,161.55

Real Wealth is in £2015. Deaths 1866-1992

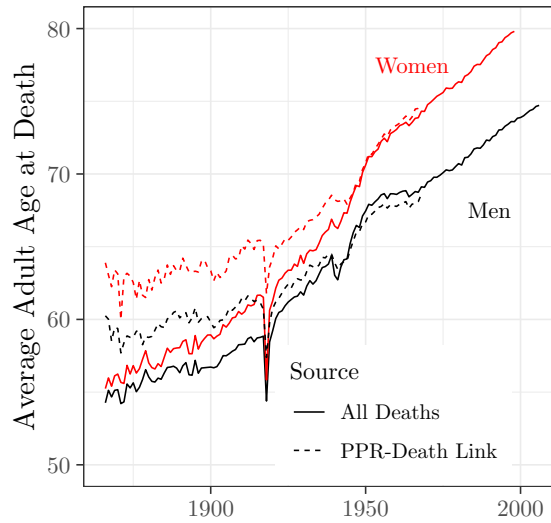


Figure 3.1: Age at Death over Time, All Deaths and Linked Probate-Deaths

$$m^e = \frac{\sum d_0^e}{\sum b^e} \quad (3)$$

where d_0 are deaths where the integer age is zero (and thus less than one years old), and b are the number of births, by year.

To analyze the determinants of infant mortality in more depth, we constructed a ‘synthetic’ individual level dataset based upon a cross tabulation of the death and birth registers. First we extracted all the death register data, by individual, on infant deaths. By comparing the counts of this individual level data, with the counts of births, we calculate how many births survived their first year of life, by ethnicity, district, and year.¹⁴ We then appended to this infant death data, a new observation for every survivor with a dummy coded as zero where a birth survives, and as one where the new born dies in their first year of life. This results in a ‘synthetic’ individual level database, not dependent on linking names, that we analyze in a standard regression framework. Figure 3.2 compare the resulting individual level estimate of the

¹⁴We assume that infant deaths are registered in the same district as their birth.

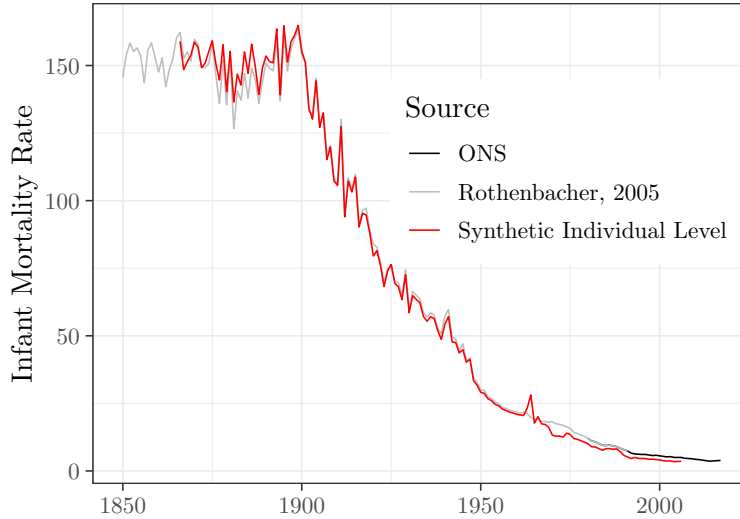


Figure 3.2: Comparison of Infant Mortality Rates

Source: Synthetic individual level data from 100% transcriptions of deaths and births, 1866-2007, Rothenbacher (2005); Office for National Statistics (2019).

infant mortality rate from the synthetic data, with that from official sources and Rothenbacher (2005). The individual rates from the synthetic data correspond closely to existing estimates.

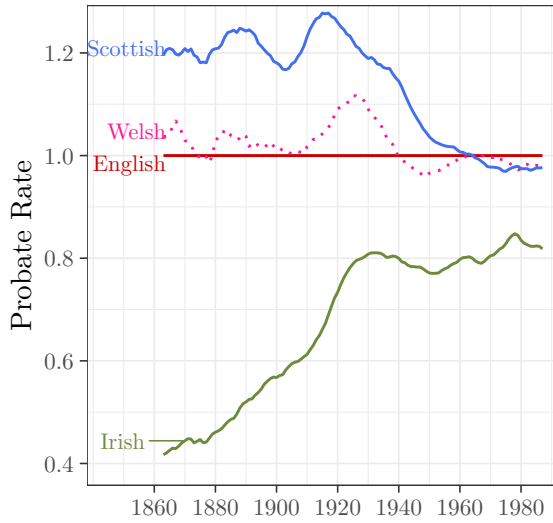
4 Results

4.1 Wealth

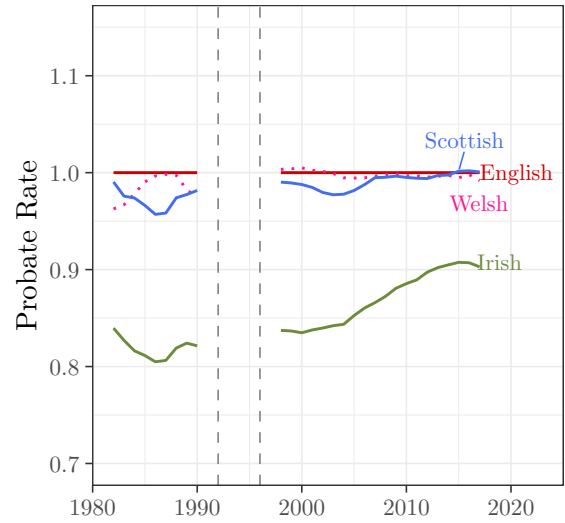
Figure 4.1 presents the pattern of wealth for the major ethnic groups of England and Wales, 1858 to 2018. Wealth is normalized so that the wealth of those with English names is set to one. The Scottish are probated at a higher rate, are richer on average, and have 50% greater representation among the top 1% of wealth holders. This advantage has declined over time. By 1960, proportions probated, and by 1990, wealth, are both approximately equal to that of the English. However the top 1% Scottish ‘effect’ is ever-present 1858 to 1992. Throughout, the Welsh, and the English, have almost exactly the same probate rate. However the Welsh are always poorer, and have a lower probability of being in the top 1%. But Welsh average wealth, by around 1990, is close to that of the English. Thus there is evidence of the convergence of wealth between ethnic groups in England and Wales, and a striking reversal of the status of non-British or Irish ethnicities.

The Irish do not share in this convergence. Throughout they have a lower probate rate, lower average wealth, and lower probabilities of being in the top 1%. The Irish ‘effect’ is persistent throughout. Proportions probated are at least 20% lower than the English, 1858 to 1990. In 2019, they are 10% less. Average wealth for the Irish is about 75% that of the English throughout, and the Irish have about 75% of the English probability of being in the top 1%.

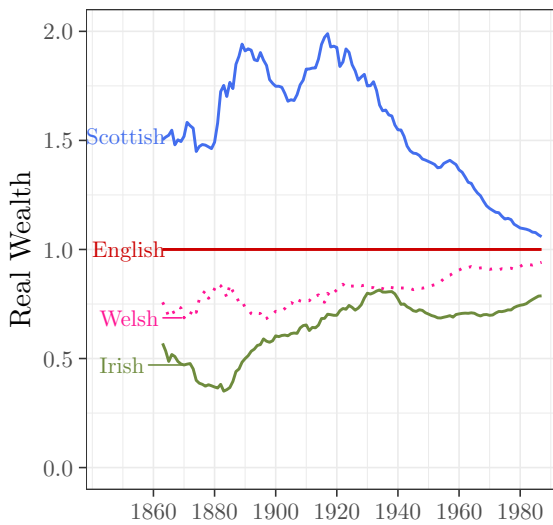
Figure 4.2 compares the distributions of wealth of the British and Irish. The two prominent peaks in all plots are a result of the attribution of inferred wealth to those who die with wealth below the probate threshold. As can be seen from panel (a), which compares the English and Irish, there is a lower share of top wealth holders amongst the Irish. The Irish underrepresentation in the top 1%, as reported in Figure 4.1 (d) is apparent at every moment of the wealth distribution. This is not the case for the Welsh and the Scottish.



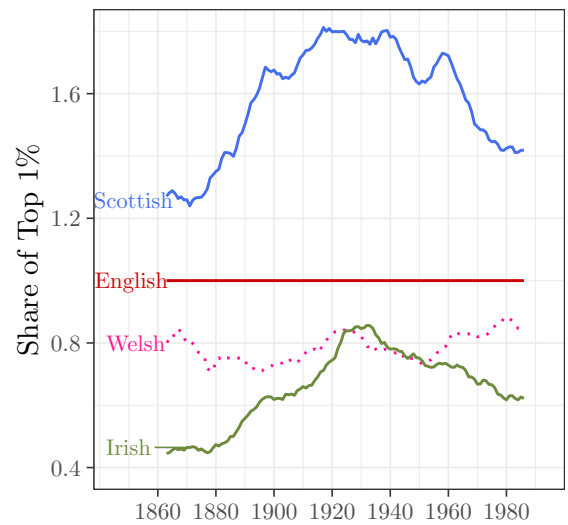
(a) Probate Rate, 1858-1992



(b) Probate Rate, 1996-2019



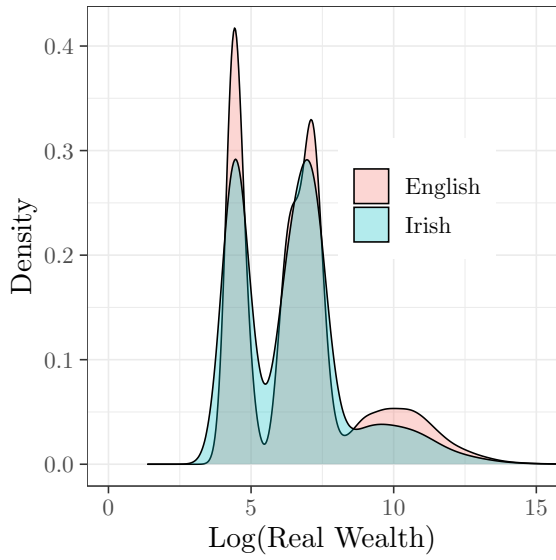
(c) Average Wealth



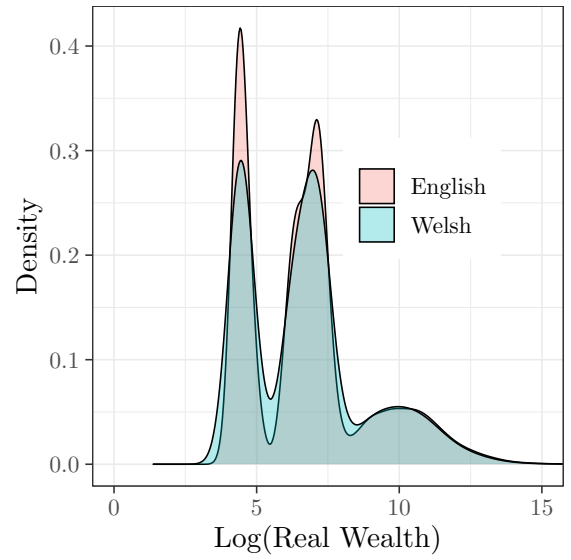
(d) Share of the Top 1%

Figure 4.1: The Wealth of the Irish and British, 1858-2018

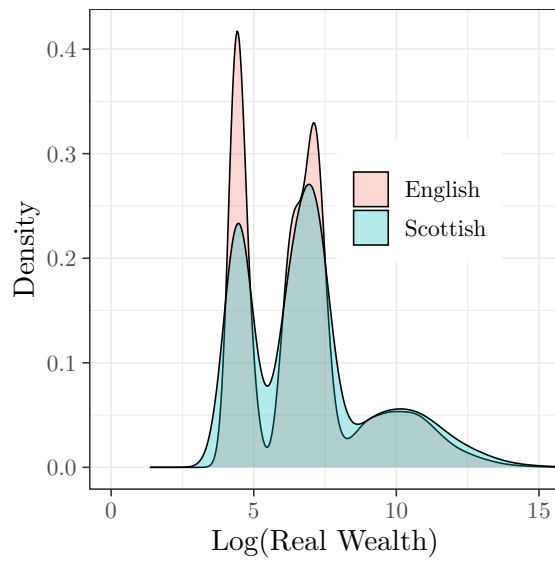
Notes: English surnames are set to one in all figures.



(a) Irish



(b) Welsh



(c) Scottish

Figure 4.2: The Distribution of Wealth, British and Irish

4.1.1 Wealth controlling for Age at Death and Place of Death

The wealth patterns reported above could have a simple explanation. They could be a result of the Irish in England being a significantly younger population than the native English. Suppose the more successful Irish spend their working lives in England, and then return home. This would lead us to observe lower wealth (and lower age at death as will be shown later), because the richer and older Irish are not in England, but in Ireland, when they die. The Irish who die in England then, are simply those unlucky few who die young. To address this we use the linked PPR-Death data, as described in section 3.3 to estimate wealth controlling for age at death.

Another confounder is the locational choice of Irish migrants. The Irish wealth effect evidenced above could simply reflect the urban character of Irish life in England during the period. Of course locational choice is endogenous to wealth but we can ask how much of the Irish wealth effect is attributable to locational sorting by including controls for place of death.

Table 4.1 reports the results of the regression

$$\log(w_i) = D_i^F + Age_i + Age_i^2 + \sum D^E + \sum D^R \quad (4)$$

where w_i is real wealth, both observed in the PPR calendars, or inferred. D represents a dummy variable for one of f , female, e , ethnic group, and R , one of the over 1,000 registration districts of death in operation over the sample period, and illustrated in appendix figure D.2.

Since wealth and age at death are endogenous, we do not assign causality to these correlations. More modestly, the test is whether controlling for age at death attenuates the ‘Irish’ effect. If it does, then that would be consistent with the Irish simply being a younger ‘at risk’ population, as measured by English wealth and death registers, with the richer, older Irish, returning to die in Ireland. If the effect is still there, controlling for age, then that is consistent with a genuine ‘Irish’ negative wealth effect.

Table 4.1 reveals that the Irish ‘effect’ is only very modestly reduced by the inclusion of age at death controls.¹⁵ Further, in all sample periods, the Irish coefficient is statically indistinguishable where age controls are used, or not. average wealth. Appendix section E decomposes the Irish wealth effect.¹⁶

The Irish wealth penalty is not a result of older Irish leaving England. Nor is it a result of locational choice.

4.2 Infant Mortality

Figure 4.3 presents the pattern of infant mortality for the British and Irish ethnic groups of England and Wales, 1866 to 2007. As with wealth, infant mortality is normalized so that of the English is set to one, by year.

English, Welsh and Scottish ethnicities have broadly similar infant mortality rates 1866 to 2007. The Irish register infant mortality rates 20 to 25% higher than the English 1866 to about 1950. Thereafter rates slowly converge by about 1990.

Infant mortality rates are much higher in urban areas during the 19th century (Woods (2000)). Is the higher infant mortality rate of the Irish a product of migration into urban slums?

To examine this we combined the birth and death data into a synthetic individual level dataset as described in section 3.4. We then ran a linear probability model of infant death on the ethnic and registration district dummies as

$$D^{ID} * 1000 = \sum D^E + \sum D^R \quad (5)$$

(note that we multiple the dummy by 1000 for ease of interpretation). Table 4.2 shows that about 50% of the Irish infant mortality effect is due to sorting between registration districts. Of

¹⁵Note that this contrasts with the effect of place on infant mortality, as reported in table 4.2.

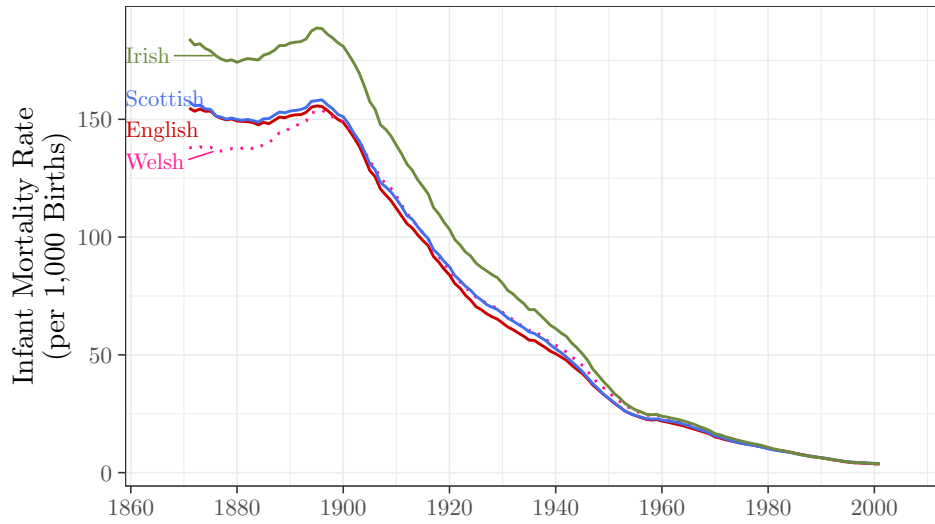
¹⁶Tables E.1 and E.2 examine the probability probated, controlling for age at death and county of death. Tables E.3 examine *probated* real wealth, controlling for age at death and county of death.

Table 4.1: Wealth and Ethncity, Linked Data: Deaths->PPR, controlling for Age at Death and District of Death

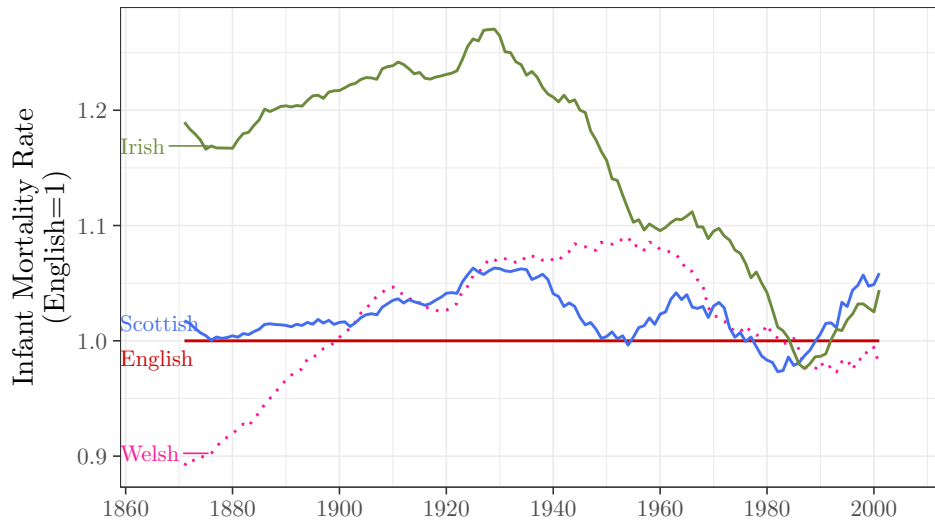
	ln(Real Wealth)					
	1866-1899		1900-49		1950-1992	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-.40*** (.002)	-.41*** (.002)	-.29*** (.001)	-.30*** (.001)	-.12*** (.001)	-.13*** (.001)
Welsh	.0002 (.01)	-.02** (.01)	.10*** (.004)	.09*** (.004)	.004 (.002)	.02*** (.002)
Scottish	.07*** (.005)	.07*** (.01)	.11*** (.003)	.11*** (.003)	-.03*** (.002)	-.03*** (.002)
Irish	-.37*** (.005)	-.34*** (.01)	-.35*** (.003)	-.31*** (.003)	-.22*** (.002)	-.19*** (.002)
Other	-.07*** (.01)	-.07*** (.01)	-.16*** (.004)	-.14*** (.004)	-.10*** (.003)	-.11*** (.003)
Age at Death Quadratic?	✓	✓	✓	✓	✓	✓
District Fixed Effects?		✓		✓		✓
Observations	3,168,149	3,168,149	7,742,186	7,742,186	10,470,387	10,470,387
R ²	.02	.03	.02	.05	.005	.02

Note:

*p<0.05; **p<0.01; ***p<0.001
 OLS, English is the omitted Group.



(a) Levels



(b) Indexed to "English"

Figure 4.3: The Infant Mortality Rate, Major Ethnicities, 1866-2007

course there could be further sorting *within* these districts that we do not observe. Given the degree of attenuation once district fixed effects are included, we speculate that the majority of this Irish mortality penalty could be due to geography.

Table 4.2: Infant Mortality and Ethnicity, controlling for Place

	Died as an Infant*1000					
	1866-1899		1900-49		1950-2007	
	(1)	(2)	(3)	(4)	(5)	(6)
Welsh	-10.69*** (0.23)	-4.82*** (0.28)	3.38*** (0.16)	0.61*** (0.18)	0.52*** (0.06)	0.32*** (0.07)
Scottish	2.05*** (0.35)	-7.89*** (0.36)	3.04*** (0.22)	-3.41*** (0.23)	0.16* (0.08)	-0.37*** (0.08)
Irish	29.04*** (0.37)	13.48*** (0.37)	18.69*** (0.22)	9.14*** (0.23)	0.98*** (0.07)	-0.14 (0.07)
Other	-9.80*** (0.71)	-22.09*** (0.73)	-12.47*** (0.40)	-18.40*** (0.41)	1.88*** (0.10)	1.00*** (0.10)
English Average	152.15		84.43		12.94	
District Fixed effects?		✓		✓		✓
Quadratic Time Trend?	✓	✓	✓	✓	✓	✓
Observations	28,720,507	28,720,507	38,297,859	38,297,859	39,622,143	39,622,143
R ²	0.0003	0.01	0.01	0.02	0.005	0.01

Note:

*p<0.05; **p<0.01; ***p<0.001

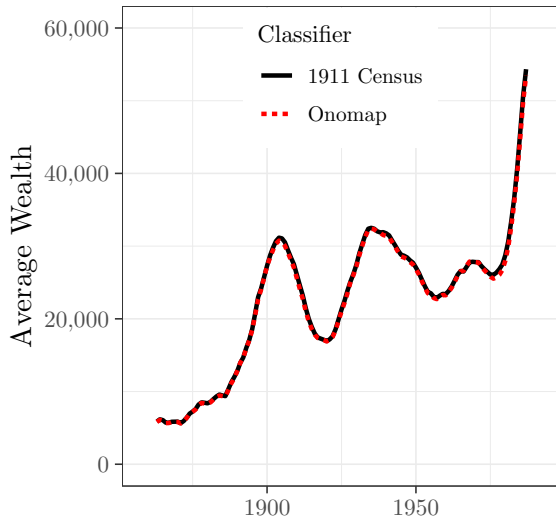
Linear Probability Model (OLS), English is the omitted Group.

5 Results using an Alternative Ethnic Classification

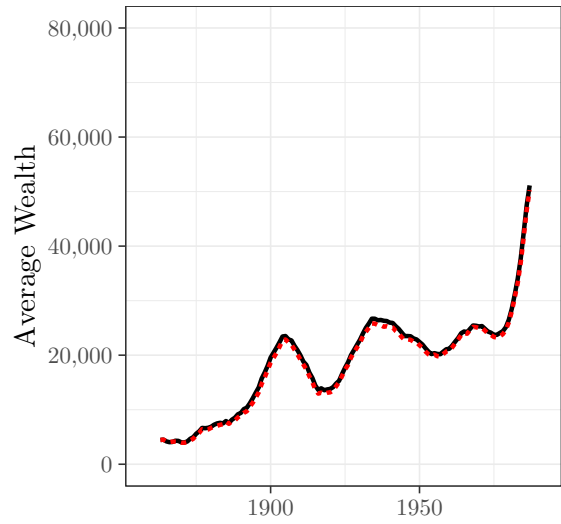
How robust are these patterns to a different method of ethnic classification? We compare the results of our 1911 ethnicity assignment with that of ‘Onomap’, a classification system developed by Paul Longley and numerous collaborators at University College London. Using billions of contemporary records from telephone directories and electoral registers, from nearly all countries in the World, a network analysis clusters surnames together based upon shared forenames (Mateos et al. (2011)). These clusters map on to known ethnocultural groups. An example for the Irish would be an observed cluster containing surnames such as *Murphy*, *McCarthy*, *Kelly*, and *O’Shea*, linked to each other through shared, distinctively Irish, forenames such as *Cormac*, *Bridget*, *Niall* and *Sorcha*.

Figure 5.1 reports the average wealth for the British and Irish ethnic groups for both classifications. They are identical for the English, the Scottish, and the Welsh. However, for the Irish the trends are different. The Onomap classifier results in wealth estimates substantially lower than that of the 1911 census assignment used here.

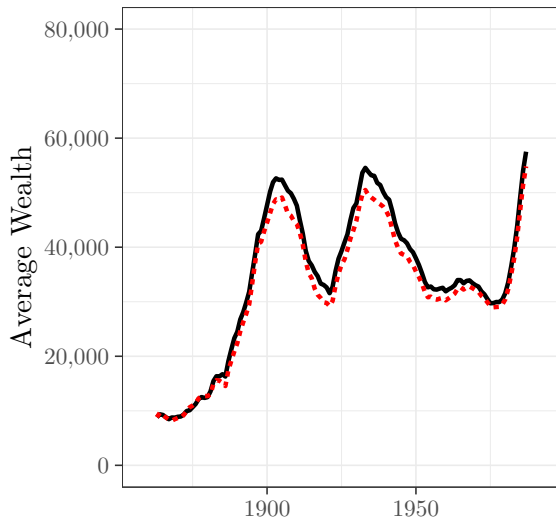
Figure 5.2 reports the infant mortality rate for the British and Irish ethnic groups for both classifications. The different classifications produce identical results.



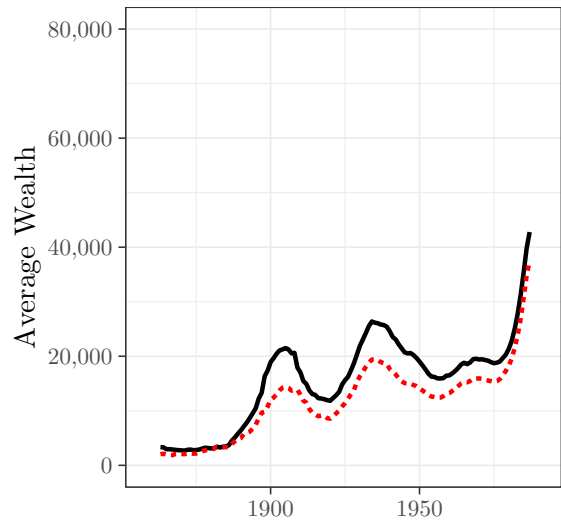
(a) English



(b) Welsh

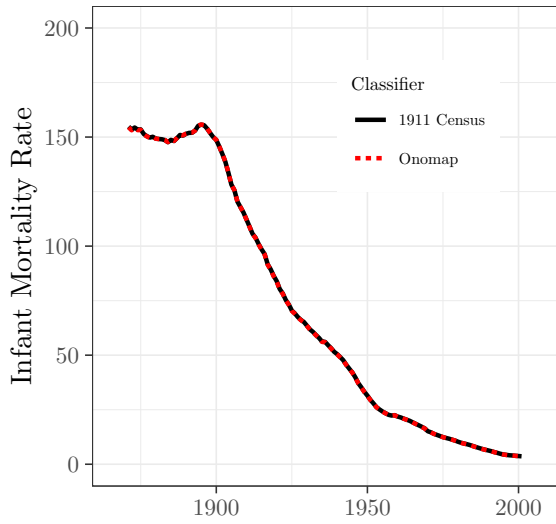


(c) Scottish

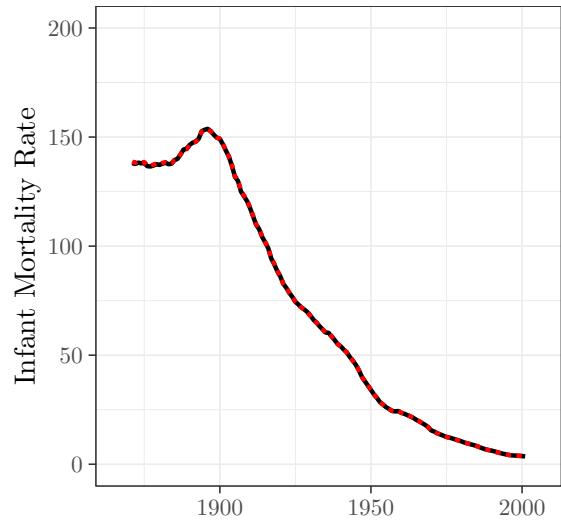


(d) Irish

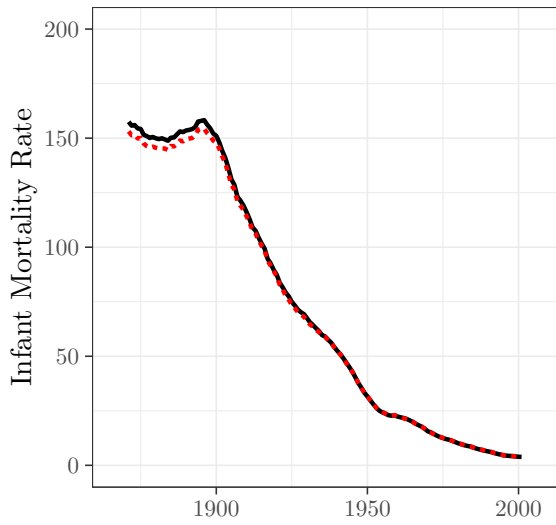
Figure 5.1: Comparing Average Wealth, 1858-1992, using Alternative Ethnic Classifier



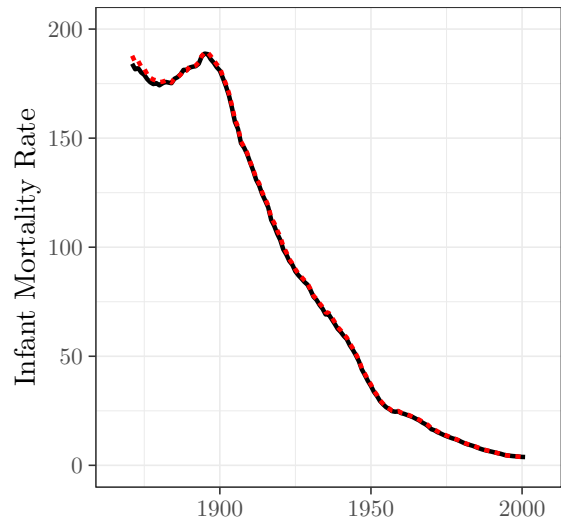
(a) English



(b) Welsh



(c) Scottish



(d) Irish

Figure 5.2: Comparing Infant Mortality Rates, 1866-2007, using Alternative Ethnic Classifier

One possible explanation for the divergence between the methods for Irish Wealth is the *contemporary* nature of the Onomap classifier. Over time, successful Irish could integrate into the English and adopt English forenames for their children. These Irish would then be classified as ‘English’ by Onomap. Of course it also could be that Onomap better classifies ethnicity than our 1911 Census classifier. (Or vice versa of course.) The 1911 census based classifier still results in Irish wealth significantly below that of the English. It may be that our choice of classifier is an overestimate of Irish wealth, relative to Onomap, and thus an *underestimate* of the true Irish-English, in England, wealth gap. The conclusions from the main analysis are unchanged.

6 Regional Differences

The Irish in England have been and remain significantly poorer at death, and until recently faced higher mortality for their infants than the English. This penalty persists even when we control for geography, as in section 4. However, it is still possible that there are different levels and trends in Irish assimilation between the different regions of England. In order to examine this, here we split England into its historic North and South division. For the South we separate out London.¹⁷

Figure 6.1 reports the trend of wealth for the British and Irish ethnicities 1860 to 1992, by region. It is evident that the majority of Irish underperformance is attributable to the Irish experience in the North. In the South (excluding London) the Irish were richer at death than the English, 1860 to 1940. After 1940 the Irish fell behind the English but the scale of the wealth-gap, at around 10%, is small relative to that observed in the North. In London the Irish are always poorer than the English (apart from a brief period around 1920). But again this wealth gap is small (less than 10%), relative to that of the North.

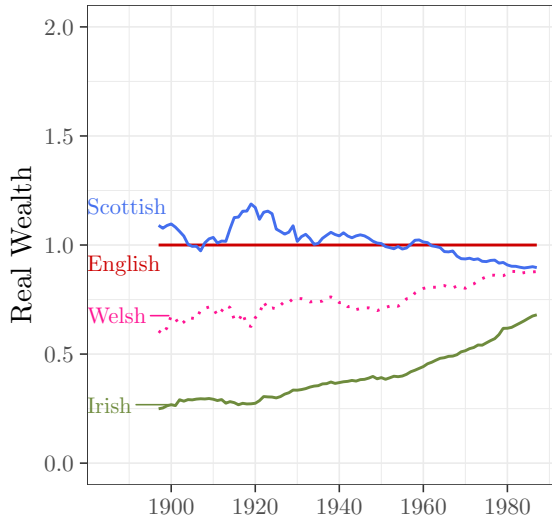
In the North, the Irish had 25% of the wealth of the English in the 19th century. This rose to about 70% by 1992. The scale of this wealth gap dwarfs that of the South, and that of London.¹⁸ It is also worth noticing here that the Scottish over-performance is not present in the North.

Figure 6.1 reveals that the overall Irish-in-England wealth ‘effect’ is driven in the main by a specific geographic penalty. It is not a simple interaction however. It is not that the North was poorer, and that this mechanically drives the observation of an Irish wealth penalty. For example, if the Irish disproportionately migrate to the poorer North this could drive the appearance of a wealth penalty overall. The birth records reveal that the Irish *did* disproportionately migrate to the North, as reported in figure 6.2. Before 1950 over 50% of Irish births were in the North compared to about 25% for the English. The Irish were more likely to be found in London and had about half the likelihood of being found in the South. The Irish were twice as likely to be found in the poorer North than the native English. Thus the North-South divide is an important element in the economic history of the Irish in England.

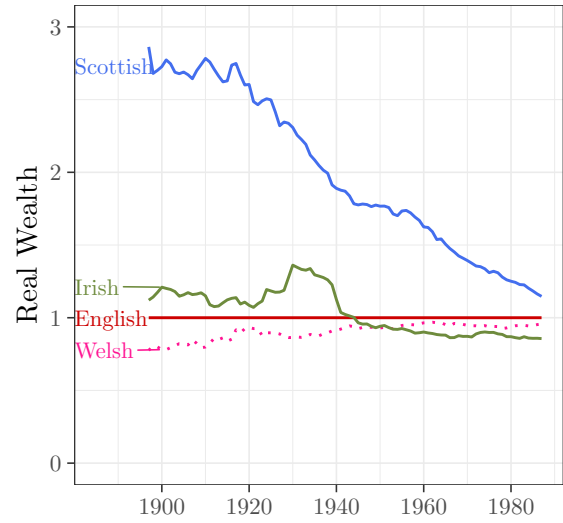
However, as revealed in figure 6.1 the Irish were far poorer relative to the English in the North than they were in the South, or in London. The specific social, economic and cultural conditions of the North resulted in Irish migrants being *much* poorer than the English. Thus, the underperformance of the Irish in England is not a result of disproportionate migration to the poorer North of England but rather an underperformance driven by the experience of the Irish in the North of England.

¹⁷The North is comprised of the counties of Cheshire, Cumberland, Durham, Lancashire, Northumberland, Westmorland, and Yorkshire.

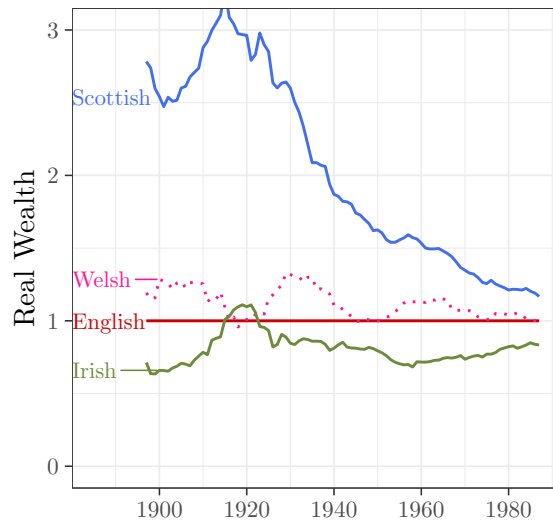
¹⁸The infant mortality rates, reported in appendix figure F.1, do not display the regional patterns of the wealth figures. Here we speculate that the urban penalty faced by migrants to both the North and South, masks the status effect picked up by the wealth data.



(a) The North



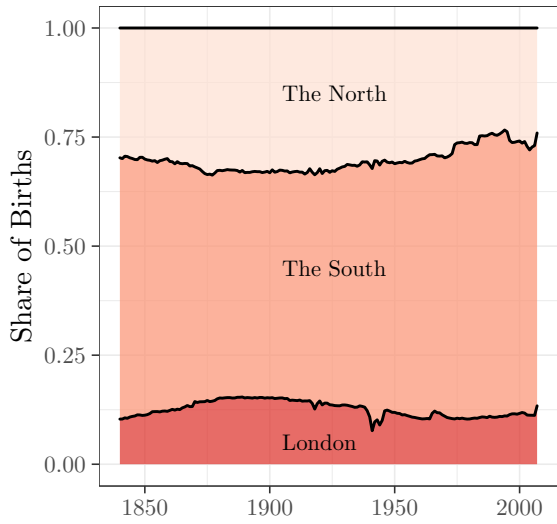
(b) The South, ex. London



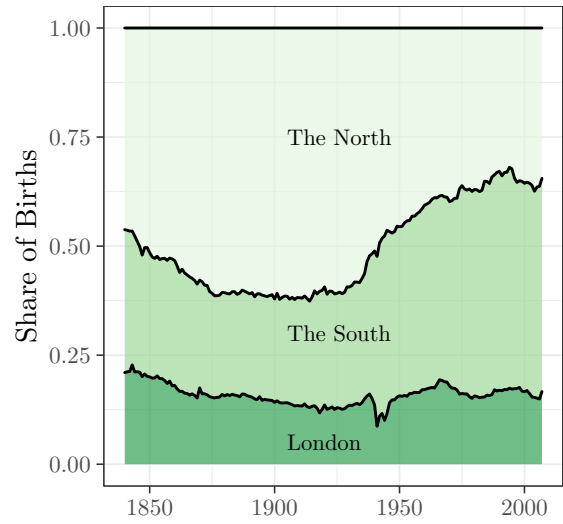
(c) London

Figure 6.1: Regional Differences in the Relative Wealth of the Irish

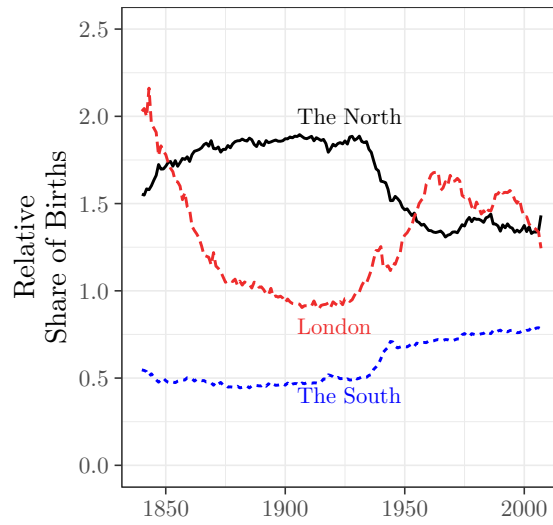
Note: The English baseline is established by region. Ethnic classification is based upon the 1911 census.



(a) English



(b) The Irish



(c) Irish relative to English

Figure 6.2: The Regional Distribution of Births, the English and the Irish Compared

7 Discussion

This paper has identified a large and persistent Irish penalty in wealth and in infant survival in England over the past century and a half. Why were these outcomes so severe for the nineteenth and twentieth-century Irish (and not the Welsh and Scottish)? Since our main objective has been to discover and describe outcomes, a rigorous analysis of what caused them is beyond the scope of the present paper. Here our discussion of likely causes is informal.

One obvious candidate is anti-Irish discrimination in the labour market and elsewhere, and the generally unwelcoming, if not outright hostile, social landscape (see for example Winder 2010). Within Britain, the Irish have long been the ‘other’ ethnic group.¹⁹ During the nineteenth and twentieth centuries crude anti-Irish stereotyping was widespread, flaring up in periods of increasing Anglo-Irish tensions such as the 1860s, the 1880s and during the Troubles of the 1970s and 1980s (Ó Tuathaigh (1981); De Nie (2004)).²⁰ And yet, despite the penchant of many for living in Irish neighbourhoods, most of the immigrants married out from early on, and the declining use of Irish forenames in the nineteenth century suggests a degree of assimilation (Smith and MacRaild (2009)).

Hard evidence of wage and job discrimination over the century and a half covered in this paper is lacking, but that is not to rule out such discrimination. The case of coal-mining, studied by MacRaild (2010), is interesting in this respect, although it refers only to the pre-1880 period. One might have assumed that Irish coal-miners would have achieved parity of status eventually, but that was not so before 1880 at least; they were still underrepresented relative to their share in the labour force in 1881, and to be found disproportionately in the lower-paid, menial categories of work. That can hardly have been because they were happy to be so. For the recent past, the evidence is mixed (compare Heath and Ridge (1983); Bruce et al. (2005); Walls and Williams (2003)).

A second likely explanation is the socio-economic character of the Irish migration flow to England. Migrants from Ireland to England may have been adversely selected, as claimed as long ago as 1776 by Adam Smith in a much-cited passage:

“The chairmen, porters, and coal-heavers in London, and those unfortunate women who live by prostitution, the strongest men and the most beautiful women perhaps in the British dominion, are said to be, the greater part of them, from the lowest rank of people in Ireland” (Smith (1776, p.161))

If migration from Ireland was indeed negatively selected, in that those who moved to England were disproportionately poorer in physical and human capital than those who remained, the patterns that we have described might reveal perhaps not so much an ‘Irish’ effect as a ‘poor’ effect. Yet in Adam Smith’s day and long after, Irish emigrants were better off and healthier than those who could not afford to leave at all. That said, the more affluent among them made their way to America, while successive cohorts of the less affluent joined an English working class that was not very upwardly mobile either. Whether this changed in the post-famine era is a question on which detailed census comparisons of the occupations of the Irish-born living in Ireland and in England could add insight from the mid-nineteenth century on. For example, the 1911 censuses suggest that the percentages with skilled occupations such as blacksmith, grocer,

¹⁹Writing in 1870, at a time when his links to the Irish community in England were closest, Karl Marx declared:

[...] in all the big industrial centres in England there is profound antagonism between the Irish proletariat and the English proletariat. The average English worker hates the Irish worker as a competitor who lowers wages and the standard of life. He feels national and religious antipathy for him. He regards him somewhat like the poor whites of the Southern States of North America regard their black slaves.

²⁰Not for nothing does one well-known survey of Irish migration to Britain between 1750 and 1922 end with a chapter on “A Culture of Anti-Irishness”, and a study of Irish migration in the interwar period that followed is entitled “Almost a Class of Helots in an Alien Land” (MacRaild (2010); Delaney (1999)). The stereotyping of the Irish made them seem more homogeneous than they really were. In Liverpool, and arguably in Glasgow too, it probably played a role in entrenching “a protective and defensive . . . ethnic affiliation” that persisted for many decades (Belchem (1999, p. 129)).

butcher, plumber, and carpenter – though not baker or tailor – were higher among the Irish who stayed at home than those who left, implying adverse selection.²¹ Such a straightforward descriptive exercise would probably lend some empirical ballast to Smith’s observation from an earlier era.²² Another way to identify selection might be to produce a set of English surnames linked to being ‘poor’, defined by their status in the 19th century, and to analyse whether they display patterns similar to the Irish. Perhaps the persistence of Irish poverty reflects, in part at least, a world of high social immobility, as Clark and Cummins (2015) argue is the case for England? In that case, the Irish may simply have become indistinguishable from the ‘poor’ English.

To sum up, this Irish status effect could reflect both poverty itself, discrimination, and assimilation, or some mix of the three. By comparing the Irish in England to the *poor* English we can explore this further. In a world where status, and wealth, persist across many generations, as is claimed by Clark and Cummins (2015) for England over the sample period of this paper, the Irish ‘penalty’ could simply reflect the typical persistence of *any* identified poor group’s status. To address this, we identify a set of poor and rich sub-groups of English, and track their relative wealth over time. Starting with all ‘rare’ English surnames, defined as having between 3 and 200 holders dying 1866-1900, we calculate average wealth for every surname by combining the sum of probated wealth with the number of non-probated (whom we assume die with £1). We then compare these surname averages with the average for all English surnames over the same period. This gives us a snapshot of who was rich, and who was poor, 1866-1900. We then define ‘Super Rich’ surnames as those that have wealth three times that of the average, ‘Rich’ as above average, ‘Poor’ have wealth 10-20% of average, and ‘Super Poor’ have wealth 10% of the average or less.

Figure 7.1 reports average wealth for these surnames during the period they were defined (1866-1900), and from 1900 to 1992. Notice that the regression to the mean is faster in the period immediately preceding when the groups were defined. This is because some rare surnames will randomly have high wealth, and some will randomly have low wealth. To measure social mobility we thus need to examine the wealth trajectories post 1900. (See Clark et al. (2014); Clark and Cummins (2015) for more detail on this idea.)

Figure 7.1 compares the Irish to this set of English wealth groups. The figure shows that the Irish, 1858 to 1992, only very modestly regress towards English mean wealth, but at a much slower rate than any of the English wealth groups. In fact 1920 to 1992, there is really no movement in the relative wealth of the Irish. Social mobility is not occurring for the Irish in England for most of the 20th century.

We cannot identify why the Irish persist as an underclass in England, poorer than even the English Victorian-defined “super poor” in 1992. If this were a result of labour market discrimination against the Irish, then we would need to also explain why the Scots, and also why almost all other ethnicities over the sample period, do not experience this (see Cummins (2022a)).

However, one possible mechanism could be the nature of the selectivity of migration from Ireland. The evidence presented here and in the wider literature is consistent with migration from Ireland to England being negatively selected. Perhaps the relentless addition of young, poorly educated immigrants to the stock of Irish in England helps to explain the persistence of Irish non-convergence, as in figure 7.1. By the same token the scale of negatively selected migration from Ireland over most of the 20th century, by increasing human capital per capita in the sending economy, may have played some part in Ireland’s rapid economic growth towards the end of the century. A population consistently pruned of the bottom quartile of its human capital distribution may find itself better primed for economic growth once the right macro conditions are satisfied. The surprisingly rapid convergence of Irish and English living standards in the 1990s and 2000s may therefore be related to the issues discussed in this paper.

²¹Personal communication from John Fitzgerald, Trinity College Dublin.

²²The Irish censuses of 1901 and 1911 can also be used to estimate return migration rates for England. English-born children, linked to Irish-born parents, cross-tabulated with the numbers of Irish-born in the English censuses, can be used to calculate such rates. These returnees could also be compared the general population, as Fernihough and Gráda (2019) do for American returnees in the 1911 census of Ireland.

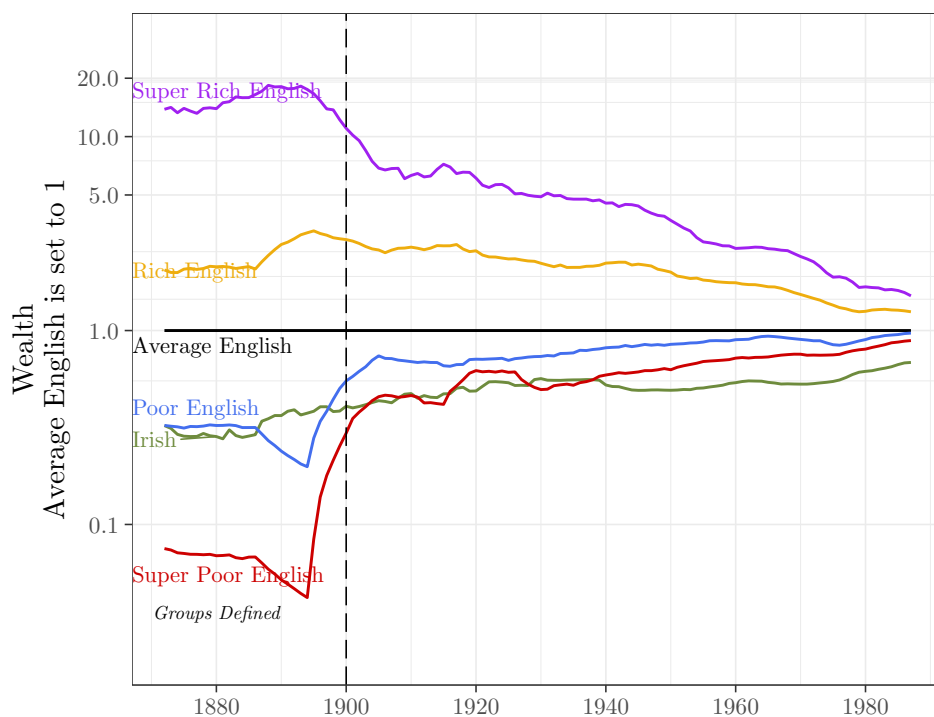


Figure 7.1: A Distinctive Irish Wealth Pattern

Notes: “Irish” and “English” are defined for a surname based on the distribution of holders’ locations of birth in the 1911 census. Taking rare English surnames who have between 3 and 200 holders dying 1866-1900, we calculate average wealth by combining the sum of probated wealth with the number of non-probated, whom we assume die with £1. We then average wealth over each surname, and compare it with the average for all English surnames. “Super Rich” surnames are those that have wealth three times that of the average, “Rich” are above average, “Poor” have wealth 10-20% of average, and “Super Poor” have wealth 10% of the average or less. The figure shows that the Irish do not regress towards the mean 1920-92, and their wealth does not track that of the English “Super Poor”. *Source:* 100% Death Register and Probate Calendar Transcriptions.

8 Conclusion

Using surnames from the universe of probate and vital registers, this paper has documented the lower wealth and higher infant mortality of the Irish, 1866 to 2018. The Irish did worse at both the end and the start of life. The Irish were poorer not because the older and richer among them returned to Ireland; controlling for age makes no difference. However, the sorting of the Irish into areas with higher infant mortality rates does explain some of that inequity. The Irish wealth penalty is in the main driven by the experience of those who migrated, and stayed, in the North of England. Now that these previous invisible inequities have been revealed future research can perhaps identify the forces that have kept the Irish as an underclass in England for so long.

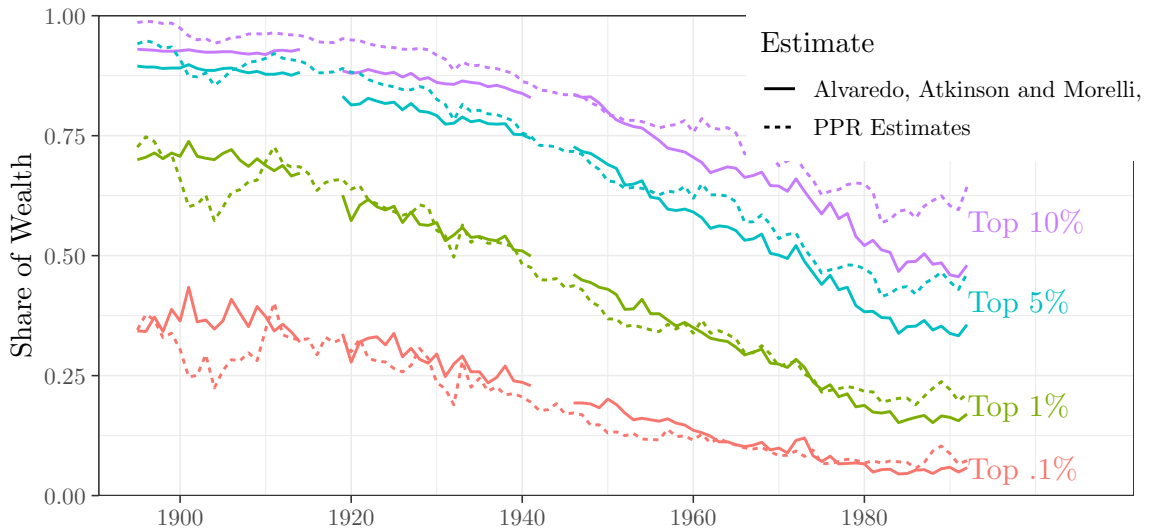
References

- Alvaredo, Facundo, Anthony B Atkinson, and Salvatore Morelli**, “Top wealth shares in the UK over more than a century,” 2017.
- , **Anthony B. Atkinson, and Salvatore Morelli**, “Top wealth shares in the UK over more than a century,” *Journal of Public Economics*, 2018, *162*, 26 – 47. In Honor of Sir Tony Atkinson (1944-2017).
- Atkinson, Anthony B.**, “Wealth and Inheritance in Britain from 1896 to the Present,” 2013.
- , **James P. F. Gordon, and Alan Harrison**, “Trends in the shares of top wealth-holders in Britain, 1923-1981,” *Oxford Bulletin of Economics and Statistics*, 1989, *51* (3), 315–332.
- , **Peter G. Backus, and John Micklewright**, “Charitable Bequests and Wealth at Death,” *The Economic Journal*, 10 2017, *127* (605), F1–F23.
- Atkinson, Anthony Barnes and Allan James Harrison**, *Distribution of personal wealth in Britain*, Cambridge Univ Pr, 1978.
- Belchem, John**, “The Liverpool–Irish enclave,” *Immigrants & Minorities*, 1999, *18* (2-3), 128–146.
- Blake, David and J. Michael Orszag**, “Annual estimates of personal wealthholdings in the United Kingdom since 1948,” *Applied Financial Economics*, 1999, *9* (4), 397–421.
- Bruce, Steve, Tony Glendinning, Iain Paterson, and Michael Rosie**, “Religious discrimination in Scotland: Fact or myth?,” *Ethnic and Racial Studies*, 2005, *28* (1), 151–168.
- Clark, Gregory and Neil Cummins**, “Intergenerational Wealth Mobility in England, 1858–2012: Surnames and Social Mobility,” *The Economic Journal*, 2015, *125* (582), 61–85.
- , – , **Daniel Diaz Vidal, Yu Hao, Tatsuya Ishii, Zach Landes, Daniel Marcin, Kuk Kuk Mo Jung, Ariel Marek, and Kevin Williams**, *The Son Also Rises: Surnames and the History of Social Mobility: Surnames and the History of Social Mobility*, Princeton University Press, 2014.
- Cummins, Neil**, “Where is the Middle Class? Inequality, Gender and the Shape of the Upper Tail from 60 million English Death and Probate Records, 1892-2016,” *Centre for Economic Policy Research Discussion Paper 13436*, 2019.
- , “Where Is the Middle Class? Evidence from 60 Million English Death and Probate Records, 1892–1992,” *The Journal of Economic History*, 2021, *81* (2), 359–404.
- , “Ethnic Wealth Inequality in England and Wales, 1858-2018,” 2022.

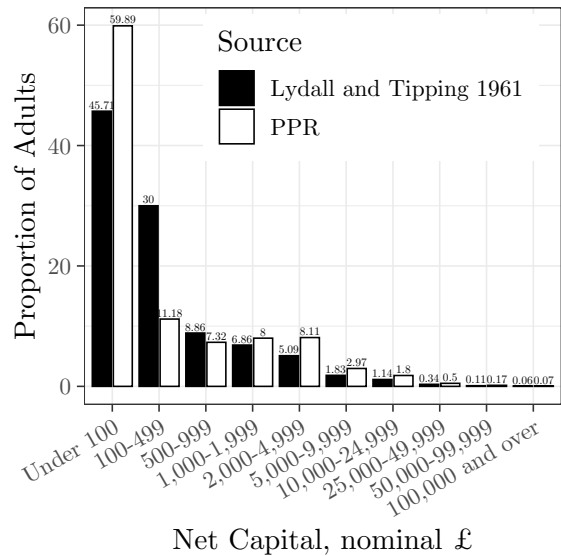
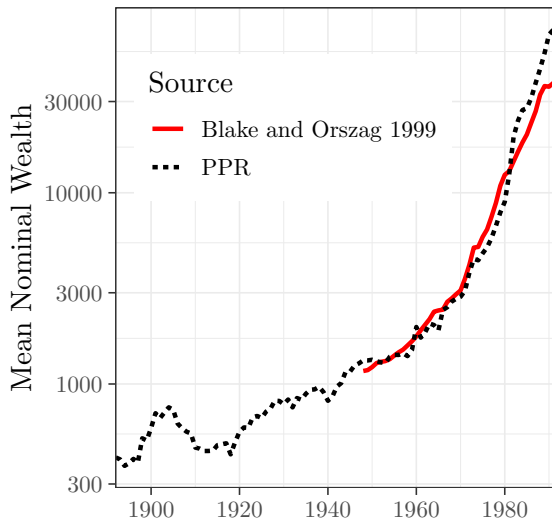
- , “The hidden wealth of English dynasties, 1892–2016,” *The Economic History Review*, 2022, *forthcoming*.
- Darwen, Lewis, Donald Macraill, Brian Gurrin, and Liam Kennedy**, “‘Unhappy and Wretched Creatures’: Charity, Poor Relief and Pauper Removal in Britain and Ireland during the Great Famine*,” *The English Historical Review*, 06 2019, *134* (568), 589–619.
- Das-Munshi, J., C. Clark, M.E. Dewey, G. Leavey, S.A. Stansfeld, and M.J. Prince**, “Born into adversity: psychological distress in two birth cohorts of second-generation Irish children growing up in Britain,” *Journal of Public Health*, 04 2013, *36* (1), 92–103.
- Delaney, Enda**, “Almost a class of helots in an alien land: The British state and Irish immigration, 1921–45,” *Immigrants & Minorities*, 1999, *18* (2-3), 240–265.
- Delaney, Liam, Alan Fernihough, and James P. Smith**, “Exporting Poor Health: The Irish in England,” *Demography*, 2013, *50* (6), 2013–2035.
- De Nie, Michael**, *The Eternal Paddy: Irish Identity and the British Press, 1798–1882*, University of Wisconsin Press., 2004.
- Fernihough, Alan and Cormac Ó Gráda**, “Across the Sea to Ireland: Return Atlantic Migration before the First World War,” Working Papers 201929, School of Economics, University College Dublin November 2019.
- GOV.UK**, “Find a will,” 2018. <https://probatesearch.service.gov.uk/#calendar>.
- , “Dealing with the estate of someone who’s died,” 2020. <https://www.gov.uk/probate-estate>.
- Harding, S and R Balarajan**, “Mortality of third generation Irish people living in England and Wales: longitudinal study,” *BMJ*, 2001, *322* (7284), 466–467.
- Heath, Anthony and John Ridge**, “Social mobility of ethnic minorities,” *Journal of Biosocial Science*, 1983, *15* (S8), 169–184.
- Hickman, Mary J.**, “Census Ethnic Categories and Second-Generation Identities: A Study of the Irish in England and Wales,” *Journal of Ethnic and Migration Studies*, 2011, *37* (1), 79–97.
- House of Commons Library**, “Probate Fees,” 2019. <http://researchbriefings.files.parliament.uk/documents/CBP-7929/CBP-7929.pdf>.
- Karagiannaki, Eleni**, “Recent Trends in the Size and the Distribution of Inherited Wealth in the UK,” *Fiscal Studies*, 2015, *36* (2), 181–213.
- Li, Y. and A. Heath**, “Minority ethnic men in British labour market (1972–2005),” *International Journal of Sociology and Social Policy*, 2008, *28* (5/6), 231–244.
- Li, Yaojun and Anthony Heath**, “Persisting disadvantages: a study of labour market dynamics of ethnic unemployment and earnings in the UK (2009–2015),” *Journal of Ethnic and Migration Studies*, 2020, *46* (5), 857–878.
- Lydall, H. F. and D. G. Tipping**, “The Distribution of Personal Wealth in Britain,” *Bulletin of the Oxford University Institute of Economics & Statistics*, 1961, *23* (1), 83–104.
- MacRaid, Donald**, *The Irish Diaspora in Britain, 1750-1939. 2nd edition*, London: Macmillan, 2010.
- Mateos, Pablo**, “A review of name-based ethnicity classification methods and their potential in population studies,” *Population, Space and Place*, 2007, *13* (4), 243–263.

- , **Paul A. Longley, and David O’Sullivan**, “Ethnicity and Population Structure in Personal Naming Networks,” *PLOS ONE*, 09 2011, 6 (9), 1–12.
- Neal, Frank**, *Black '47: Britain and the Famine Irish*, London: Macmillan, 1997.
- Office for National Statistics**, “2011 Census analysis: Immigration Patterns of Non-UK Born Populations in England and Wales in 2011,” 2013. www.ons.gov.uk: link.
- , “Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland,” 2018. www.ons.gov.uk: link.
- , “Deaths by single year of age tables - UK,” 2019.
- , “Deaths registered in England and Wales,” 2021. www.ons.gov.uk: link.
- , “Vital statistics in the UK: births, deaths and marriages,” 2021.
- Ó Gráda, Cormac**, “The population of Ireland 1700-1900: a survey,” in “Annales de démographie historique” JSTOR 1979, pp. 281–299.
- Ó Tuathaigh, M. A. G.**, “The Irish in Nineteenth-Century Britain: Problems of Integration,” *Transactions of the Royal Historical Society*, 1981, 31, 149–173.
- Rothenbacher, Franz**, *The European population since 1945*, Palgrave Macmillan, 2005.
- Schrier, Arnold.**, 1958. *Ireland and the American Emigration 1850-1900*, University of Minnesota Press., 1958.
- Schurer, Kevin and Edward Higgs**, “Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911: Special Licence Access. [data collection].,” 2021. 2nd Edition. UK Data Service. SN: 7856.
- **and Matthew Woollard**, “1881 Census for England and Wales, the Channel Islands and the Isle of Man (Enhanced Version) [computer file],” 2000.
- Smith, Adam**, “An inquiry into the wealth of nations,” *Strahan and Cadell, London*, 1776.
- Smith, Malcolm T. and Donald M. MacRaidl**, “The Irish in the Mining Industry in England and Wales: Evidence from the 1881 Census,” *Irish Economic and Social History*, 2009, 36 (1), 37–62.
- Turner, John D.**, “Wealth concentration in the European periphery: Ireland, 1858-2001,” *Oxford Economic Papers*, 2010, 62 (4), 625–646.
- Walls, Patricia and Rory Williams**, “Sectarianism at work: Accounts of employment discrimination against Irish Catholics in Scotland,” *Ethnic and Racial Studies*, 2003, 26 (4), 632–661.
- Williamson, Jeffrey G.**, “The Impact of the Irish on British Labor Markets During the Industrial Revolution,” *The Journal of Economic History*, 1986, 46 (3), 693–720.
- Woods, Robert**, *The Demography of Victorian England and Wales* Cambridge Studies in Population, Economy and Society in Past Time, Cambridge University Press, 2000.
- Wrigley, Edward. A. and Roger S. Schofield**, *The Population History of England 1541-1871*, Cambridge Universtiy Press, 1981.

A Extra Background Material on Data



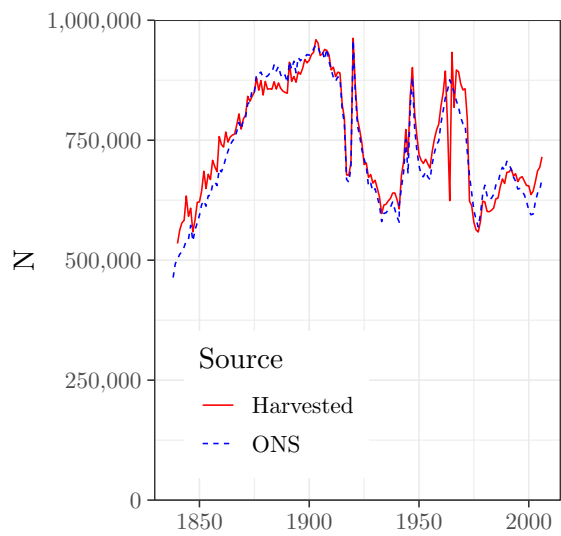
(a) Comparing Different Estimates of Top Wealth Shares, England 1892-1992



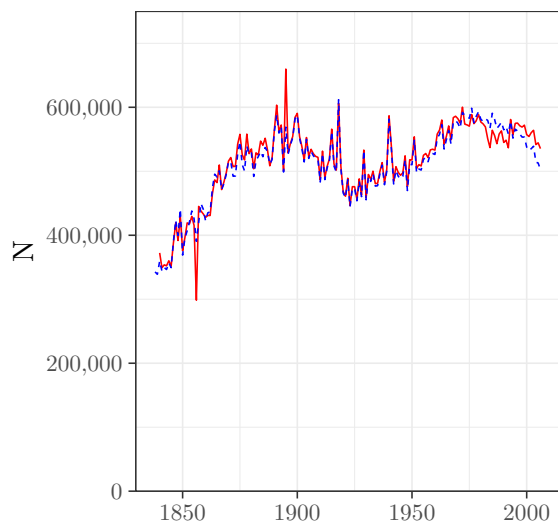
(b) Comparing Average Wealth with Blake and Orszag (1999) (c) Comparison of Net Capital with Lydall and Tipping (1961), by Wealth Band, 1950s

Figure A.1: The PPR Calendar Wealth Data, Compared with Existing Estimates

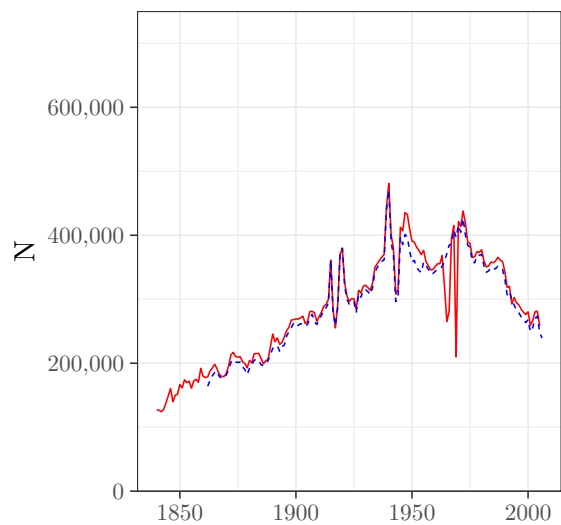
Notes: See Cummins (2021) for a detailed account of the source, construction and validation of the PPR data. Sources: PPR wealth data, Alvaredo et al. (2017) table D1, Blake and Orszag (1999, Table 12) (sum of columns 'net financial wealth', 'housing wealth' and 'consumer durable assets'). These aggregate sums were converted to a per adult measure using population data from Office for National Statistics (2018). Source for figure c: Lydall and Tipping (1961, p.89). Note that the PPR covers England, the Lydall and Tipping (1961) estimates cover Great Britain. Both estimates exclude pension wealth. These figures are also reported in Cummins, 2021.



(a) Births



(b) Deaths



(c) Marriages

Figure A.2: Data Collection Verification, 'Harvested' versus Official Count Comparison
Notes: The source for the Official Counts is Office for National Statistics (2021b).

B The Proportion with ‘Significant Wealth’, 1996-2018

The PPR Calendar data was supplemented by a database of the number of deaths and the number of probates, by surname, 1996-2018. Every probate over this period is listed, by name, on <https://probatesearch.service.gov.uk/#calendar>. It was necessary to enter an exact surname on the webpage to return the count of that surname for a given year. From a 100% sample of the 1881 census ((Schurer and Woollard, 2000)) and the 100% samples of births, marriages and deaths, 1838-2007, and the probate Calendar 1892-1992, a master-list of 3,535,375 surnames was created. Of these surnames many were mistakes so a second list was created filtering the master list by the criteria that the name appeared at least 5 or more times in the death registers, 1983-2007. This resulted in 92,812 surnames which were searched individually for every year 1996 to 2020, a total of 2,320,300 searches for each of the 25 years. (As the probate process can take a few months to a year and those years are this incomplete, I do not report the post 2018 rate here.) Each surname from this master-list was entered into <https://probatesearch.service.gov.uk/#calendar> and the count recorded (GOV.UK, 2018).

As reported in table B.1 the threshold estate value above which probate was legally required has been £5,000 from 1984 to today, 2020. In recent years however, the *de facto* reality is that financial institutions have exercised discretion in releasing monies to relatives and beneficiaries from the bank accounts of the recently deceased. In 2020, banks apply their own discretion upon which accounts need probate and which don’t. The value they apply as their probate limit could range from £5,000 to £50,000.²³

It is not clear from existing academic literature or the archives of official Govt. websites advising on probate (<https://webarchive.nationalarchives.gov.uk>) when exactly the nominal probate went from being a flat £5,000 across all institutions, to a discretionary amount that varies in the range £5-£50 thousand, and is institution specific. In 2007-8 (see Atkinson et al. (2017, F8) and as late as 2010 (See Karagiannaki (2015, p.187)), there is evidence that the £5,000 probate threshold was generally applied.²⁴

Before 1994, at least, and probably until at least 2010, the assumption that the non-probated estates were worth precisely less than £5,000 appears to be well justified. However, for post-

²³The current official Government advice on probate states “Contact each asset holder (for example a bank or mortgage company) to find out if you’ll need probate to get access to their assets. Every organization has its own rules.” GOV.UK (2020). A list of the institution specific probate limits are reported here: <https://www.co-oplegalservices.co.uk/media-centre/articles-may-aug-2018/bank-limits-for-probate/>. A news article from 1994 states “Although the Act does not specifically apply to banks and to building societies, they usually apply their discretion in a similar way, and will normally only pay out above the pounds 5,000 limit with a grant of probate.” <https://www.independent.co.uk/news/business/why-the-bereaved-must-wait-rules-governing-the-release-of-money-when-a-person-dies-can-cause-1420519.html>. A 2017 blog post by a probate professional <https://www.todayswillsandprobate.co.uk/guest-writers/obtaining-up-to-50k-without-grant-probate/> discusses the change.

²⁴Atkinson et al. (2017) state “We have been told by Her Majesty’s Revenue & Customs (HMRC) that the ‘small estate’ category probably accounts for the large majority of estates that do not go through probate ” (p.F8).

Years	Nominal Probate Threshold	Source
1858-1900	£10	Turner 2010 p.628
1901-1931	£50	Turner (2010) p.628
1932-1964	£100	Atkinson and Harrison (1978) p.36
1965-1974	£500	Atkinson and Harrison (1978) p.36
1975-1984	£1,500	Atkinson and Harrison (1978) p.36
1984→	£5,000	Turner (2010) p.628, Alvaredo et al. (2018) p.29 Atkinson et al. (2017) p.F8, Karagiannaki (2015) p.187

Table B.1: The Minimum Probate Threshold, 1858-2017

2010, and in particular more recent years, this assumption is not reasonable. Therefore, we can only interpret the probate rate 1996-2018 as being an indicator of wealth that was significant enough for the asset holders (e.g the banks or building societies) to demand an act of probate before transferring the monies. As this could be anywhere between £5,000-50,000, the probate rate after 2010 can only be interpreted as a measure of significant wealth, and not wealth above the legal probate threshold. So I report this measure separately here and do not include it in the main analysis.

Table B.2 reports the count of probates, the sum of adult deaths and the proportion probated from 1996 to 2020. The proportion of adult deaths (deaths to those aged 20 and above) requiring an act of probate to deal with their financial assets at death is consistently around 50%. (Note that the 45-47% recorded in 2016-8 may be underestimated due to the lag in recording probates.) This is consistent with the calculations of Karagiannaki (2015) in her analysis of inherited wealth, who estimates a proportion probated of about 50% for the period 2002-2007 (p.187). A figure of 50% is also reported for 2016 in House of Commons Library (2019, p.7).

Table B.2: Proportion Probated, 1996-2018

	$N_{Probates}$	$N_{AdultDeaths}$	Prop. Probated
1996	266,236	556,003	0.48
1997	270,153	551,125	0.49
1998	267,581	546,765	0.49
1999	268,320	546,980	0.49
2000	260,342	531,734	0.49
2001	257,968	526,436	0.49
2002	258,379	529,468	0.49
2003	261,600	533,201	0.49
2004	250,165	508,443	0.49
2005	251,295	507,230	0.50
2006	246,889	496,696	0.50
2007	247,885	498,258	0.50
2008	250,171	503,390	0.50
2009	242,546	485,806	0.50
2010	246,748	488,040	0.51
2011	240,566	479,335	0.50
2012	248,151	494,422	0.50
2013	249,000	502,187	0.50
2014	242,478	496,853	0.49
2015	250,743	525,073	0.48
2016	242,379	520,610	0.47
2017	248,864	528,838	0.47
2018	241,124	537,228	0.45

Source: Office for National Statistics (2019) and probatesearch.service.gov.uk

C Extra Detail

C.0.1 Irish Names

The PPR Calendar data was processed via an OCR (Optical Character Recognition) engine. The process in general worked very well and the resulting data set passed multiple data-quality tests (Cummins (2019)). Amongst the Irish, names beginning with “O” are commonplace, and non-existent within other populations. As the OCR process and the algorithms used to extract

surnames may have missed this "", I inspected all possible candidate "O" stemmed names in the PPR calendar data. This check turned up numerous oddities. For example, there are 37,613 deaths 1838-2007 for people with the surname "O'Brien" yet only 5 probates recorded, 1858-1992, for this surname. Yet, there are 3,175 probates recorded for the name "Brien" but only 2,304 deaths. I cross-checked all Irish names and assigned any possible stemmed names to the most common occurrence, as measured by the count of all deaths to that name, 1838-2007. Mechanically this was done by summing all deaths in the death data, all probates in the PPR Calendar data and in seopting all 5805 Irish names for anonommologies. This meant that all "Briens" were updated to "O'Brien", "Neill" to "O'Neill", but all "O'Sullivans" were updated to "Sullivan", "O'Daly" to "Daly". Surnames were only updated where both the stemmed and non-stemmed version were of Irish ancestry. (therby grouping over Irish doesn't make any difference to the results.

Surname	N	Updated Surname	N
O'KELLY	448	KELLY	98809
O'SULLIVAN	8085	SULLIVAN	48079
O'RYAN	118	RYAN	40263
BRIEN	2366	O'BRIEN	39108
MCGOUGH	1937	GOUGH	30224
O'CONNOR	21934	CONNOR	28858
MCCARROLL	514	CARROLL	28031
O'CARROLL	433	CARROLL	28031
O'BYRNE	411	BYRNE	25229
MCQUINN	293	QUINN	25153
MCFLYNN	17	FLYNN	21959
O'FLYNN	588	FLYNN	21959
O'FARRELL	912	FARRELL	21937
O'DONOVAN	1528	DONOVAN	20505
O'DUFFY	30	DUFFY	20451
MCCAIN	326	CAIN	17078
O'BOYLE	859	BOYLE	16204
O'CALLAGHAN	3130	CALLAGHAN	15492
MCKENNY	462	KENNY	14276
O'MAHONEY	748	MAHONEY	13101
O'DALY	51	DALY	12957
O'REGAN	909	REGAN	12947
O'DRISCOLL	1301	DRISCOLL	12774
CANN	9045	MCCANN	11692
MAHON	5040	MCMAHON	11522
O'FLANAGAN	198	FLANAGAN	11343
KENNA	899	MCKENNA	11255
O'DOHERTY	413	DOHERTY	11219
LOUGHLIN	1959	MCLOUGHLIN	10835
MCEGAN	48	EGAN	10683
MCCA VANAGH	18	CAVANAGH	8535
NALLY	397	MCNALLY	8365
MCMULLEN	3673	MULLEN	8244
O'REILLY	4966	REILLY	8179
O'LEARY	7021	LEARY	8098
O'KANE	693	KANE	7524
MCKAVANAGH	5	KAVANAGH	6693
MCKEATING	252	KEATING	6543
MCCA HILL	78	CAHILL	6541
O'SHEA	5481	SHEA	6340
O'GRADY	3345	GRADY	6310
MCGLYNN	1487	GLYNN	6152
GUINNESS	369	MCGUINNESS	5938
O'DONOGHUE	2188	DONOGHUE	5532

N is the number of deaths, 1838-2007, Continued on next page

Surname	N	Updated Surname	N
NULTY	442	MCNULTY	5461
O'FLAHERTY	935	FLAHERTY	5014
KEOWN	423	MCKEOWN	4906
COY	2924	MCCOY	4816
O'KEEFE	3462	KEEFE	4812
O'ROURKE	3927	ROURKE	4733
O'HANLON	1327	HANLON	4704
MCGARVEY	360	GARVEY	4686
O'TOOLE	2553	TOOLE	4617
MCGROGAN	112	GROGAN	4511
O'HAGAN	1361	HAGAN	4460
GARRY	1368	MCGARRY	4034
MCTIGHE	583	TIGHE	3421
MCMULLIN	793	MULLIN	3375
MALLEY	2124	O'MALLEY	3295
MCCALLAN	188	CALLAN	2779
CARTY	2746	MCCARTY	2762
MCGILLIGAN	85	GILLIGAN	2722
MCSHERRY	410	SHERRY	2714
O'DONOHUE	224	DONOHUE	2680
O'MAHONY	776	MAHONY	2572
KEEFFE	778	O'KEEFFE	2442
CLUSKEY	195	MCCLUSKEY	2441
O'DONOHUE	174	DONOHUE	2057
O'LOUGHLIN	1319	LOUGHLIN	1959
MCCORRY	177	CORRY	1895
SHAUGHNESSY	1547	O'SHAUGHNESSY	1880
MCDADE	683	DADE	1763
O'RIORDAN	540	RIORDAN	1527
MCDEVITT	345	DEVITT	1445
MCMACKIN	48	MACKIN	1397
MCGLENNON	198	GLENNON	1393
HALLORAN	893	O'HALLORAN	1318
MCGEOGHEGAN	62	GEOGHEGAN	1254
MARA	916	O'MARA	1161
O'HANRAHAN	35	HANRAHAN	1128
MCCARROL	33	CARROL	1095
MCCOLGAN	342	COLGAN	1039
CAFFERY	893	MCCAFFERY	999
MEARA	345	O'MEARA	849
KERNAN	504	MCKERNAN	844
MCLAFFERTY	43	LAFFERTY	769
MCGAHAN	462	GAHAN	740
O'BEIRNE	326	BEIRNE	729
MCCREEDY	212	CREEDY	701
CARRON	467	MCCARRON	697
CUSKER	70	MCCUSKER	651
MCCULLY	566	CULLY	598
MCMACKEN	0	MACKEN	507
MCCONVEY	65	CONVEY	487
MCCASHIN	14	CASHIN	472
MCCALVEY	22	CALVEY	458
O'BRYNE	35	BRYNE	417
GREAVY	26	MCGREAVY	381
O'HERLIHY	51	HERLIHY	374
MCTEER	152	TEER	374
KITTRICK	33	MCKITTRICK	372

N is the number of deaths, 1838-2007, Continued on next page

Surname	N	Updated Surname	N
O'RIELLY	49	RIELLY	350
CRUDDEN	57	MCCRUDDEN	346
O'RORKE	213	RORKE	328
MCCOLLUM	180	COLLUM	318
O'HEHIR	43	HEHIR	302
CLOY	64	MCCLOY	294
MCKERNEY	91	KERNEY	290
NERNEY	157	MCNERNEY	272
CUMISKEY	247	MCCUMISKEY	264
QUEENEY	157	MCQUEENEY	244
CUDDEN	83	MCCUDDEN	225
MCCANNY	6	CANNY	192
CARTIN	159	MCCARTIN	192
MCGAVIGAN	8	GAVIGAN	180
MCCOMISKEY	57	COMISKEY	163
MONAGLE	10	MCMONAGLE	157
CLENAGHAN	31	MCCLENAGHAN	139
ANANEY	0	MCANANEY	118
CRICKARD	36	MCCRICKARD	100
O'RAHILLY	21	RAHILLY	97
GRANAGHAN	45	MCGRANAGHAN	84
MCPOLIN	38	POLIN	82
MCLOUGHNEY	16	LOUGHNEY	79
SYOCK	0	SYMCOCK	70
O'CALLAGHAM	0	CALLAGHAM	68
CLARNAN	0	MCCLARNAN	67
CLAFFERTY	4	MCCLAFFERTY	60
NIFFE	0	MCNIFFE	59
ALHONE	0	MCAHONE	56
ELRUE	0	MCELRUE	45
MCTEGGART	10	TEGGART	43
AREAVEY	0	MCAREAVEY	42
ADOREY	0	MCADOREY	39
ILHONE	0	MCILHONE	39
GUONE	0	MCGUONE	35
GURREN	16	MCGURREN	34
GAGHEY	0	MCGAGHEY	33
ENIRY	0	MCENIRY	28
ILHATTON	0	MCILHATTON	23
ILLMURRAY	0	MCILLMURRAY	23
LOUGHIN	0	MCLOUGHIN	21
ELEARNEY	0	MCELEARNEY	17
GENNITY	0	MCGENNITY	17
SHEFFREY	0	MCSHEFFREY	17
ALENEY	0	MCALENEY	16
KEEFRY	0	MCKEEFRY	16
STRAVOCK	0	MCSTRAVOCK	15
ALISKEY	0	MCALISKEY	12
CUSKEY	0	MCCUSKEY	12
MCCAHERTY	0	CAHERTY	11
GAVOCK	0	MCGAVOCK	11
ILMAIL	0	MCILMAIL	11
MANNIMAN	0	MCMANNIMAN	11
MURPHY-CONNOR	0	MURPHY-O'CONNOR	11
ANAW	0	MCANAW	10
PHILOMEY	0	MCPHILOMEY	9
CARTER-GRATH	0	CARTER-MCGRATH	8

N is the number of deaths, 1838-2007, Continued on next page

Surname	N	Updated Surname	N
ANOY	0	MCANOY	8
COY-HILL	0	MCCOY-HILL	8
ILVAR	0	MCILVAR	8
ELHENNY	0	MCELHENNY	7
ERLAINE	0	MCERLAINE	7
CROSBIE-DONNELL	0	CROSBIE-MCDONNELL	6
AVINCHEY	0	MCAVINCHEY	6
CALLISKEY	0	MCCALLISKEY	6
GLEISH	0	MCGLEISH	6
CARROLL-ARDLE	0	CARROLL-MCARDLE	5
ALERNON	0	MICALERNON	5
ASTOCKER	0	MCASTOCKER	5
ATASNEY	0	MCATASNEY	5
NAIR-WILSON	0	MCNAIR-WILSON	5
ATAKNEY	0	MCATAKNEY	4
CAGHY	0	MCCAGHY	4
CUNE-COLBERT	0	MCCUNE-COLBERT	4
ERLEANE	0	MCERLEANE	4
GLEENON	0	MCGLEENON	4
ILKENNY	0	MCILKENNY	4
MENAMAN	0	MCMENAMAN	4
BARRY-CALLAGHAN	0	BARRY-O'CALLAGHAN	3
DILLON-NALLY	0	DILLON-MCNALLY	3
ANEANEY	0	MCANEANEY	3
ANENNY	0	MCANENNY	3
CONIGLEY	0	MCCONIGLEY	3
DOWELL-POLKE	0	MCDOWELL-POLKE	3
GUGGON	0	MCGUGGON	3
KEAGNEY	0	MCKEAGNEY	3
KEEFREY	0	MCKEEFREY	3
KLIZUK	0	KLIMCZUK	2
BRIDE-HARROW	0	MCBRIDE-HARROW	2
CONNELLOGUE	0	MCCONNELLOGUE	2
CUE-SMITH	0	MCCUE-SMITH	2
DERMOTT-PAINE	0	MCDERMOTT-PAINE	2
ELHENNON	0	MCELHENNON	2
ELVANNA	0	MCELVANNA	2
GEOUCH	0	MCGEOUCH	2
GOWAN-SCANLON	0	MCGOWAN-SCANLON	2
INRUE	0	MCINRUE	2
SARSTEDT-CARTHY	0	SARSTEDT-MCCARTHY	2
BINGHAM-GUINNESS	0	BINGHAM-MCGUINNESS	0
FITZPATRICK-GOUGH	0	FITZPATRICK-MCGOUGH	0
HANNAN-DWYER	0	HANNAN-O'DWYER	0
ALARNEY	0	MALARNEY	0
ALERNEY	0	MICALERNEY	0
ALORAN	0	MALORAN	0
ANARNEY	0	MCANARNEY	0
ANESPY	0	MCANESPY	0
GUICKIN	0	MCGUICKIN	0
KIVERIGAN	0	MCKIVERIGAN	0

Table C.1: Adjusted Irish Stem Names (Mc and O')

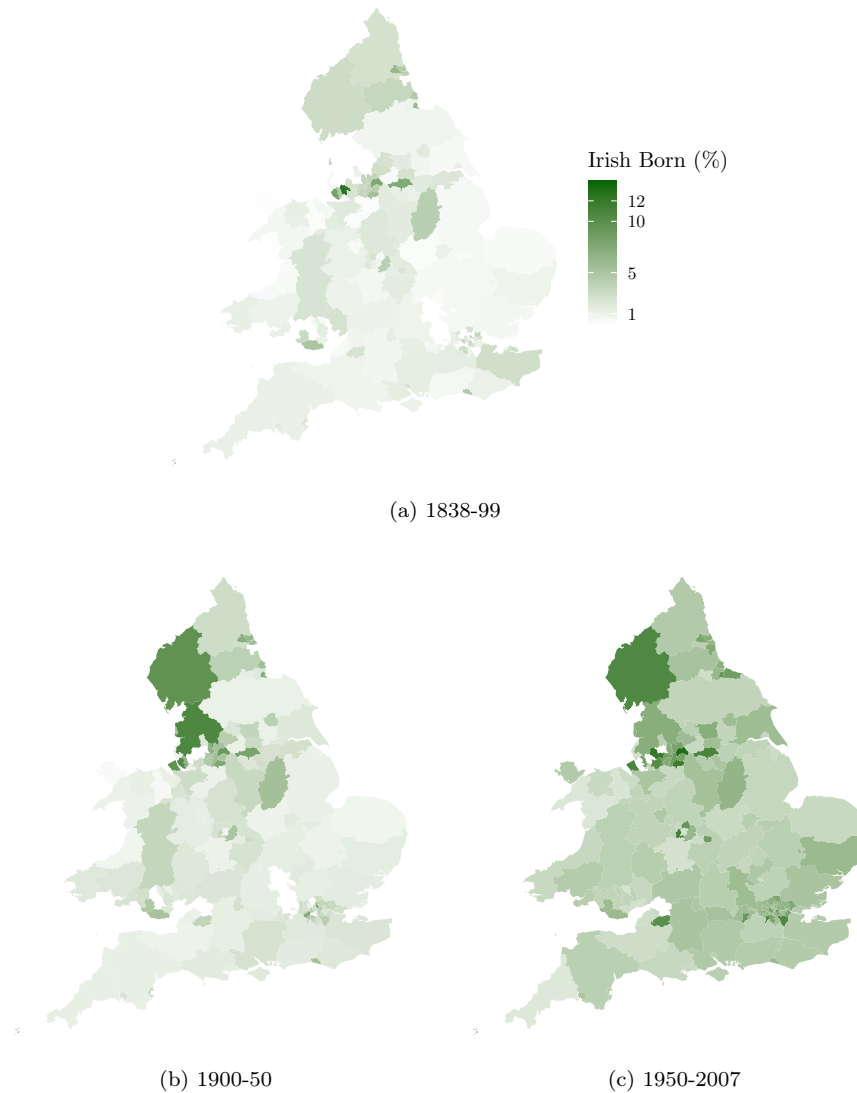


Figure D.1: The Spatial Distribution of the Irish 1838-2007
 Source: 100% sample of Birth Registers.

D Where were the Irish in England?

Where were there Irish? Figure D.1 reports the spatial distribution of the proportion Irish, by registration district, aggregated to local authority area unit of 2018 to ensure spatial consistency over time. Figure D.2 plots the location of the centroid of each registration districts used in this paper's analysis.

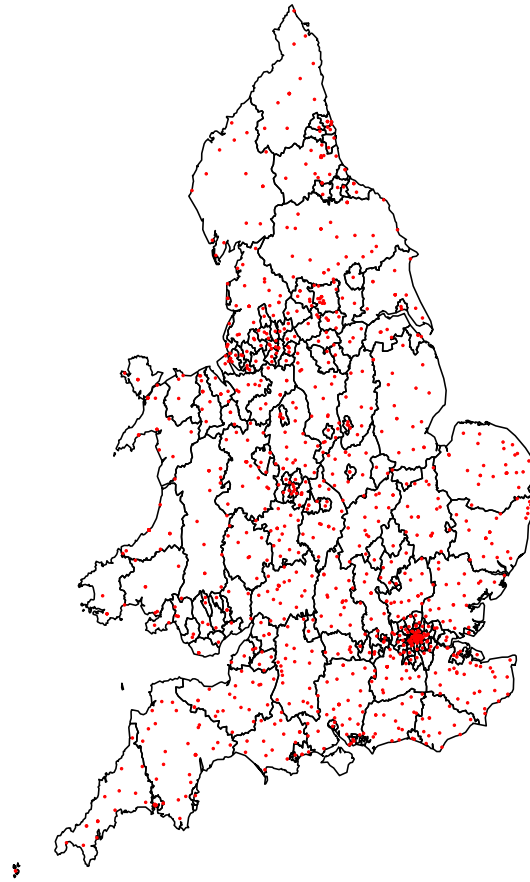


Figure D.2: Registration Districts

Notes: Here we plot the centroid of the registration districts, overlaid with the borders of local authority units of 2018. As many registration districts merge and split (sometimes multiple times) over the sample period, the dots are intended to convey the granularity of the spatial units in a summary way.

E Wealth Regressions

To investigate whether the “Irish” effect on probated wealth is robust when controlling for age at death, we use the linked PPR-Death data to estimate two models. First we look at the extensive margin, the probability of probate (of achieving ‘significant’ wealth at death).

$$Prob(p_i) = \alpha + D_i^F + Age_i + Age_i^2 + \sum D^E \quad (6)$$

where p_i is a categorical variable indicating whether an individual i was probated, α is a constant, D^F is a categorical variable code to one where an individual i has a typically female first name, Age is age at death, and D^E are categorical variables indicating ethnicity of an individual’s surname. The results of this regression are reported in table E.1. Table E.2 controls for district of death.

Table E.1: Probability Probated and Ethnicity, Linked Data: Deaths->PPR, controlling for Age at Death

	Probated (1/0)*100					
	1866-1899		1900-49		1950-1992	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-6.11*** (.03)	-6.18*** (.03)	-5.41*** (.03)	-5.63*** (.03)	-2.10*** (.02)	-.88*** (.02)
Welsh	.05 (.10)	.14 (.10)	1.96*** (.09)	2.55*** (.09)	-.70*** (.05)	-.91*** (.05)
Scottish	.80*** (.08)	1.11*** (.08)	.97*** (.07)	1.22*** (.07)	-1.63*** (.03)	-1.79*** (.03)
Irish	-5.96*** (.08)	-5.37*** (.08)	-9.33*** (.07)	-8.59*** (.07)	-4.89*** (.04)	-5.37*** (.04)
Other	-.86*** (.13)	-.50*** (.13)	-4.38*** (.10)	-4.06*** (.10)	-2.25*** (.05)	-2.25*** (.05)
Age at Death Quadratic?		✓		✓		✓
Observations	3,168,203	3,168,203	7,742,653	7,742,653	14,331,999	14,331,999
R ²	.01	.02	.01	.02	.003	.01

Note: *p<0.05; **p<0.01; ***p<0.001
Linear Probability Model (OLS), English is the omitted Group.

Table E.3 reports the results of the regression

$$\log(w_i) = \alpha + D_i^F + Age_i + Age_i^2 + \sum D^E \quad (7)$$

where w_i is *probated* real wealth. Table E.4 controls for county of death.

Table E.2: Probability Probated and Ethnicity, Linked Data: Deaths->PPR, controlling for Age at Death and District of Death

	Probated (1/0)*100					
	1866-1899		1900-49		1950-1992	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-6.18*** (.03)	-6.32*** (.03)	-5.63*** (.03)	-5.84*** (.03)	-.88*** (.02)	-1.22*** (.02)
Welsh	.14 (.10)	-.15 (.11)	2.55*** (.09)	2.16*** (.10)	-.91*** (.05)	-.31*** (.05)
Scottish	1.11*** (.08)	1.23*** (.08)	1.22*** (.07)	1.33*** (.07)	-1.79*** (.03)	-1.57*** (.03)
Irish	-5.37*** (.08)	-4.79*** (.08)	-8.59*** (.07)	-7.64*** (.07)	-5.37*** (.04)	-4.55*** (.04)
Other	-.50*** (.13)	-.65*** (.13)	-4.06*** (.10)	-3.30*** (.10)	-2.25*** (.05)	-2.77*** (.05)
Age at Death Quadratic?	✓	✓	✓	✓	✓	✓
District Fixed Effects?		✓		✓		✓
Observations	3,168,203	3,168,203	7,742,653	7,742,653	14,331,999	14,331,999
R ²	.02	.03	.02	.04	.01	.08

Note:

*p<0.05; **p<0.01; ***p<0.001

Linear Probability Model (OLS), English is the omitted Group.

Table E.3: Probated Wealth and Ethnicity, controlling for Age at Death

	1866-1899		log(Real Wealth) 1900-49		1950-2007	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-.119*** (.003)	-.163*** (.006)	-.189*** (.001)	-.275*** (.002)	-.181*** (.001)	-.238*** (.002)
Welsh	-.265*** (.006)	-.125*** (.016)	-.145*** (.002)	-.038*** (.006)	-.043*** (.003)	.019*** (.005)
Scottish	.254*** (.008)	.295*** (.012)	.194*** (.003)	.224*** (.004)	.157*** (.003)	.171*** (.004)
Irish	-.009 (.011)	.097*** (.018)	-.145*** (.004)	-.087*** (.006)	-.218*** (.004)	-.181*** (.005)
Other	.490*** (.014)	.616*** (.022)	.237*** (.005)	.339*** (.007)	.307*** (.006)	.326*** (.007)
Age at Death Quadratic?		✓		✓		✓
Observations	1,004,139	345,756	4,691,333	2,146,999	3,703,560	1,941,926
R ²	.006	.013	.006	.022	.007	.013

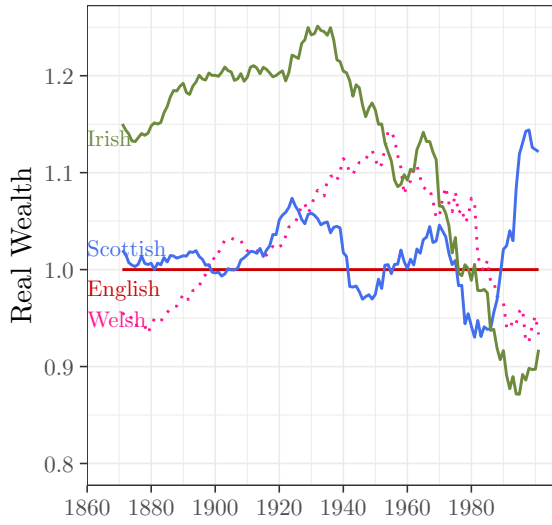
Note: *p<0.05; **p<0.01; ***p<0.001
 OLS, English is the omitted Group.

Table E.4: Probated Wealth and Ethnicity, controlling for Age at Death and County

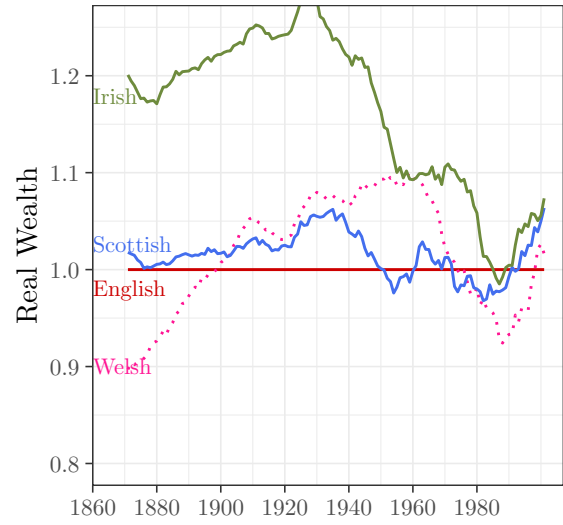
	log(Real Wealth)					
	1866-1899		1900-49		1950-2007	
	<i>OLS</i>	<i>felm</i>	<i>OLS</i>	<i>felm</i>	<i>OLS</i>	<i>felm</i>
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-.163*** (.006)	-.278*** (.010)	-.275*** (.002)	-.287*** (.002)	-.238*** (.002)	-.245*** (.002)
Welsh	-.125*** (.016)	.015 (.031)	-.038*** (.006)	.041*** (.006)	.019*** (.005)	.056*** (.006)
Scottish	.295*** (.012)	.345*** (.021)	.224*** (.004)	.251*** (.004)	.171*** (.004)	.198*** (.004)
Irish	.097*** (.018)	.066* (.031)	-.087*** (.006)	-.065*** (.006)	-.181*** (.005)	-.149*** (.006)
Other	.616*** (.022)	.583*** (.036)	.339*** (.007)	.306*** (.007)	.326*** (.007)	.333*** (.007)
Age at Death Quadratic?	✓	✓	✓	✓	✓	✓
County Fixed Effects?		✓		✓		✓
Observations	345,756	122,704	2,146,999	2,047,462	1,941,926	1,865,598
R ²	.013	.043	.022	.032	.013	.025

Note: *p<0.05; **p<0.01; ***p<0.001
 OLS, English is the omitted Group.

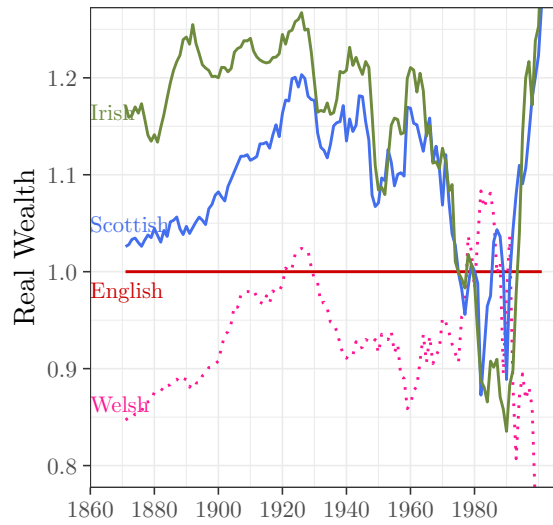
F Extra Empirical Results



(a) The North



(b) The South, ex. London



(c) London

Figure F.1: Regional Differences in the Relative Infant Mortality of the Irish
Note: The English baseline is established by region.