

The Irish in England

January 5, 2024

Abstract

We use the universe of probate and vital registers, from England between 1838 and 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, being on average 50% poorer than the English. Infant mortality was about 25% higher for the Irish between the 1830s and the mid-twentieth century but has subsequently equalized. Sorting, both to urban areas, and to the North of England, 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.

Introduction

Even before the Great Famine of the 1840s, the Industrial Revolution had attracted hundreds of thousands of Irish immigrants to Britain. In 1841 over 0.4 million of them 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. Mass emigration to North America also preceded the Famine. Between 1825 and 1845 nearly 0.9 million made the crossing, the great majority never to return (Williamson (1986); Neal (1997); Darwen et al. (2019); Connolly (2022)).

As Figure 2a illustrates, the annual exodus from Ireland peaked at about 100,000 during and after the Great Famine and the net outflow remained positive thereafter, with small breaks during the wars and in the 1970s, and a shift to net inward migration from the 1990s on (interrupted between 2008 and 2014 by crash of the Celtic Tiger).

Mainly as a result, the population of the island of Ireland (at just over 7m) still falls short today 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 2b). These missing 35 million ‘Irish’ are instead to be found elsewhere, primarily in North America and Britain.

For most of the past 150 years, England has been richer, per capita, than Ireland. England’s inhabitants have lived longer lives, and infants have had a higher chance of surviving their first year, than those living in Ireland, as evident in figure 1.¹ Within England, the Irish have long

¹Adding additional, comparable demographic and macro data for the entire period covered in this paper is problematic for the reasons set out in de Bromhead et al. (2023). According to census data the proportion of the population of Ireland aged 5 and over who could neither read nor write fell from 38.7 per cent in 1861 to

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 2c). It is reflected in the number of Irish-born in Britain, which rose from 419,256 in 1841 to 727,326 in 1851, 805,717 in 1861 before falling thereafter 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 2d). 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 history of the Irish in Britain was not, by and large, one 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)).²

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 such as low status white-collar occupations and foremen 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 focus here is not just on the Irish-born but on what used to be described a century ago as “the Irish race”, i.e. both those from Ireland and those of Irish extraction. 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 between 1838 and 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.³

We measure status in two ways; wealth at death, and infant mortality. In this way we capture ethnic inequality both at the start and end of life. The results are stark. From at least

11.9 per cent in 1911. Comparable data are unavailable for England, but the proportion of grooms (brides) who signed the register at marriage rose from two-thirds (half) in 1840 to 97 percent in 1900 (Mitch (1983)). This suggests considerable, but incomplete convergence in literacy in the second half of the nineteenth century.

²The 2021 census of England and Wales tells a more positive story, however, as discussed below.

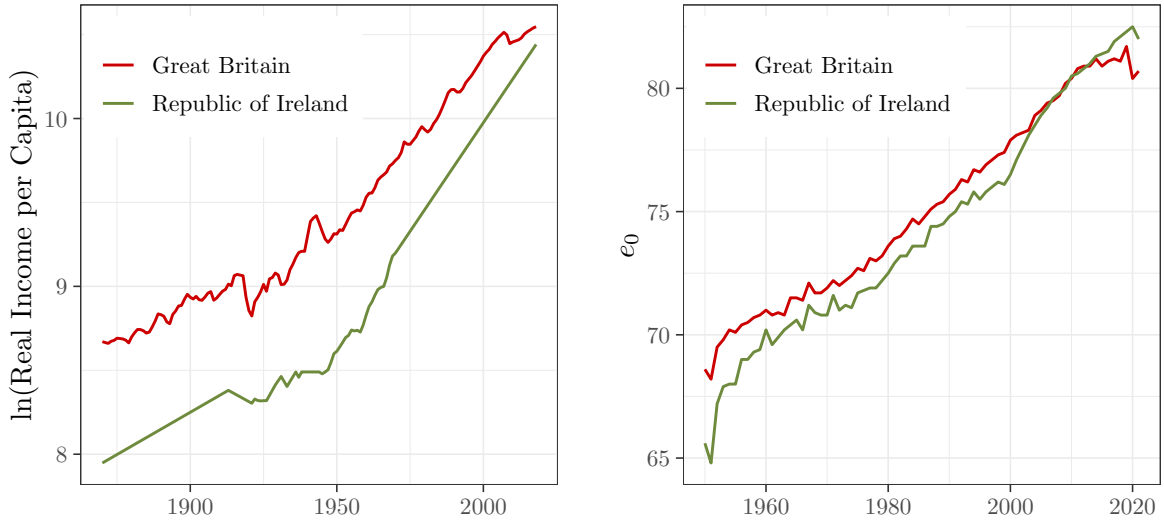
³For robustness, we also calculate ethnicity in this way using Onomap, a contemporary classification system based on billions of global records (Appendix). We also check our results using varying thresholds for the proportion of the holders of a surname born in Ireland in the 1911 census of England and Wales (Appendix).

the mid-19th century until very recently the ‘Irish’ in England have persisted as an underclass. We document the lower wealth, and higher infant mortality, of those with distinctively Irish surnames. Using linked data we show that this Irish effect is robust to age controls. Lower Irish wealth is therefore not an artifact of the return migration of richer, older Irish to Ireland. The ‘Irish’ were always poorer than the English, and this pattern is persistent, although not constant, between the 1850s and the end of the last millennium.⁴ We show that the Irish wealth penalty is driven by the experience of the Irish in the North of England. Half of the Irish infant mortality effect is sorting into higher mortality districts.

Only in the 1980s did Irish infant mortality fully converge with the that of the native English, and both the average wealth of the Irish at death and the wealth of the wealthiest among them still lags behind the English today. However, data on education and socioeconomic status in the latest (2021) census suggests that full convergence between the Irish and the English in terms of wealth at death is – finally – not far away.

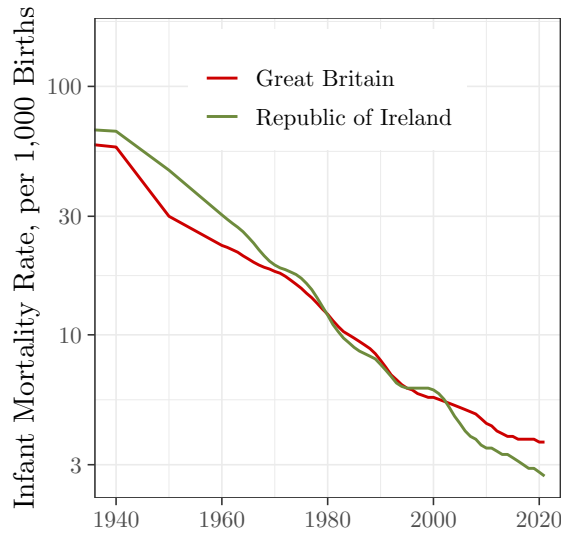
The paper proceeds as follows. The data are presented in section and the methodology in section which describes in detail the process and accuracy of the surname ethnicity assignment, and the construction of the three wealth measures, and the infant mortality rate. Section presents the results. What 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 . Section concludes.

⁴Earlier 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.



(a) Real Income per Capita

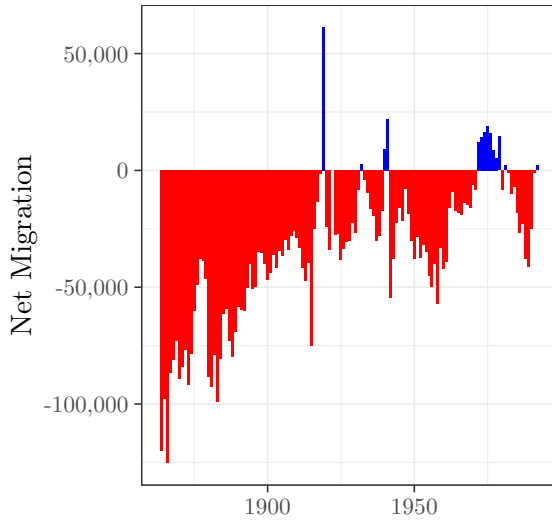
(b) Life Expectancy at Birth



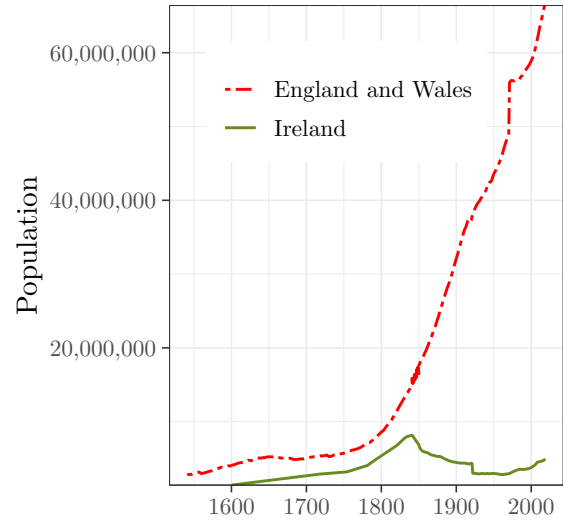
(c) Infant Mortality Rate

Figure 1: Comparison of Income per capita, life expectancy and Infant Mortality, Ireland and Great Britain, 1870-2020

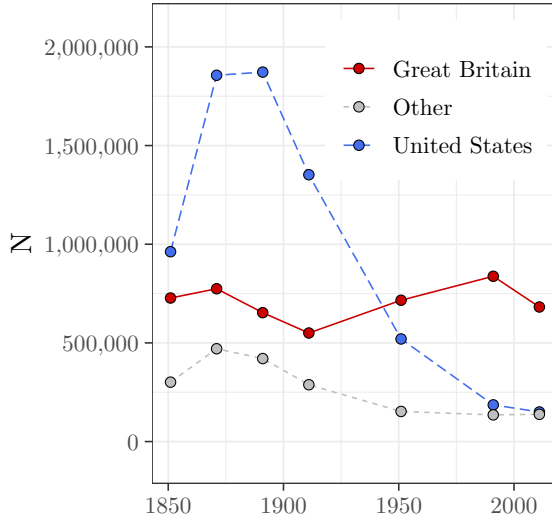
Source: Panel A, Maddison database, Central Statistics Office Ireland (CSOI); Panel B: Our World in Data (link); Panel C: CSOI (link). This figure describes English and Irish life expectancies at birth and infant mortality ratios, as well as British GDP per capita and a measure of Irish income, for the period since independence. The income data were taken from the Maddison database, adjusted for the assumption that Irish income per capita was about 90 per cent of British c. 2020 (as implied by corrected consumption data). The estimates of Irish infant mortality in Panel C are too low before the 1940s for reasons given in de Bromhead et al. (2023). Clearly the demographic data tell a more optimistic tale than the national accounts.



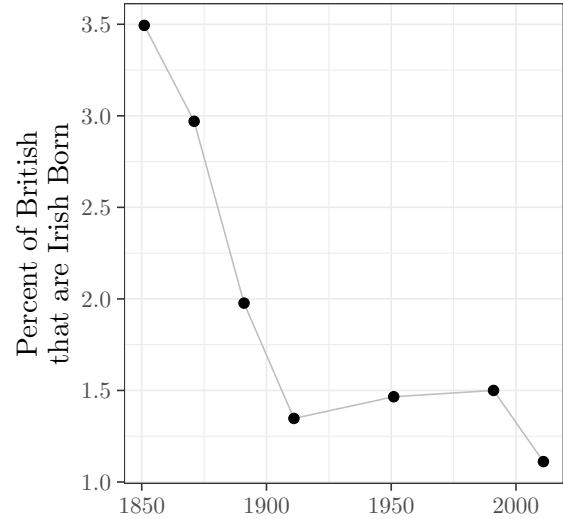
(a) Net Migration, Ireland, 1850-1992



(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 2: 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.

Data

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.⁵ 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).⁶

The PPR Calendar data were 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 £50,000.⁷

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 13, reported in the appendix, compares the numbers collected versus that recorded by the official records (from Office for National Statistics (2021b)), by year. They match very closely.

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.

Methodology

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.⁸ 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.

This has the unavoidable drawback of imposing an ethno-cultural identity on some who might not self-identify as Irish, such as the distant descendants of mixed marriages. It is certainly the case that the accuracy of surnames as an indicator of ethnicity is likely to decline from generation to generation. Moreover, except at the outset, most of those identified as ‘Irish’ in this analysis were born in England rather than in Ireland. The frequency of “Patrick” and “Bridget” as forenames amongst second generation Irish migrants of the 19th century was significantly lower than that of the first generation (declining to 1.5% from 10% for both names), suggesting at

⁵The 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).

⁶ Appendix figure 12 for a reproduction of some of these comparisons over time, from Cummins, 2021.

⁷See appendix section for more detail on this.

⁸Overviews of the use of surnames to infer ethnicity, in the social sciences and genetics, are given in (Mateos, 2007; Mateos et al., 2011).

minimum some aspirational assimilation (Smith and MacRaild (2009) compare Connor (2021)).
9

We use the 19 million de-anonymized individual adult records in the special access version of the 1911 census to examine the distribution of place of birth for the over 500,000 surnames that it includes (Schurer and Higgs (2021))¹⁰. We confine our focus to adults so that the fertility of recently arrived immigrants does not skew the immigrant surnames to be categorized incorrectly as “English”.¹¹ Table 1 ranks the top 25 countries by numbers of birth listed in 1911. Nearly 90% of those enumerated were born in England, 6.2% were born in Wales, 1.8% in Scotland, 1.8% in Ireland. All other countries each represent far less than 1%.

Table 1: Distribution of 1911 Adult Census Population by Country of Birth

Country	N	%
England	16,164,030	88.01
Wales	1,145,890	6.24
Scotland	333,534	1.82
Ireland	325,508	1.77
Russia	59,694	0.33
Germany	50,380	0.27
India and Pakistan	39,903	0.22
France	32,582	0.18
United States	25,650	0.14
Isle of Man	24,627	0.13
Australia	16,572	0.09
Italy	15,578	0.08
Poland	15,576	0.08
Canada	12,302	0.07
Austria	10,377	0.06
Switzerland	8,825	0.05
Netherlands	6,908	0.04
South Africa	6,743	0.04
At Sea	5,734	0.03
Sweden	5,207	0.03
Norway	4,687	0.03
Belgium	4,469	0.02
New Zealand	4,102	0.02
Denmark	3,923	0.02
Malta	3,363	0.02

Source: 1911 Census

Based on table 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’ birth countries. We therefore 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 was before 1947), Germany, France, Italy, Poland and the Netherlands. How can we know whether a given surname corresponds to a origin country? Table 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, Brown (Scottish), Cohen (Russian), Murphy (Irish), Durand (French), Van Gelder (Dutch), Becker (German), Posner (Russian and Polish) would all incorrectly be classified as English. Historical migration patterns here skew a simple rule.

To attribute a surname to an ethnic origin we therefore cross-reference the complete matrix of surnames by country of birth (as represented by the example surnames in table 2), with the average proportions born in England from table 1. Then we implement a two-step process to classify each surname. First, we attribute to each an ethnic origin based upon the most frequent

⁹In table 3 we find 63% of those with the surname “Murphy” born in England in the 1911 census, with 26% born in Ireland, 67% of Kellys born in England, 17% in Ireland, and 65% of “Ryans” born in England, 25% in Ireland, and 63% of “O’Neills” born in England, 26% in Ireland. A proportion of these will be English women who take their Irish husband’s surname upon marriage. Similarly, some Irish women will lose their Irish surname upon marriage. Thus, we interpret the proportion of a surname born in Ireland as fuzzy indicator of “Irishness” and adopt varying thresholds in assignment.

¹⁰We count those aged 20 and above as adults. We include both men and women.

¹¹An example of how such a process would occur 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.

Table 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	2,549	.889	.045	.004	.013	.000	.000	.003	.001	.000	.000	.000
Davies	118,823	.356	.601	.005	.003	.000	.000	.001	.000	.000	.000	.000
MacDonald	4,344	.632	.022	.237	.045	.000	.001	.012	.001	.000	.000	.000
Murphy	12,436	.627	.049	.018	.263	.000	.000	.005	.000	.000	.000	.000
Cohen	6,446	.412	.004	.004	.006	.382	.022	.000	.001	.000	.100	.006
Becker	425	.527	.012	.012	.014	.056	.256	.005	.012	.000	.007	.002
Singh	138	.022	.000	.000	.000	.000	.000	.971	.000	.000	.000	.000
Durand	152	.539	.007	.007	.007	.000	.000	.059	.263	.000	.000	.000
Ferrari	117	.316	.017	.000	.000	.000	.000	.000	.000	.385	.000	.000
Posner	140	.386	.000	.000	.007	.307	.036	.007	.000	.000	.214	.000
Van Gelder	48	.667	.000	.000	.000	.021	.000	.000	.000	.000	.000	.208

Note: Calculated from the 1911 census. **Bolded** text indicates assigned ethnicity.

country of birth. Where there is a country other than England or Wales which accounts for 20% or over of the births of that surname, we update the ethnic origin to that country. For example, Murphy is attributed as “English” in step one, as 62.7% of the holders of the Murphy surname in the 1911 census of England and Wales are born in England. But in step two, we update Murphy to “Irish” because 26.3% of Murphys are born in Ireland, as 26.3% is above the 20% threshold. This procedure works to attribute correctly all of the example surnames in table 2). For robustness, we also construct assignments based on 15%, and 25% thresholds.¹²

Table 3 presents the matrix of the proportions of adults enumerated in the 1911 census of England and Wales by birth country for a set of popular Irish surnames. We took the top 15 Irish surnames as listed at <https://forebears.io/ireland/surnames>, a website that has aggregated a considerable volume of data on contemporary global surname distributions. Our 20% threshold categorizes 12 of these 15 surnames as “Irish” while our 15% threshold categorizes 14 of the 15 as “Irish”.

Table 3: Irish Surnames from the 1911 Census

Surname	N	Country of Birth										
		England	Wales	Scotland	Ireland	Russia	Germany	India	France	Italy	Poland	Netherlands
Murphy	12,436	.627	.049	.018	.263	.000	.000	.005	.000	.000	.000	.000
Kelly	16,388	.667	.025	.037	.174	.001	.000	.004	.001	.000	.000	.000
Byrne	3,637	.604	.015	.016	.329	.000	.000	.006	.000	.001	.000	.000
Walsh	11,053	.742	.028	.028	.167	.001	.000	.003	.000	.000	.000	.000
Ryan	6,055	.646	.046	.014	.248	.000	.000	.006	.000	.000	.000	.000
O’Brien	5,793	.569	.076	.014	.295	.000	.000	.008	.001	.000	.000	.000
O’Connor	2,833	.590	.034	.016	.315	.000	.001	.005	.001	.000	.000	.000
O’Sullivan	646	.303	.084	.005	.573	.000	.000	.014	.000	.000	.000	.000
Doyle	4,188	.640	.034	.021	.273	.000	.001	.004	.000	.000	.000	.000
O’Neill	3,016	.631	.043	.025	.261	.000	.001	.005	.001	.000	.000	.000
Lynch	3,957	.668	.043	.019	.228	.000	.001	.005	.001	.001	.000	.000
McCarthy	5,324	.622	.109	.011	.213	.000	.000	.004	.001	.000	.000	.000
Brennan	2,346	.652	.021	.015	.277	.000	.001	.004	.000	.000	.000	.000
Dunne	910	.502	.012	.010	.436	.000	.001	.010	.000	.000	.000	.000
Murray	12,050	.725	.025	.108	.095	.000	.001	.007	.001	.000	.000	.000

Note: Calculated from the 1911 census for the Top 15 Irish Surnames listed at forebears.io

This historical method to infer ethnicity can be expected to lose a degree of contemporary ‘ethnic accuracy’ over time. In other words the informational content of surnames in relation

¹²Upon inspection, it was apparent that this method incorrectly assigned many Welsh surnames as “English” (e.g. Edwards and Hughes). This is because of the very unequal population sizes of the two neighboring countries. 42% of Hughes and 28% of Edwards are born in Wales. We therefore updated a surname to “Welsh” if more than 25% of the holders of a surname were born in Wales. As 6.2% of the population of England and Wales was born in Wales (table 1), the 25% cutoff here implies that the holders of a “Welsh” surname are at least 400% more likely to have been born in Wales than the average English.

to ethnic origin is likely to be less in 2000 than it is in 1900. Through inter-marriage, surnames that we assign as ‘Irish’ in 1911 are likely to have an increasing English ancestry, as those who we assign as ‘English’ will have more Irish ancestry, over time. But this is precisely what we want to observe. If assimilation is rapid and complete, we should see full convergence in socio-economic outcomes. To complement our historical ‘Irish’ analysis we also examine the status of the ‘Irish’ using a modern assignment of ethnicity, Onomap, in appendix section .

Our approach has another limitation. Since we rely on distinctively ‘Irish’ surnames, we necessarily exclude the minority of Irish people with ‘non-Irish’ surnames. They would have accounted for perhaps one-fifth of emigration from Ireland to England, and would most likely have fared better than the majority. Our results, then, refer to the four-fifths or so of Irish of Gaelic/Catholic background.

Wealth Calculations

We first analyze the relative wealth of the Irish using three measures: 1. the probate rate, 2. average wealth, and 3. the representation of a group in the top 1% of wealth-holders.¹³ 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.¹⁴ 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

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

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.

¹³We 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.

¹⁴As age at death is only recorded in the death registers from 1866. Therefore to calculate N_{20} for each ethnicity we calculated $\frac{N_{20}^{1866-76}}{N}$ for all deaths 1866-76, then used this ratio to infer N_{20}^e for ethnicity e by calculating $N_{20}^e = N * \frac{N_{20}^{1866-76}}{N^{1866-76}}$.

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 for 1858-1992, and 12,486,026, for whether an individual is probated for 1996 to 2018.

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.¹⁵ 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 year¹⁶, 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 4 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.¹⁷

Table 4: 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

¹⁵As noted in section , 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.

¹⁶For example there are 285 “Elizabeth Jones” dying in 1905.

¹⁷Note 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.

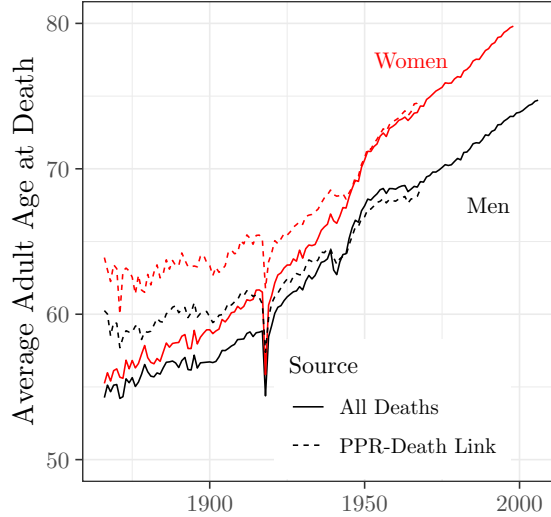


Figure 3: Age at Death over Time, All Deaths and Linked Probate-Deaths

Figure 3 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.

Infant Mortality

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

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

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.¹⁸ 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 4 compare the resulting individual level estimate of the 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.

¹⁸We assume that infant deaths are registered in the same district as their birth.

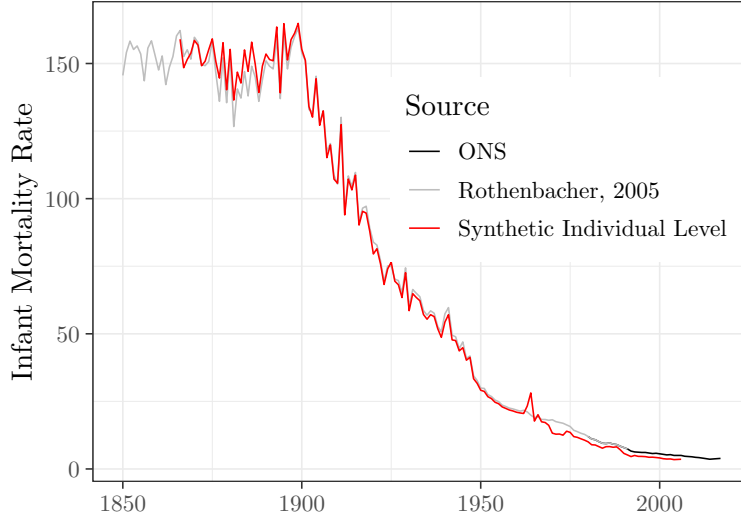


Figure 4: Comparison of Infant Mortality Rates

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

Results

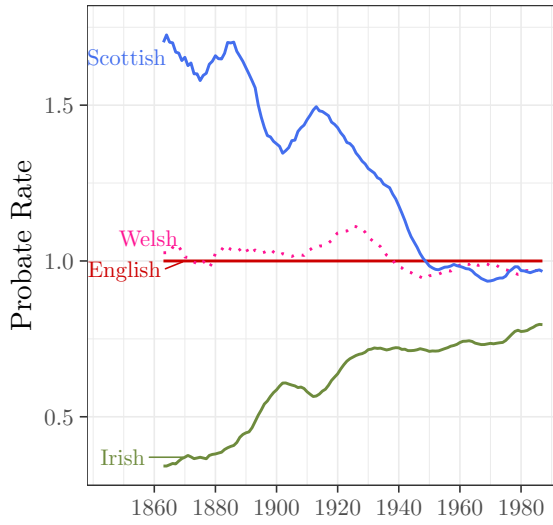
Wealth

Figure 5 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.

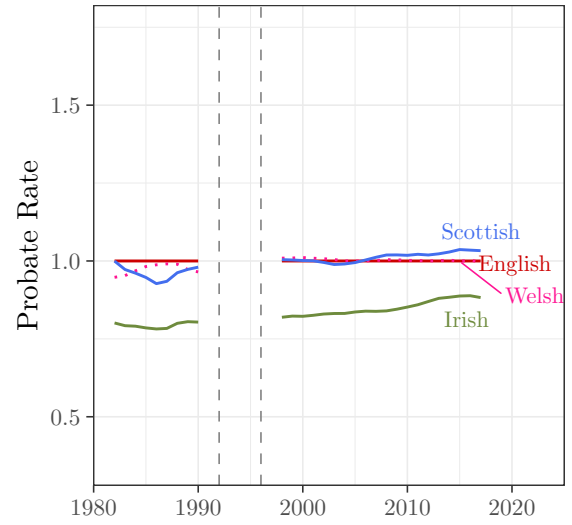
The Irish in England did not share in this convergence, at least until very recently. Throughout they had lower probate rates, lower average wealth, and lower probabilities of being in the top 1%. The Irish ‘effect’ was persistent throughout. The disadvantage was not constant over time, however. Between 1858 to 1990 the proportion of Irish probated was always at least 20% lower than the English, but it rose sharply between the 1860s and the 1920s, and then stagnated for half a century before rising again in the 1970s and 1980s. Average wealth for the Irish fluctuated less. The ratio rose between the 1870s and the Great War, but stuck thereafter at about 75% that of the English, while the relative Irish probability of being in the top 1% rose up to the early 1920s and declined thereafter to about one-third c. 1980.

We therefore can periodize the speed of Irish-in-England wealth convergence into three distinct phases. The first, 1860-1900, is a period of consistent year-upon-year convergence. After 1900, and excluding some oscillations associated with the first World War, and a significant rise in the probate rate 1920-30, there is no change in the relative wealth status of the Irish to the English all the way to 1980. After 1980 we see significant but not complete convergence.

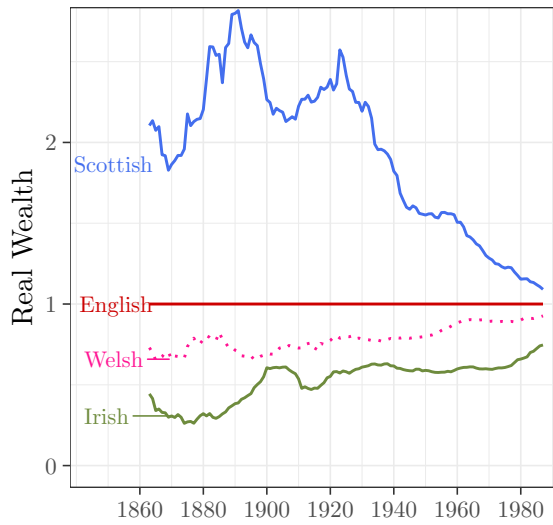
Figure 6 compares the wealth distributions of wealth of the British and Irish. The two prominent peaks in all plots are a result of the attribution of inferred wealth to those who



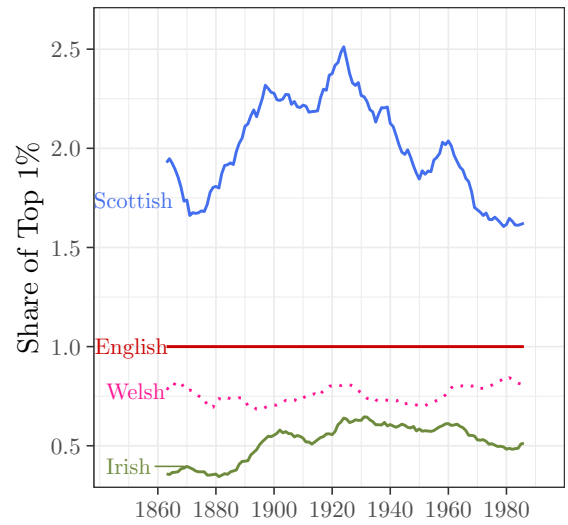
(a) Probate Rate, 1858-1992



(b) Probate Rate, 1996-2019



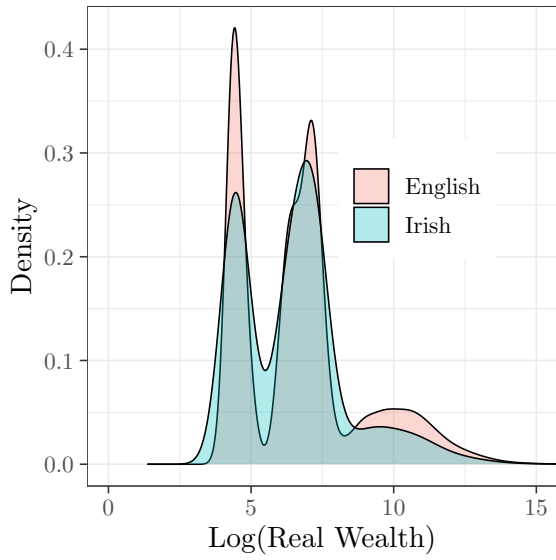
(c) Average Wealth



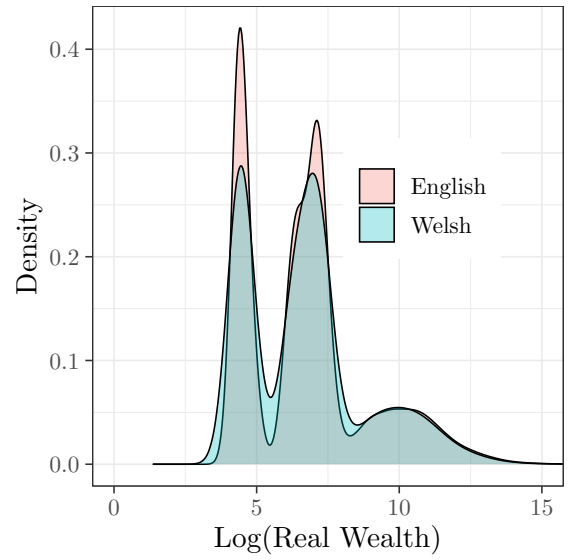
(d) Share of the Top 1%

Figure 5: The Wealth of the Irish and British, 1858-2018

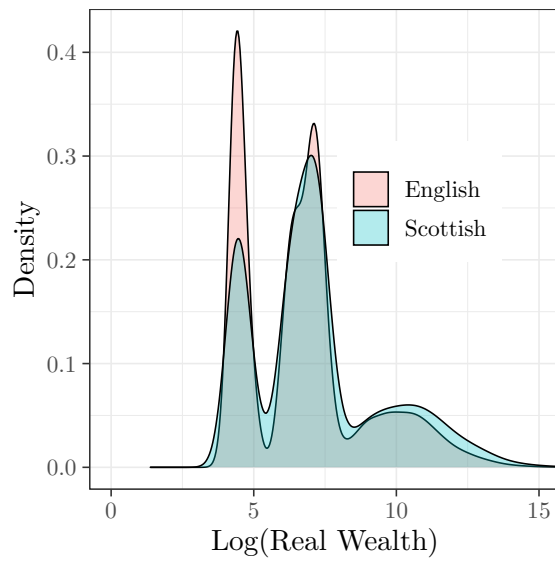
Notes: English surnames are set to one in all figures.



(a) Irish



(b) Welsh



(c) Scottish

Figure 6: The Distribution of Wealth, British and Irish

die with wealth below the probate threshold. As can be seen from panel (a), which compares the English and Irish, the share of top wealth holders amongst the Irish is lower. The Irish underrepresentation in the top 1%, as reported in Figure 5 (d) is apparent at every moment of the wealth distribution. This is not the case for the Welsh and the Scottish.

Wealth controlling for Age at Death and Place of Death

Might the wealth patterns reported above have been the product of the more successful among the Irish spending their working lives in England and then returning home? While there has always been return migration, neither the representativeness nor the motivations of the returnees are known. However there is a presumption that the paucity of welfare networks for the elderly in Ireland in the past made it less likely for the very poor to return Malcolm (2006, 107–9). Moreover, the numbers of elderly returnees were usually modest: between 1946 and 1991, total net migration of those aged 65 and over averaged 1,400, while in 1991 the total number who had returned from England and Wales at some point when aged 65 or above was 6,265 (Malcolm (2006, 38,40)). These would have been very small proportions of the Irish population in England, first generation and other, at any one time.

The following exercise adds further perspective. 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 older and richer Irish were not in England, but in Ireland, when they died. The Irish who died in England could be a younger population with mechanically lower wealth because of life-course effects. To assess this we use the linked PPR-Death data, as described in section 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 5 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 ??.

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 5 reveals that the Irish ‘effect’ is only very modestly reduced by the inclusion of age at death controls.¹⁹ Further, in all sample periods, the Irish coefficient is statically indistinguishable where age controls are used, or not. average wealth. Appendix section decomposes the Irish wealth effect.²⁰

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

Infant Mortality

Infant mortality rates are a sensitive indicator of a population’s material living standards and health (Huck (1995); Baird et al. (2011)). Figure 7 presents the pattern of infant mortality for

¹⁹Note that this contrasts with the effect of place on infant mortality, as reported in table 6.

²⁰Tables 10 and 11 examine the probability of being probated, controlling for age at death and county of death. Tables 12 examine *probated* real wealth, controlling for age at death and county of death.

Table 5: Wealth and Ethncity, Linked Data: Deaths linked to 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	.04*** (.01)	.02** (.01)	.14*** (.004)	.13*** (.004)	.01*** (.003)	.03*** (.003)
Scottish	.09*** (.01)	.08*** (.01)	.05*** (.005)	.04*** (.005)	-.09*** (.003)	-.09*** (.003)
Irish	-.43*** (.01)	-.39*** (.01)	-.43*** (.004)	-.38*** (.004)	-.25*** (.002)	-.22*** (.002)
Other	-.08*** (.01)	-.08*** (.01)	-.20*** (.01)	-.17*** (.01)	-.11*** (.003)	-.12*** (.003)
Age at Death Quadratic?	✓	✓	✓	✓	✓	✓
District Fixed Effects?		✓		✓		✓
Observations	3,155,344	3,155,344	7,716,522	7,716,522	10,430,698	10,430,698
R ²	.02	.03	.02	.05	.005	.02

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

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 from 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 were much higher in urban areas during the 19th century (Woods (2000)). Was 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 . We then ran a linear probability model of infant death (D^{ID}) on the set of ethnic (E) and registration (R) district dummies (D) as

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

(note that we multiple the infant death dummy (D^{ID}) by 1000 for ease of interpretation). Table 6 shows that about 50% of the Irish infant mortality effect is due to sorting between registration districts. Of 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 suspect that the majority of this Irish mortality penalty could be due to geography.

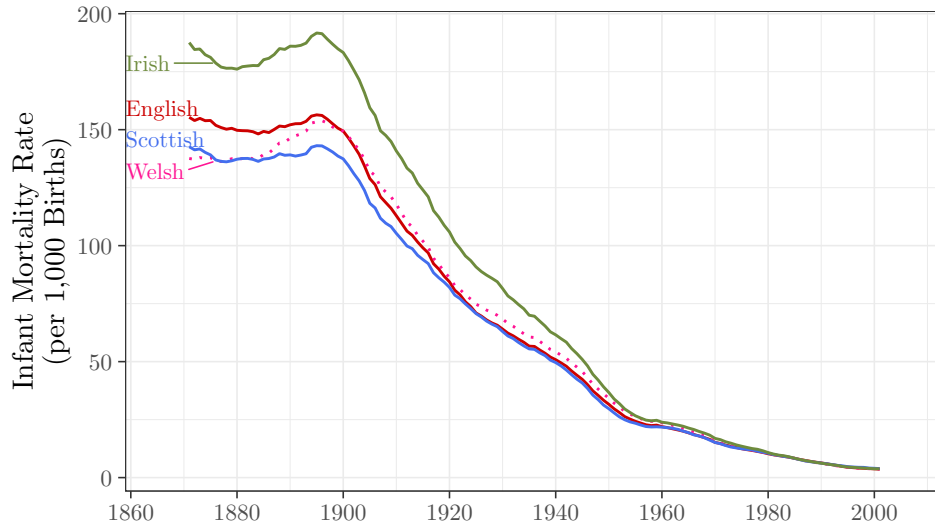
Table 6: Infant Mortality and Ethnicity, controlling for Place

	Died as an Infant*1000					
	1866-1899		1900-49		1950-2007	
	(1)	(2)	(3)	(4)	(5)	(6)
Scottish	-13.54*** (0.71)	-20.62*** (0.71)	-3.93*** (0.42)	-7.39*** (0.42)	-0.18 (0.13)	-0.42** (0.13)
Other	-12.65*** (0.93)	-28.12*** (0.95)	-15.01*** (0.48)	-22.61*** (0.50)	2.08*** (0.11)	1.21*** (0.12)
Irish	30.70*** (0.50)	13.63*** (0.51)	19.37*** (0.29)	9.52*** (0.30)	0.98*** (0.09)	-0.12 (0.09)
Welsh	-11.47*** (0.24)	-4.90*** (0.29)	3.10*** (0.16)	0.58** (0.18)	0.50*** (0.07)	0.36*** (0.07)
English Average	152.75		84.82		12.94	
District Fixed effects?		✓		✓		✓
Quadratic Time Trend?	✓	✓	✓	✓	✓	✓
Observations	28,685,192	28,685,192	38,237,313	38,237,313	39,317,853	39,317,853
R ²	0.0002	0.01	0.01	0.02	0.005	0.01

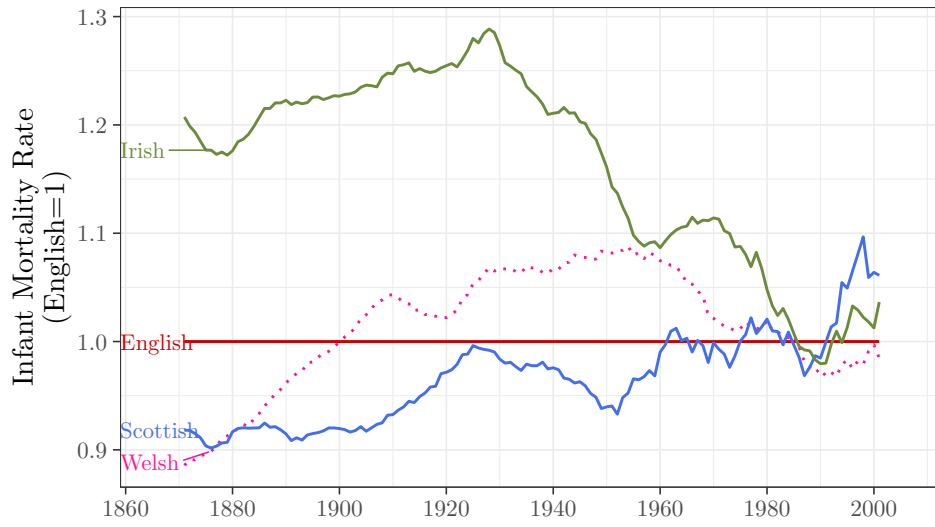
Note:

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

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



(a) Levels



(b) Indexed to "English"

Figure 7: The Infant Mortality Rate, Major Ethnicities, 1866-2007

Regional Differences

The Irish in England have been significantly poorer at death, and until recently faced higher mortality for their infants than the English. The Irish were disproportionately urban but as we show in section , this does not explain all of the disadvantage. The possibility remains 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.²¹

Figure 8 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.

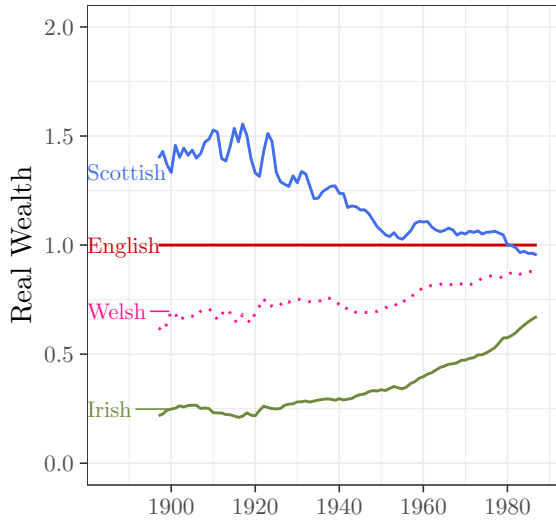
In the North, the Irish had 25% of the wealth of the English in the 19th century. This rose to about 70% by 1992. The scale of this wealth gap dwarfs that of the South, and that of London.²² It is also worth noticing here that the Scottish over-performance is not present in the North.

Figure 8 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 9. 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.

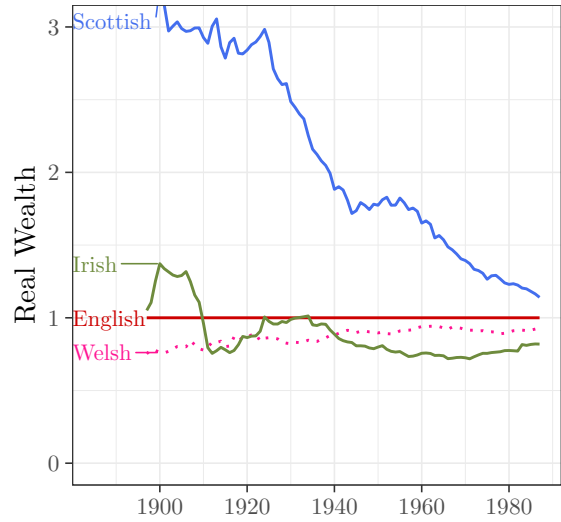
However, as revealed in figure 8 the Irish were far poorer relative to the English in the North than they were in the South, or in London. The specific social, economic and cultural conditions of the North resulted in Irish migrants being *much* poorer than the English. Thus, the underperformance of the Irish in England is a result not only of disproportionate migration to the poorer North but even more so an underperformance driven by the specific experience of the Irish in the North.

²¹The North is comprised of the counties of Cheshire, Cumberland, Durham, Lancashire, Northumberland, Westmorland, and Yorkshire.

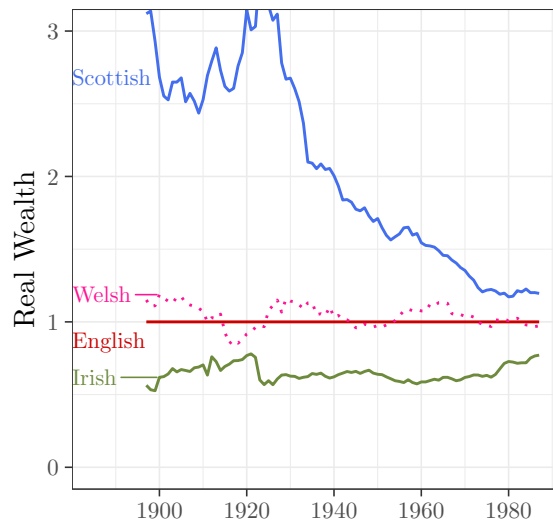
²²The infant mortality rates, reported in appendix figure 19, do not display the regional patterns of the wealth figures. Here we speculate that the urban penalty faced by migrants to both the North and South, masks the status effect picked up by the wealth data.



(a) The North



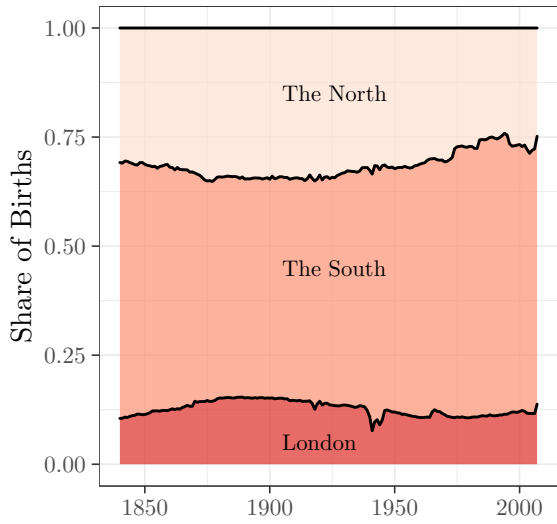
(b) The South, ex. London



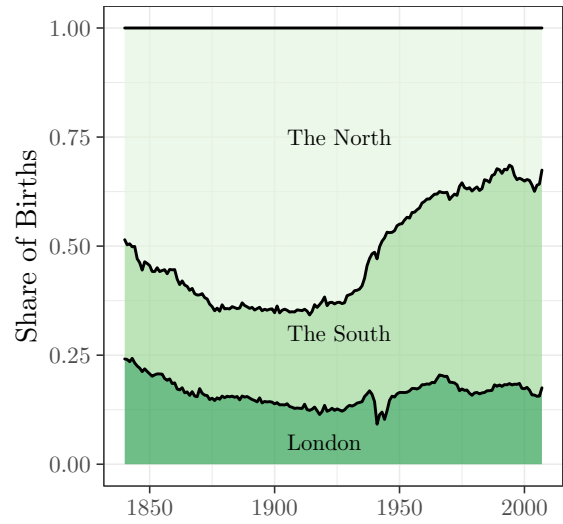
(c) London

Figure 8: Regional Differences in the Relative Wealth of the Irish

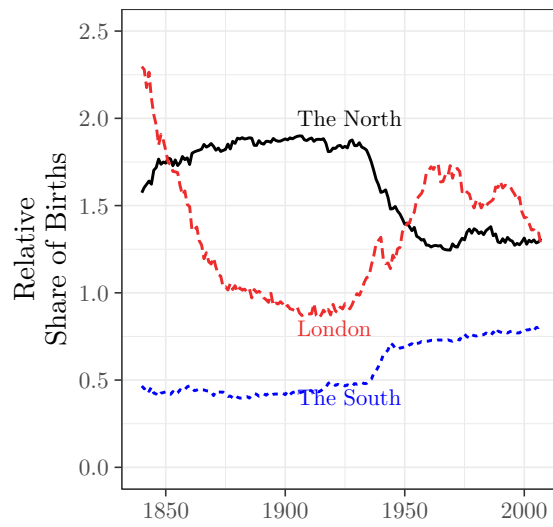
Notes: The English baseline is established by region. Ethnic classification is based upon the 1911 census. See figure 9 and 20 for the breakdown of the share of births by ethnic origin of surname.



(a) English



(b) The Irish



(c) Irish relative to English

Figure 9: The Regional Distribution of Births, the English and the Irish Compared

Discussion

This paper has identified a large and persistent, though also time-varying, 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? What explains why some periods were better for convergence than others? Why during much of the twentieth century did the gap in infant mortality rates decline, while that in probate wealth did not? Our main objective in this paper has been to discover and describe outcomes rather than to explain them. We conclude with an informal discussion of some possible explanations, but leave a formal analysis of their causes for another day.

First, as infant mortality is measured contemporaneously, and wealth is measured at death, it could be that the incomplete convergence of Irish wealth relative to English by 2020, and the convergence of infant mortality by the 1980s, are telling the same story but with a lag on wealth. This remains to be seen, but very recent data (from the 2021 census (Office for National Statistics (2023))) on the self-declared health, educational status, and occupational profile of those of Irish ethnicity broadly corroborates.²³

Second, could the Irish simply have been sending their wealth home? This seems less likely. 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 per head 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.²⁴

Third, there is the issue of selection, a key factor in the broader literature on the economics of migration. Over two centuries ago, in a much-cited passage, Adam Smith wrote of migration from Ireland to England:

“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. Certainly, Irish immigrants to England continued to be overwhelmingly working-class long after Adam Smith’s time; in 1911, nearly four-fifths of Irish-born male workers were unskilled labourers and 86 per cent of females were in ‘domestic service’ (Glynn (1981, p.58)). Moreover, in 1911 the censuses also suggest that the percentages with skilled occupations such as blacksmith, grocer, 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.²⁵ Such a straightforward descriptive exercise for earlier censuses would probably lend more empirical ballast to Smith’s observation from an earlier era.²⁶

²³Once differences in age-distributions are controlled for the health status of the Irish and the British are almost identical, while the proportion of Irish with third-level or higher qualifications is higher at all ages. See Office for National Statistics (2023).

²⁴For 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).

²⁵Personal communication from John Fitzgerald, Trinity College Dublin.

²⁶The 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

There was selection too among emigrants to England relative to those who made their way to America; the more affluent made their way to America, while successive cohorts of the less affluent joined an English working class that was not upwardly mobile either. And this may help explain why those who arrived in the United States were less predominantly working class, and why they were more upwardly mobile than their compatriots in England and their descendants (Collins and Zimran (2019); Anbinder et al. (2019, 2022)).

However, the fact that the Irish found themselves in lower status occupations, poorer at death, and facing higher infant mortality rates for their children, is not proof of negative selection.

There also remains the possibility of anti-Irish discrimination in the labour market, in health services, and the generally unwelcoming, if not outright hostile, social landscape (see for example Winder 2010). A recent study of the treatment of Irish defendants tried in London’s Old Bailey finds that during the nineteenth century they were more harshly treated by juries than others in terms of likelihood of conviction and sentencing (Bindler et al. (2023)). In the case of coal-mining, studied by MacRaild (2010), one might have assumed that the Irish would have achieved parity of status eventually, but that was not so, at least before 1880; 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. Within Britain, the Irish have long been the ‘other’ ethnic group. Writing in 1870, at a time when his links to the Irish community in England were closest, Karl Marx declared:²⁷

[...] 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.

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, 162-3), De Nie (2004)). Not 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)). But 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)). 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)).

In sum, this Irish status effect could reflect selection, discrimination, and slow 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

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. Finally, there is the option of using the socio-economic ranking of surnames in terms of occupational categories and location (as in Connor (2020)) as a means of proxying selection. But surname rankings may vary over time, while accessible data survive only for 1901 and 1911, so we simply note this possibility here.

²⁷As cited in Marx and Engels (Marx and Engels (1971, p.254)), ‘confidential communication’, 28 March 1870.

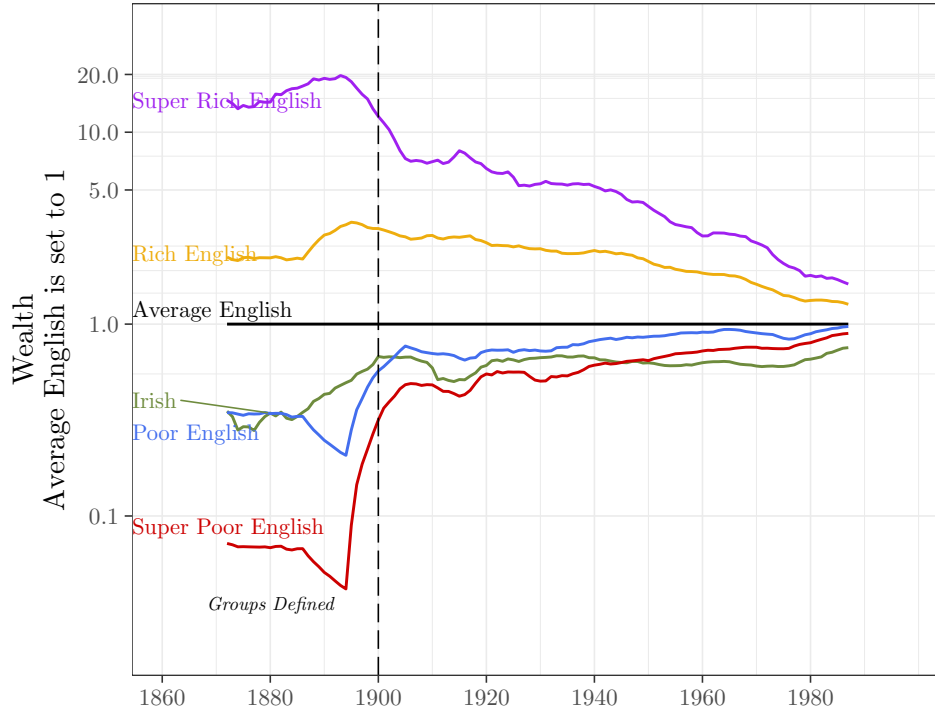


Figure 10: A Distinctive Irish Wealth Pattern

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.

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 10 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 10 compares the Irish to this set of English wealth groups. It shows that between 1858 and 1992 the the Irish only very modestly regressed towards English mean wealth, but at a much slower rate than any of the English wealth groups. In fact between 1920 and 1992, there was really no movement in the relative wealth of the Irish. Social mobility was 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)). Or were the Irish special in this respect?

We conclude with a final reflection on the nature of the selectivity of migration from Ireland. As noted earlier, the evidence presented here and in the wider literature is consistent with migration from Ireland to England for most of the period under review being negatively selected. Perhaps the relentless addition of young, poorly educated immigrants for much of the twentieth century to the stock of Irish in England helps to explain the persistence of Irish non-convergence, as in figure 10. 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.

In recent decades, the profile of Irish immigrants to England has changed, and data in the 2021 census of England and Wales reflect this (Office for National Statistics (2023)). While the self-reported health of the middle-aged and elderly among the ethnic Irish (self-declared, but presumably mostly Irish-born) continues to lag behind that of the English, the proportion of Irish with third-level education considerably exceeds the proportion of English, particularly at younger ages. This suggests that in decades to come the economic wellbeing of Irish in England, as reflected in probate data, may fully converge to, or surpass, that of the English.

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).
- Anbinder, Tyler, Cormac Ó Gráda, and Simone A. Wegge**, "Networks and Opportunities: A Digital History of Ireland's Great Famine Refugees in New York," *The American Historical Review*, 12 2019, *124* (5), 1591–1629.
- , —, **and** —, "'The Best Country in the World': The Surprising Social Mobility of New York's Irish-Famine Immigrants," *Journal of Interdisciplinary History*, 2022, *53* (3), 407–38.
- 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.
- Baird, Sarah, Jed Friedman, and Norbert Schady**, “Aggregate Income Shocks And Infant Mortality In The Developing World,” *The Review of Economics and Statistics*, 2011, 93 (3), 847–856.
- Belchem, John**, “The Liverpool–Irish enclave,” *Immigrants & Minorities*, 1999, 18 (2-3), 128–146.
- Bindler, Anna, Randi Hjalmarsson, Stephen Machin, and Melissa Rubio-Ramos**, “Murphy’s Law or Luck of the Irish? Disparate Treatment of the Irish in 19th Century Courts,” *CEPR Discussion Papers*, 2023.
- 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.
- 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.
- Collins, William J. and Ariell Zimran**, “The economic assimilation of Irish Famine migrants to the United States,” *Explorations in Economic History*, 2019, 74, 101302.
- Connolly, Sean**, *On Every Tide: The Making and Remaking of the Irish World*, London: Little, Brown, 2022.
- Connor, Dylan Shane**, “Class Background, Reception Context, and Intergenerational Mobility: A Record Linkage and Surname Analysis of the Children of Irish Immigrants,” *International Migration Review*, 2020, 54 (1), 4–34.
- , “In the Name of the Father? Fertility, Religion, and Child Naming in the Demographic Transition,” *Demography*, 10 2021, 58 (5), 1793–1815.
- 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.

- de Bromhead, Alan, Alan Fernihough, and Cormac Ó Gráda**, “175 years of mismeasuring Ireland?: some traps set by historical data,” *Journal of the Statistical and Social Inquiry Society of Ireland*, 2023, 51, 128–155. Publisher Copyright: © 2022, Statistical and Social Inquiry Society of Ireland. All rights reserved.
- 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.
- Glynn, Sean**, “Irish Immigration to Britain, 1911–1951: Patterns and Policy,” *Irish Economic and Social History*, 1981, 8 (1), 50–69.
- 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>.
- Huck, Paul**, “Infant Mortality and Living Standards of English Workers During the Industrial Revolution,” *The Journal of Economic History*, 1995, 55 (3), 528–550.
- 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.
- MacRaild, Donald**, *The Irish Diaspora in Britain, 1750-1939. 2nd edition*, London: Macmillan, 2010.
- Malcolm, Elizabeth.**, *Elderly Return Migration from Britain to Ireland: A Preliminary Study*, Dublin: Council for the Elderly, 2006.

- Marx, Karl and Friedrich Engels**, *Ireland and the Irish question*, Moscow: Progress Publishers, 1971.
- 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.
- Mitch, David**, “The Spread of Literacy in Nineteenth-Century England,” *The Journal of Economic History*, 1983, 43 (1), 287–288.
- 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.
- , “Ethnic group differences in health, employment, education and housing shown in England and Wales’ Census 2021,” 2023. www.ons.gov.uk: link.
- 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. MacRaild**, “Paddy and Bidy no more: An evolutionary analysis of the decline in Irish Catholic forenames among descendants of 19th century Irish migrants to Britain,” *Annals of Human Biology*, 2009, 36 (5), 595–608. PMID: 19688622.
- Turner, John D.**, “Wealth concentration in the European periphery: Ireland, 1858-2001,” *Oxford Economic Papers*, 2010, 62 (4), 625–646.
- 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 University Press, 1981.

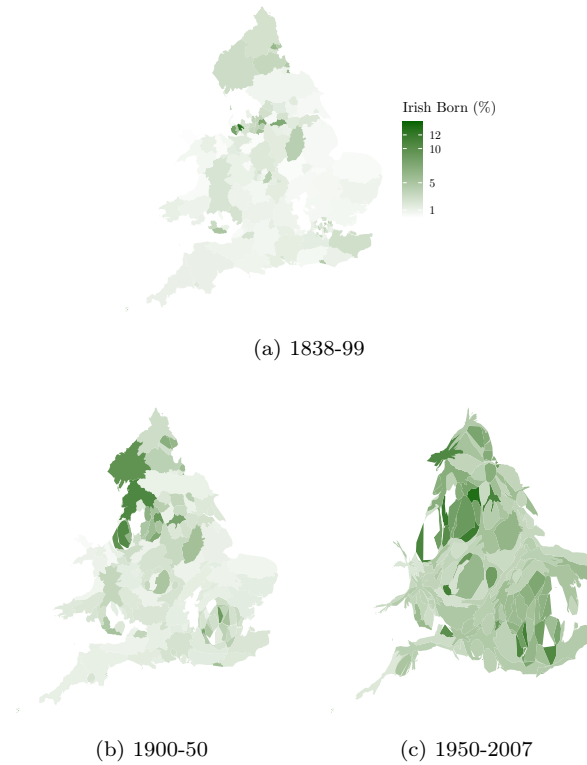


Figure 11: The Spatial Distribution of the Irish 1838-2007

Source: 100% sample of Birth Registers.

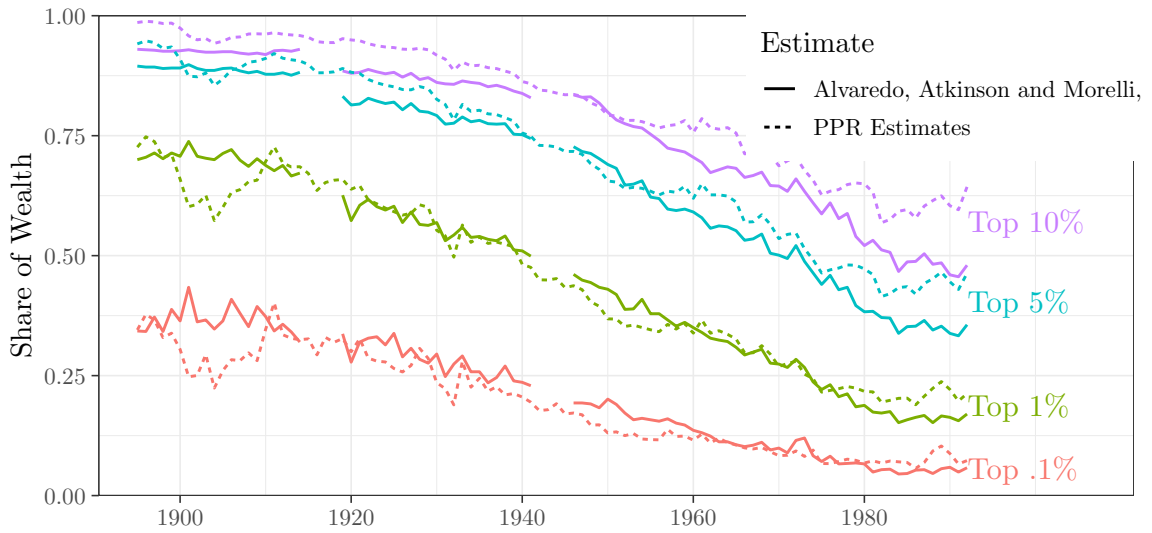
Where were the Irish in England?

Where were there Irish in England? Figure 11 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. The map is drawn as a Cartogram where the spatial units are distorted so they are proportional to total population.

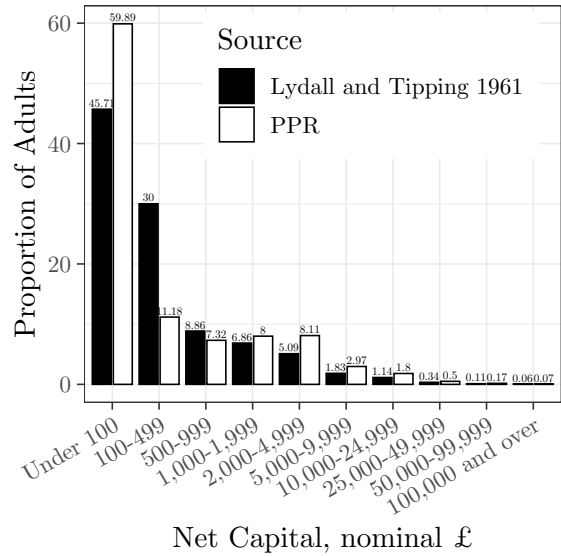
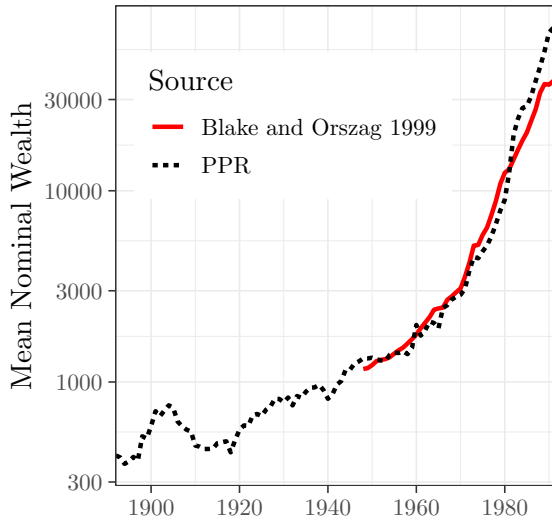
ONLINE APPENDIX to “The Irish in England”

Neil J Cummins and Cormac Ó Gráda

Extra Detail on Data



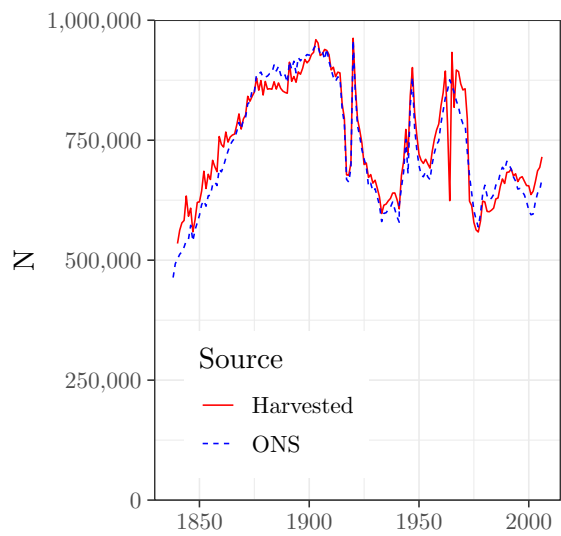
(a) Comparing Different Estimates of Top Wealth Shares, England 1892-1992



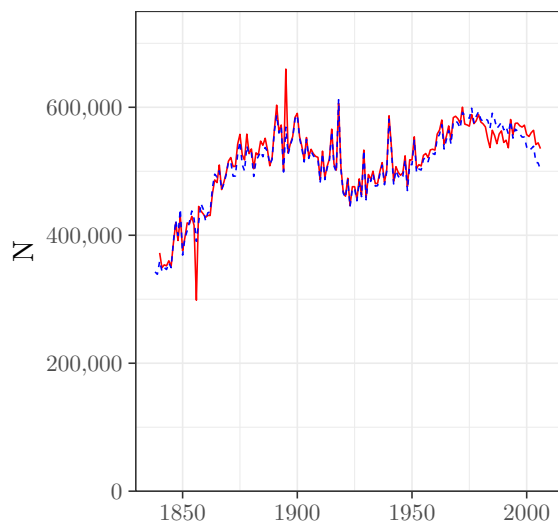
(b) Comparing Average Wealth with Blake and Orszag (1999) (c) Comparison of Net Capital with Lydall and Tipping (1961), by Wealth Band, 1950s

Figure 12: The PPR Calendar Wealth Data, Compared with Existing Estimates

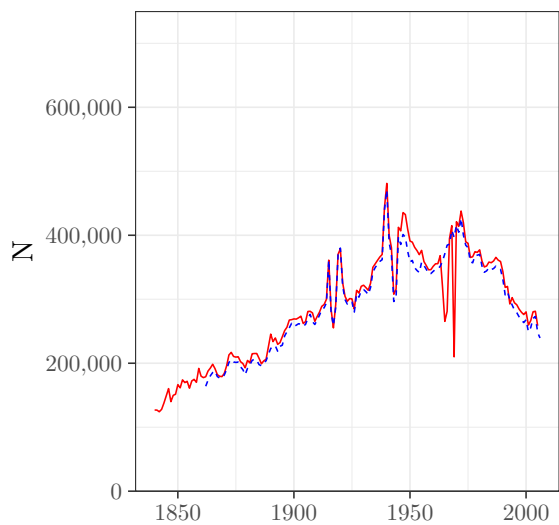
Notes: See Cummins (2021) for a detailed account of the source, construction and validation of the PPR data. Sources: PPR wealth data, Alvaredo et al. (2017) table D1, Blake and Orszag (1999, Table 12) (sum of columns 'net financial wealth', 'housing wealth' and 'consumer durable assets'). These aggregate 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 13: Data Collection Verification, 'Harvested' versus Official Count Comparison
Notes: The source for the Official Counts is Office for National Statistics (2021b).

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, we 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 7 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 £50,000.²⁸

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.²⁹

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-

²⁸The 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.

²⁹Atkinson 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 7: 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 8 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 8: 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

Other Extra Detail

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 inspecting all 5,805 Irish names for anomalies. 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 (thereby 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
MCCAVANAGH	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
MCCAHILL	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
NULTY	442	MCNULTY	5461

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

Surname	N	Updated Surname	N
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
MCMAKEN	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
O'RIELLY	49	RIELLY	350

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

Surname	N	Updated Surname	N
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	MCALHONE	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	MICALISKEY	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
ANOY	0	MCANOY	8

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

Surname	N	Updated Surname	N
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	MCALARNEY	0
ALERNEY	0	MICALERNEY	0
ALORAN	0	MICALORAN	0
ANARNEY	0	MCANARNEY	0
ANESPY	0	MCANESPY	0
GUICKIN	0	MCGUICKIN	0
KIVERIGAN	0	MCKIVERIGAN	0

Table 9: Adjusted Irish Stem Names (Mc and O')

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 10. Table 11 controls for district of death.

Table 10: 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	-6.12*** (.03)	-6.20*** (.03)	-5.42*** (.03)	-5.64*** (.03)	-2.11*** (.02)	-.88*** (.02)
Welsh	.54*** (.11)	.62*** (.11)	2.80*** (.10)	3.44*** (.10)	-.64*** (.05)	-.85*** (.05)
Scottish	1.25*** (.13)	1.61*** (.13)	-.60*** (.11)	-.39*** (.11)	-2.89*** (.06)	-3.01*** (.06)
Irish	-6.99*** (.10)	-6.28*** (.10)	-11.14*** (.10)	-10.23*** (.09)	-5.53*** (.05)	-6.09*** (.05)
Other	-1.46*** (.18)	-.88*** (.18)	-5.29*** (.12)	-4.82*** (.12)	-2.45*** (.06)	-2.45*** (.06)
Age at Death Quadratic?		✓		✓		✓
Observations	3,155,398	3,155,398	7,716,988	7,716,988	14,274,538	14,274,538
R ²	.01	.02	.01	.02	.002	.01

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

Table 12 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 13 controls for county of death.

Table 11: 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	-6.20*** (.03)	-6.33*** (.03)	-5.64*** (.03)	-5.85*** (.03)	-.88*** (.02)	-1.23*** (.02)
Welsh	.62*** (.11)	.40** (.13)	3.44*** (.10)	3.09*** (.11)	-.85*** (.05)	-.19*** (.05)
Scottish	1.61*** (.13)	1.53*** (.13)	-.39*** (.11)	-.40*** (.11)	-3.01*** (.06)	-2.75*** (.05)
Irish	-6.28*** (.10)	-5.59*** (.11)	-10.23*** (.09)	-9.16*** (.09)	-6.09*** (.05)	-5.09*** (.04)
Other	-.88*** (.18)	-.97*** (.18)	-4.82*** (.12)	-3.78*** (.12)	-2.45*** (.06)	-2.99*** (.06)
Age at Death Quadratic?	✓	✓	✓	✓	✓	✓
District Fixed Effects?		✓		✓		✓
Observations	3,155,398	3,155,398	7,716,988	7,716,988	14,274,538	14,274,538
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 12: 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 13: Probated Wealth and Ethncity, 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.

Results using Different Thresholds for Assigning Ethnicity from the 1911 Census

Our favored measure of ethnicity is assigned for an individual surname based on the distribution of place of birth of the holders that surname in the 1911 census of England and Wales. Using a two step rule, 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 20% 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 2).

Here we compare our average wealth results based upon the 20% threshold used in the paper with assignments based on 15% and 25% thresholds. Figure 14 reports the average wealth for the Irish and Scottish ethnic groups for these varying thresholds. We exclusively focus on these two groups as they are the major 19th century migrant groups to England and Wales, and thus the threshold dynamics are expected to be similar.

For the Irish the 15% and 20% cutoffs produce almost identical results with some modest divergence in the 1860s, and early 20th century. However both the Irish and the Scottish, a 25% threshold produces higher estimates of wealth for most of the sample period. A 15% rule also produces slightly lower estimates of wealth for the Scottish.

There are 506,441 “Irish” with 19,868 surnames assigned using a 15% cut-off, 360,912 with 18,316 surnames using a 20% cut-off, and 206,091 with 17,201 surnames using a 25% cutoff. Thus the counts per surname decline as the assignment threshold rises, from 25.49 per surname using a 15% rule, to 19.7 for a 20% rule, and 11.98 for a 25% rule. Therefore employing the higher cut-off has the undesired effect of selecting for more unusual, smaller names. This is potentially problematic if rarer names themselves are correlated with wealth.

Figure 15 demonstrates that for the English, rarer names are indeed correlated with higher wealth. This figure reports wealth 1858-1992, stratified for English surnames by frequency range in the universe of deaths, 1838-1900. Thus utilizing a higher selection cutoff in assigning ethnicity may result in an unrepresentative sample as it selects for rarer names, which are powerfully correlated with wealth.

Thus, as a sanity check we calculated average wealth by year for the top 15 Irish surnames reported previously in table 3. We do this as these top 15 Irish surnames are commonly accepted as being “Irish”. Figure 16 reports the average wealth for this group of commonly accepted Irish aligns almost perfectly with that generated from our 15% and 20% threshold cutoffs. In the next section we compare our estimates with those using an entirely different method of attributing surnames to ethnicity, and find our cutoff to be broadly in line with that too. Thus we interpret our 15% and 20% cutoffs for defining ethnicity as representing the average experience of the Irish in England, with the 25% representing a positively selected right tail of the Irish in England status distribution.

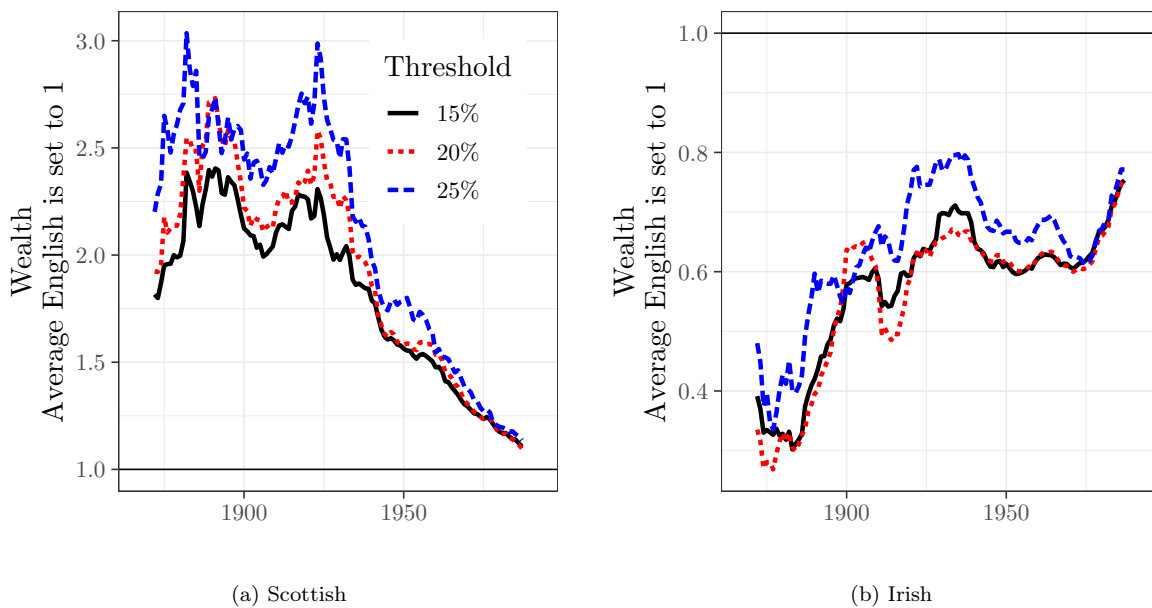


Figure 14: Comparing Average Wealth, 1858-1992, using Varying Thresholds for a Surname's Ethnicity

Note: We assign ethnicity using the individual records of the 1911 census of England and Wales. For each surname we inspect the distribution of birth of the holders of that surname. Using a two step rule, 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 20% or over of the births of that surname, we update the ethnic origin to that country. For robustness we here compare that 20% rule with a 15% rule, and a 25% rule.

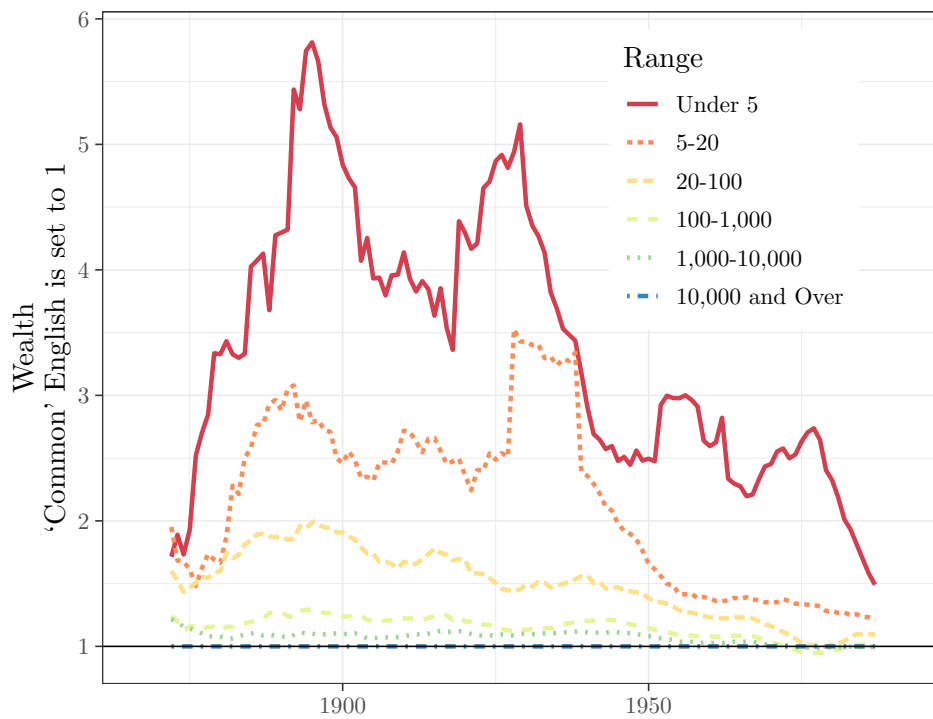


Figure 15: The Variation in Average Wealth by the 1858-1900 Frequency of a Surname, 1858-1992

Notes: Range refers to Frequency range of a surname in the universe of death records, 1838-1900. The figure illustrates that rarer surnames are richer than common surnames. 'Common' English are surnames in 10,000 and over range.

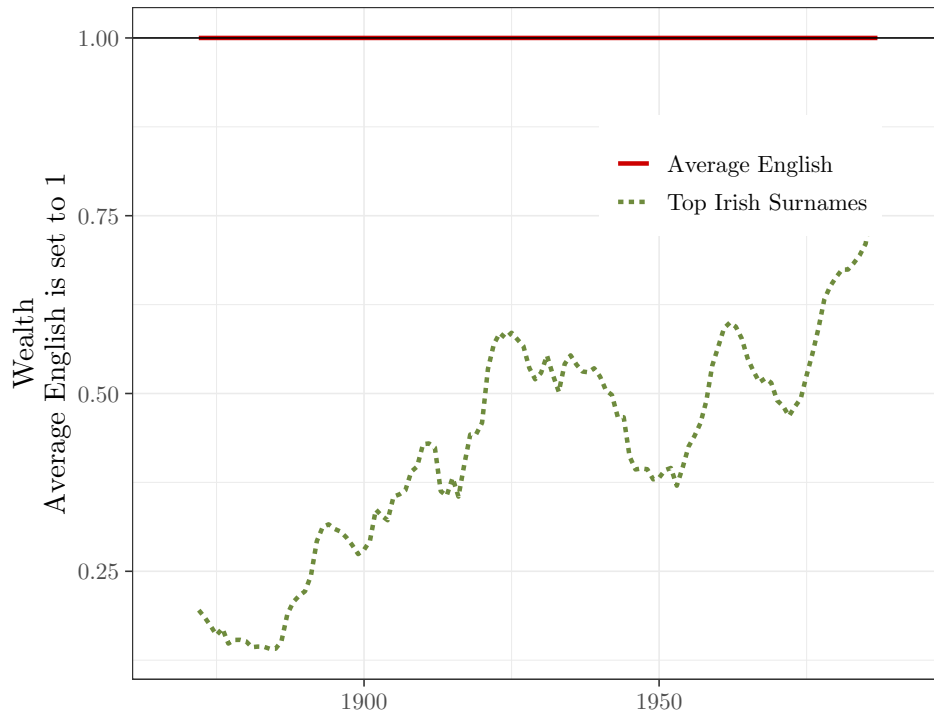


Figure 16: Wealth for the Average English compared with those holding one of the Top 15 Irish Surnames

Notes: We took the top 15 Irish surnames as listed at <https://forebears.io/ireland/surnames>, a website that has aggregated a considerable volume of data on contemporary global surname distributions.

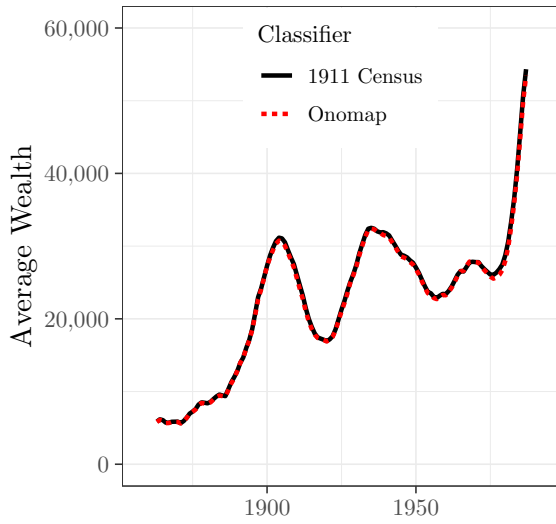
Results using an Alternative Ethnic Classification

How robust are these patterns to a entirely 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*.

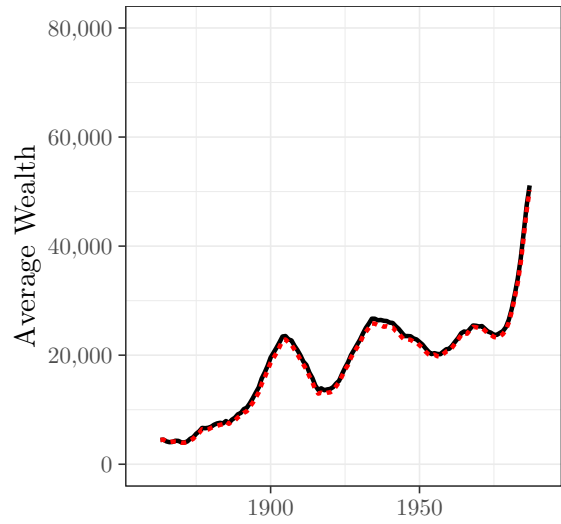
Figure 17 reports the average wealth for the British and Irish ethnic groups for both classifications. They are identical for the English, the Scottish, and the Welsh. However, for the Irish the trends are different. The Onomap classifier results in wealth estimates substantially lower than that of the 1911 census assignment used here.

Figure 18 reports the infant mortality rate for the British and Irish ethnic groups for both classifications. The different classifications produce identical results.

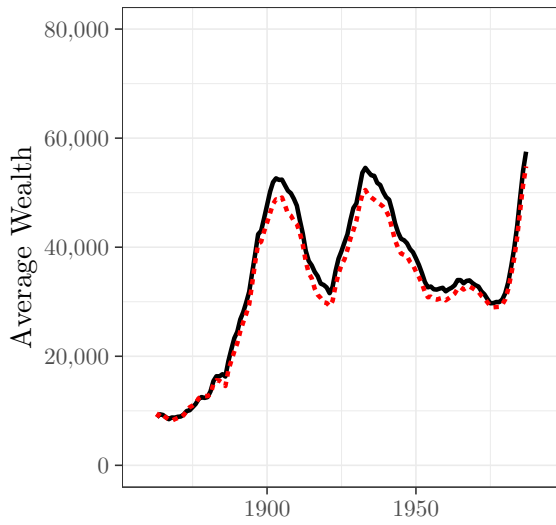
One possible explanation for the divergence between the methods for Irish Wealth is the *contemporary* nature of the Onomap classifier. Over time, successful Irish could integrate into the English and adopt English forenames for their children. These Irish would then be classified as ‘English’ by Onomap. 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.



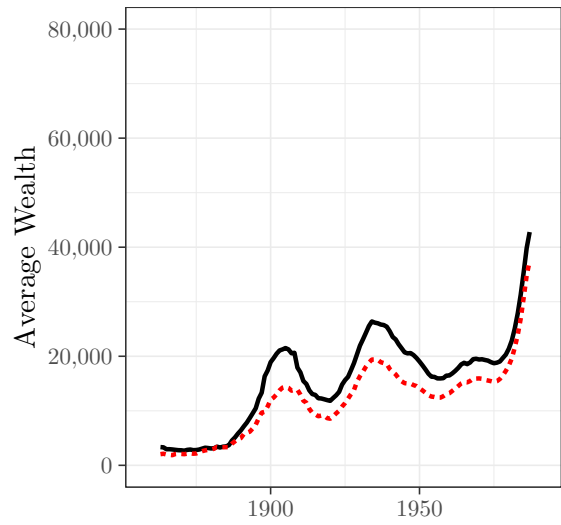
(a) English



(b) Welsh

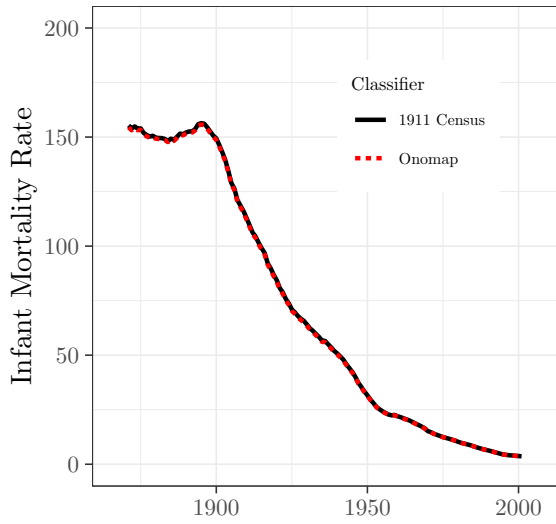


(c) Scottish

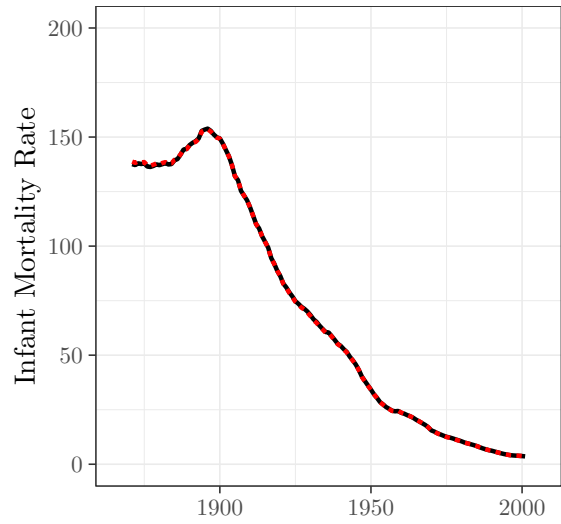


(d) Irish

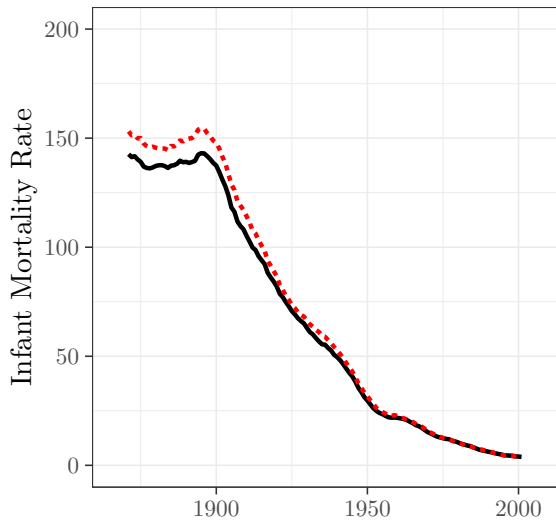
Figure 17: Comparing Average Wealth, 1858-1992, using Alternative Ethnic Classifier



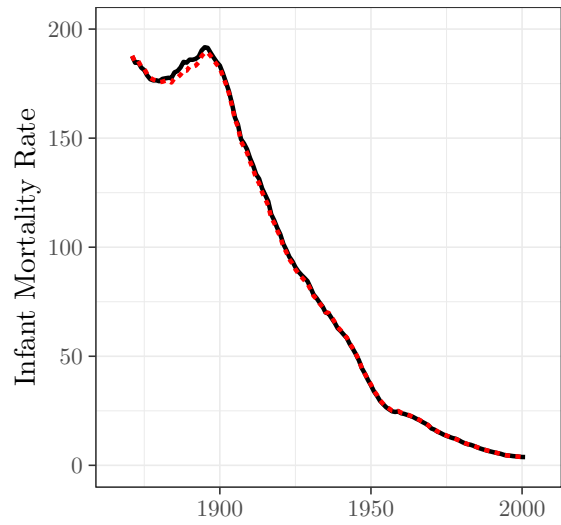
(a) English



(b) Welsh



(c) Scottish



(d) Irish

Figure 18: Comparing Infant Mortality Rates, 1866-2007, using Alternative Ethnic Classifier

Extra Tables and Figures

Table 14: Summary Statistics, Wealth

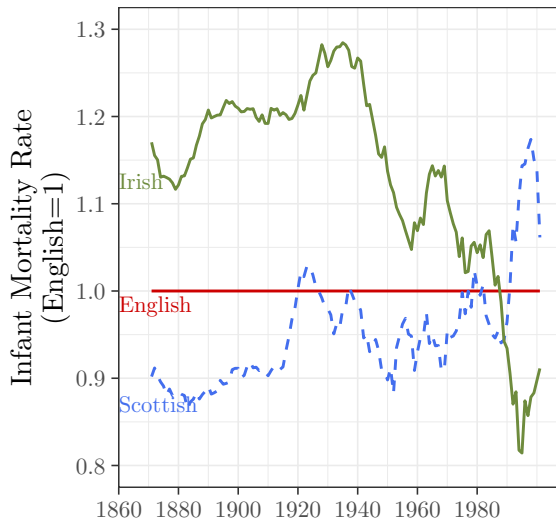
Ethnicity	N			Real Wealth				
	Deaths	Probates	Probate Rate	Min.	Max.	Mean	Median	Top 1% Prop.
All	70,045,750	15,246,247	0.22	0	1,257,371,575	23,856	895	0.010
English	61,354,562	13,379,209	0.22	0	1,257,371,575	23,967	895	0.010
Welsh	5,700,027	1,225,467	0.21	0	165,769,426	19,543	895	0.008
Scottish	813,189	219,782	0.27	0	113,345,260	43,683	1,204	0.020
Irish	1,648,773	258,650	0.16	0	105,187,746	14,932	849	0.005

Source: Universe of Deaths 1837-2007 and Probates 1858-1992.

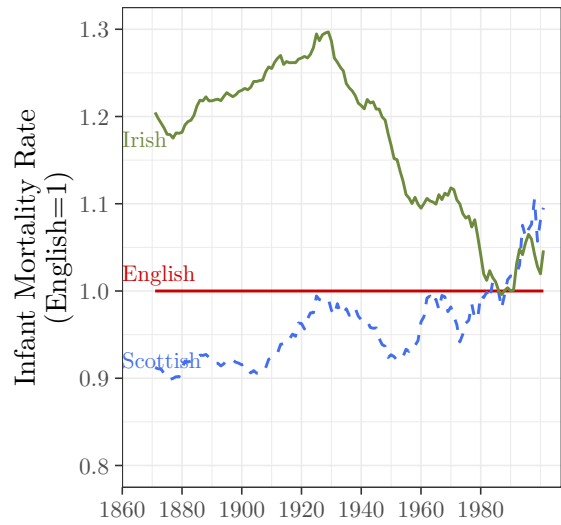
Table 15: Summary Statistics, Probate Rate, 1996-2018

Ethnicity	N		
	Deaths	Probates	Probate Rate
All	11,739,085	5,601,612	0.48
English	10,165,342	4,874,073	0.48
Welsh	942,959	454,052	0.48
Scottish	216,224	104,944	0.49
Irish	414,560	168,543	0.41

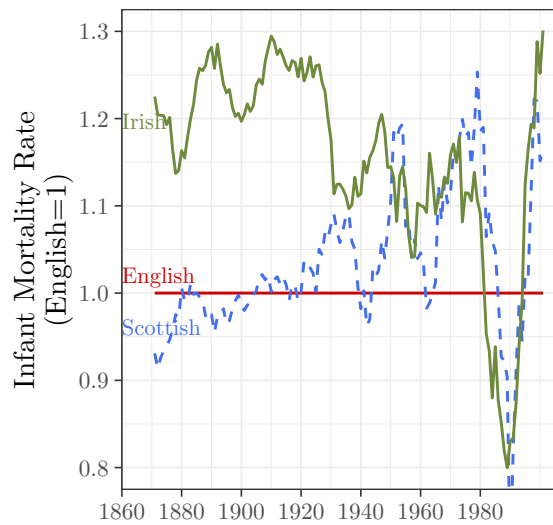
Source: Database of the number of deaths and the number of probates, by surname, 1996-2018.



(a) The North



(b) The South, ex. London



(c) London

Figure 19: Regional Differences in the Relative Infant Mortality of the Irish
Note: The English baseline is established by region.

Table 16: Summary Statistics, Infant Mortality, 1866-2006

Ethnicity	1866-1929			1930-2006		
	Births	Infant Deaths	IMR	Births	Infant Deaths	IMR
All	54,062,848	7,035,768	130.1	56,449,908	1,230,969	21.8
English	47,024,384	6,125,985	130.3	44,359,607	1,002,965	22.6
Welsh	4,447,423	557,552	125.4	4,159,910	99,061	23.8
Scottish	515,531	60,800	117.9	970,571	18,061	18.6
Irish	1,052,308	163,358	155.2	2,214,129	47,564	21.5

Source: Universe of Births and Deaths, 1866-2006. IMR is the infant mortality rate, per 1,000 births.

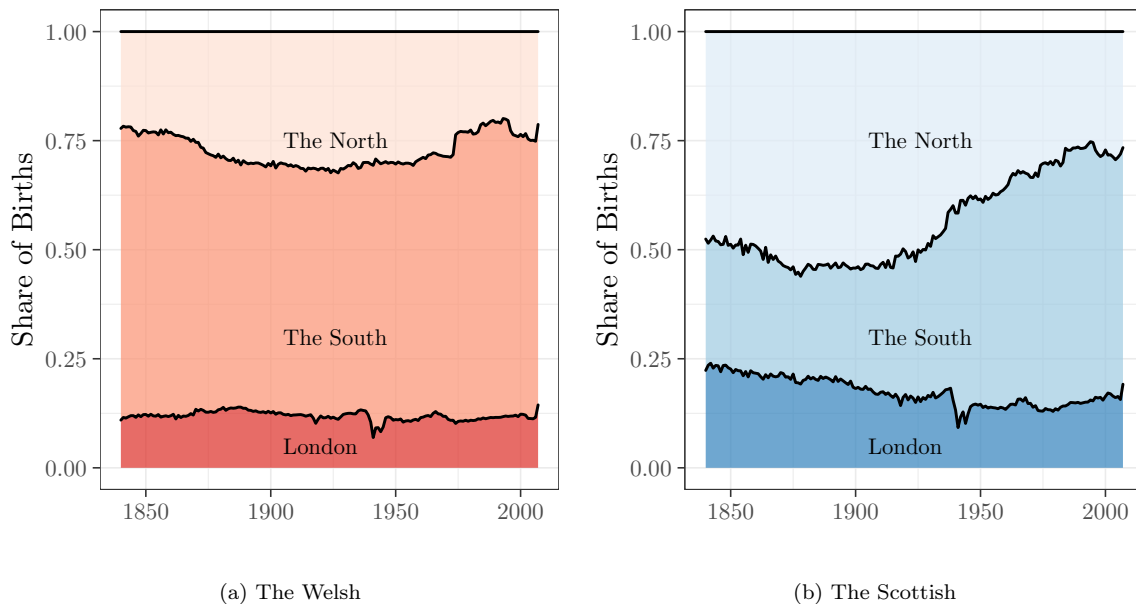


Figure 20: The Regional Distribution of Births, the Welsh and the Scottish