MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Konrad-Zuse-Strasse 1 · D-18057 Rostock · Germany · Tel +49 (0) 3 81 20 81 - 0 · Fax +49 (0) 3 81 20 81 - 202 · www.demogr.mpg.de

MPIDR Working Paper WP 2023-050 I December 2023
Revised March 2025
https://doi.org/10.4054/MPIDR-WP-2023-050

Leaving for Life: Using Online Crowd-Sourced Genealogies to Estimate the Migrant Mortality Advantage for the United Kingdom and Ireland During the 18th and 19th Centuries

Elena Maria Pojman | empojman@psu.edu
Duke Mwedzi
Orlando Olaya Bucaro
Stephanie Zhang
Michael Y. C. Chong
Monica Alexander
Diego Alburez-Gutierrez

Leaving for Life: Using Online Crowd-Sourced Genealogies to Estimate the Migrant Mortality Advantage for the United Kingdom and Ireland During the 18th and 19th Centuries

Elena Maria Pojman,[1] Duke Mwedzi,[2] Orlando Olaya Bucaro,[3] Stephanie Zhang,[4] Michael Y. C Chong,[5] Monica Alexander,[6] Diego Alburez-Gutierrez[7]

[1] Department of Sociology and Criminology, The Pennsylvania State University, University Park, PA, USA, empojman@psu.edu, https://orcid.org/0000-0001-8982-6324

[2] Department of Global Development, Cornell University, Ithaca, NY, USA, dem327@cornell.edu, https://orcid.org/0000-0001-7394-2427

[3] International Institute for Applied Systems Analysis, Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/OEAW, University of Vienna), Vienna, AT, olayabucaro@iiasa.ac.at, https://orcid.org/0000-0002-9847-1374

[4] Department of Sociology, University of California, Los Angeles, CA, USA szhang19@g.ucla.edu, https://orcid.org/0000-0001-9804-6522

[5] Department of Statistical Sciences, University of Toronto, Toronto, Canada, myc.chong@mail.utoronto.ca, https://orcid.org/0000-0003-1158-3822

[6] Departments of Statistical Sciences and Sociology, University of Toronto, Toronto, Canada, monica.alexander@utoronto.ca, https://orcid.org/0000-0002-8135-3435

[7] Kinship Inequalities Research Group, Max Planck Institute for Demographic Research, Rostock, DE, alburezgutierrez@demogr.mpg.de, https://orcid.org/0000-0002-9823-5179

E.M.P., D.M., O.O.B., and S.Z. contributed equally to this work and should all be considered first authors.

Corresponding author: empojman@psu.edu

19 March 2025

Acknowledgement: We are grateful for helpful feedback from attendees at the 2023 Southern Demographic Association and 2024 Population Association of America Annual Meetings. This project was funded by the 2023 Population and Social Data Science Summer Incubator Program of the Max Planck Institute for Demographic Research. We thank all participants and MPIDR researchers for their suggestions.

E.M.P – conceptualization, formal analysis, investigation, formal analysis, methodology, project administration, validation, visualization, writing – original draft; D.M. – conceptualization, formal analysis, investigation, methodology, visualization, project administration, writing – original draft; O.O.B. – conceptualization, data curation, formal analysis, investigation, methodology, project

administration, validation, visualization; S.Z. – conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing – original draft; M.Y.C.C. – methodology, writing – review & editing; M.A. – supervision, methodology, writing – review & editing; D.A.G. – supervision; writing – review & editing.

Abstract

Demographic studies consistently find a mortality advantage among migrants, but a lack of longitudinal data tracking individuals across national borders has limited the study of historical international migration. To address this gap, we use the crowd-sourced online genealogical database Familinx to estimate the migrant mortality advantage for migrants from the United Kingdom and Ireland between 1750 and 1910. We compare age at death for non-migrants and migrants to Canada, the United States, South Africa, New Zealand, and Australia using mixed-effects regression models that account for unobserved factors shared between siblings. Results suggest an overall expected migrant advantage of 5.9 years, 95% CI [5.7, 6.2] even after accounting for between-family variation, with migrants estimated to live an additional 2.6 [1.1, 4.0] to 8.7 [6.3, 11.2] years depending on the country of destination. This study contributes to the understanding of the migrant mortality advantage in a historical context and shows the potential for online genealogies to contribute to demographic research.

Keywords: *crowd-sourced genealogies, migrant mortality advantage, United Kingdom, Ireland, sibling effects*

Introduction

Across different time periods and contexts, migrants are found to live longer than their counterparts in both the origin and destination countries. This effect has been found to be robust across a range of contexts (Abraído-Lanza et al., 1999; Aldridge et al., 2018; Swerdlow, 1991) despite possible statistical censoring issues such as the undercounting of return migrants (Pablos-Méndez, 1994). One hypothesized reason for the migrant mortality advantage is the healthy migrant effect, which argues that healthy individuals positively select into migration (Razum, 2008). If this is true, it is unclear the *extent* of the migrants' mortality advantage over non-migrants, especially as migrant destinations may vary greatly from one another and from the origin country. Additionally, the mortality outcomes of migrants may not be independent of their family members' outcomes, even among family members who did not migrate. Siblings are especially similar because of shared genetic and environmental factors. However, data concerns have typically limited researchers' ability to study how between-family effects may contribute to the migrant mortality advantage.

A common strategy in many migrant mortality studies is to compare migrants with non-migrants in the destination country. However, this strategy risks obscuring the relationship between migration and mortality, as it rests on the assumption that non-migrants in the destination serve as an appropriate comparison group. Instead, we use the online crowd-sourced genealogical dataset Familinx (Kaplanis et al., 2018) to study the migrant mortality advantage during a period of large emigration from the United Kingdom and Ireland from the mid-18th century until the early 21st century. In addition to being able to compare migrants and non-migrants who are born in the same country, the genealogical structure of the data allows comparison between siblings. We test whether the migrant mortality advantage holds for this migration flow using mixed effects models which allow us to distinguish between the effects on mortality that stem from shared health

advantages within families versus the direct effect of migration on mortality.

The current study is guided by the following questions: (1) What is the size of the migrant mortality advantage for migrants from the United Kingdom and Ireland between 1750-1910? (2) How does this mortality advantage vary across the destination of migration? We focus on Canada, the United States, South Africa, New Zealand, and Australia as destination countries because of their geographical variation. As British colonies, these countries also facilitated virtually free migration from the UK and Ireland, mitigating the effects of selection bias in more recent migration flows due to legal restrictions on migration. Our results suggest an overall expected migrant advantage of 5.9 years, 95% CI [5.7, 6.2] even after accounting for between-family variation. The results also suggest that the migrant mortality advantage is heterogeneous by destination, ranging from 2.6 [1.1, 4.0] years in South Africa to 8.7 [6.3, 11.2] years in New Zealand. Though historical in coverage, these findings offer a novel contribution to the literature on the migrant mortality advantage by offering a new context in which migrants out-lived their non-migrant peers. Additionally, they show the potential utility of crowd-sourced genealogical data in demographic research.

Background

Mechanisms Driving the Healthy Migrant Hypothesis

The size of the migrant mortality advantage is partially dependent on the group against whom migrants are compared. Table 1 shows two common approaches of previous studies, including data requirements and interpretations. The first, more common approach is to compare migrants to non-migrants in the destination country, and the second approach is to compare migrants to non-migrants from the same origin country.

[INSERT TABLE 1 ABOUT HERE]

Studies using the first approach consistently find a mortality advantage for immigrants relative to the native-born population (Guillot et al., 2023; Mehta et al., 2016; Razum, 2008). Research on more recent migration flows have argued that this advantage is paradoxical, as migrants tend to be of lower socioeconomic status than individuals in the destination country (Abraído-Lanza et al., 1999; Bakhtiari, 2022). Several explanations have been offered for this seemingly paradoxical finding. First, the *healthy migrant hypothesis* suggests that relatively healthy individuals are more likely to migrate (Abraído-Lanza et al., 1999; Feliciano, 2020). However, measurement issues and other data limitations complicate our understanding of this relationship. Return migrants may be negatively selected or undercounted in population statistics, leading to numerator-denominator bias and rendering them "statistically immortal" as their deaths are not recorded by the destination country (Pablos-Méndez, 1994; Puschmann et al., 2017). Additionally, observed demographic rates in either the origin or destination country may suffer from migration censorship, that is, an overrepresentation of individuals who die before they can migrate, artificially decreasing the average age of death for non-migrants relative to migrants (Kasakoff & Adams, 1995; Ruggles, 1992). Despite these challenges, studies still find a mortality advantage even after accounting for many of these potential biases (Abraído-Lanza et al., 1999).

In the second approach, migrants are compared to members of the origin country who did not migrate. Non-migrants in the origin country are likely to share early-life conditions and other characteristics that may be associated with their mortality outcomes. Comparing migrants to their non-migrant peers in the origin country also makes it possible to account for alternative mechanisms that may shape the migrant mortality advantage, such as the role of unobserved factors shared between siblings (e.g., early-life environment or parental resources) which may contribute

to similarities in mortality between migrants and their non-migrant siblings. This approach is more difficult to apply due to its data requirements but offers a better comparison between migrants and non-migrants in the origin country due to shared environments and life conditions prior to migration.

Though migrants may be healthier than their counterparts in the destination country, it is less clear how much healthier they are than those who do not migrate, that is, the family and neighbors they leave behind. Migrants are not drawn from the population at random and may have lived longer regardless of whether they migrated. On one hand, emigrants may be positively selected on a host of factors that are typically associated with a longer life: for example, socioeconomic status, health status, or survival to migration age. On the other hand, emigrants may have moved due to a lack of economic and social opportunities in their home country and thus benefitted from better conditions in their destination which increased their lifespan.

Scholars have argued that mortality differences that emerge due to migration are driven by conditions in the destination rather than the origin country (Hatton, 2021; 2004). Recent research confirms this argument; immigrants to the United States in the early 20th century saw a mortality *dis*advantage compared to the native-born population in the destination country because of higher infectious disease exposure in U.S. cities (Bakhtiari, 2022). As such, the mechanisms driving the migrant mortality advantage are unclear. It may be that migrants are healthier and otherwise more positively selected compared to their peers, or, alternatively, migrating allows individuals to avoid poor conditions in their home countries that would have negative effects on their mortality.

Migrant Mortality Advantages and Family Effects

Individuals are not randomly selected into migration. Rather, selection into migration occurs at both the population level, where some individuals are more or less likely to migrate based on their socioeconomic circumstances, and within families, as siblings may have different abilities, economic opportunities, or familial responsibilities (Abramitzky et al., 2012; Mourits & Puschmann, 2023). For example, individuals who do not expect to inherit their family's land, by virtue of gender or birth order, may have an increased propensity to migrate. Families may also select the healthiest member to migrate as a risk diversification strategy (Stark & Bloom, 1985).

If migrant siblings possess mortality advantages because they are positively selected from the population, their non-migrant siblings may also be similarly advantaged, as health-protective factors are correlated within families (Mourits & Puschmann, 2023). Siblings are more similar to each other than they are to random members of the population due to shared genetic and environmental factors (Piraino et al., 2014), though evidence regarding whether the socio-environmental or genetic component is more influential for mortality is mixed (Cournil & Kirkwood, 2001; Gudmundsson et al., 2000; Piraino et al., 2014).

Siblings also compete for scarce resources amongst themselves (Donrovich et al., 2014; Lam & Marteleto, 2008), which could diminish the magnitude of shared mortality advantages between siblings. Evidence for the effect of sibling size on mortality is mixed (Baranowska-Rataj et al., 2017; Sonneveldt et al., 2013), though overcrowding due to having many siblings may be a key mechanism driving poor outcomes, especially for larger families (Hatton & Martin, 2008). In light of these findings, it is important to consider both the role of unobserved factors shared between siblings and individual characteristics when considering the relationship between migration and mortality.

Historic Emigration from the United Kingdom and Ireland

The Demographic Transition together with the Industrial Revolution set the stage for mass emigration from Europe starting in the 18th century (Hatton & Williamson, 1994; Richards, 2018). Emigration served as a "safety valve" in the face of overcrowding and few economic opportunities, relieving the pressure caused by a higher population due to the Demographic Transition (Blanc & Wacziarg, 2025; Hatton & Williamson, 1994). The emigrants themselves were a group most poised to benefit from migration: typically young, single, unskilled men or young couples with small children (Hatton & Williamson, 1994; Horn, 1998; Thompson, 2009; Tomlins, 2001). However, it is also important to note that many migrants were coerced to move, including convicts and indentured servants (Richards, 2018; Tomlins, 2001).

In the United Kingdom, mortality decreased sharply while fertility remained stable until the mid-19th century before declining thereafter, resulting in a large population increase as the country moved through the second and third stages of the Demographic Transition (Friedlander & Okun, 2022). This population growth exceeded agricultural labor demands, despite increasing agricultural productivity (Richards, 2018). Economic circumstances worsened for many, with higher rent costs and declining access to land (Hatton & Williamson, 1994; Horn, 1998; Richards, 2018). Coupled with rapid urbanization during the Industrial Revolution from 1750–1850, urban areas were subject to overcrowding, poor sanitation, and poor health conditions, resulting in decreased lifespan, particularly for low socioeconomic status individuals (Taylor, 1988). These poor conditions prompted the out-migration of an estimated 19 million people in total from the UK and Ireland between the 18th and 19th centuries (Richards, 2018).

The sheer size of this flow, combined with the lack of legal restrictions on immigration to these countries, makes it an ideal case to study differences between migrants and non-migrants (Hatton, 2021). Colonial relationships with overseas territories facilitated mass migration for

Europeans to colonies in North America, Australasia, and Africa, with the vast majority settling in present-day United States, Canada, Australia, and New Zealand (Hatton, 2021). The varied destinations of the flow also make it possible to test whether the migrant mortality advantage holds across different contexts.

Possible Drivers of Mortality Differences Among Migrants

It is well documented that individuals select into migration, but migrants also select into different destinations based on a variety of socioeconomic factors. During the 19th Century, British migrants to Australia and New Zealand tended to be more skilled on average than those moving to Canada and the United States (Hatton 2021). This may have been due to cost differences, with fare from London to New York costing a third of the fare to Australia. As such, migrants from professional and commercial backgrounds were more likely to migrate to the Pacific, while agricultural and manual workers tended to migrate to North America (Hatton 2021). There is also evidence that British migrants to South Africa tended to be from poorer backgrounds and were motivated to move by the desire for greater economic security and independence (Tosh, 2017).

Additionally, infectious disease may have also contributed to differences in mortality rates among migrants. Prior work has shown that differences in childhood experience contributes to susceptibility to disease as an adult (Alter and Oris, 2005). Combined with existing differences in selection into migration destination by class and background, susceptibility to disease may have differed by location. Although infectious diseases such as tuberculosis, measles and smallpox were often brought by European settlers, there were some infectious diseases such as Malaria that they would not have had resistance to (Haines and Shlomowitz, 1998). Migrants who moved to densely populated urban areas may also have been more vulnerable to infectious disease mortality

(Bakhtiari, 2022). In short, a migrant's background could have influenced their susceptibility to disease as well as migration destination, resulting in disparate health outcomes across destination.

Another possible driver of mortality differences among migrants is variations in exposure to violent conflict between European migrants and indigenous populations in colonies. Contexts that saw more prolonged and large-scale violent conflict would have experienced more casualties for British and Irish migrants and thus, increased mortality. South Africa in particular is notable for large scale conflicts in the 18th and 19th centuries that led to heavy casualties for British settlers (Webb, 2015; Evans et al, 2018; Beckett, 2019). The US similarly experienced multiple protracted wars between the 17th and 19th Centuries between European settlers and various indigenous groups (Nester, 2004; Draper, Kerby and Baguley, 2023).

By contrast, where indigenous resistance was less intense or less confrontational, British and Irish migrants had less exposure to violence and thus, lower risk of mortality. In the Pacific, indigenous groups tended to favor guerilla tactics and economic sabotage, rather than open conflict (Kerkhove, 2020; Duckett White, 2023). Where open conflict did take place, it was small in scale and resulted in few casualties for migrants (Carpenter, 2015). In Canada, conflict was also small in scale and localized, as European settlers made greater use of diplomacy and treaties (Dixon, 2014). As a result, British and Irish migrants to Canada and the Pacific experienced fewer casualties than other contexts, which would be reflected in better mortality outcomes.

Current Study

We focus on the flow of migrants from the United Kingdom and Ireland to the destination countries of Canada, the United States, South Africa, Australia, and New Zealand for several reasons. Migrants moving from the United Kingdom and Ireland made up a large percentage of the flows

of this time period (Hatton, 2021). This period is also unique because the United Kingdom's colonial relationship with its territories facilitated virtually free international movement for European migrants on a scale that is no longer possible due to legal and immigration restrictions. As such, this period is an ideal case study for studying mortality differences between migrants and non-migrants as there was less selection of migrants on characteristics such as socioeconomic status and education level (Hatton, 2004, 2021).

We contribute to the continued debate on the migrant mortality advantage using a genealogical dataset that can identify historic migration flows and transnational kin ties. We account for unobserved similarities between siblings that may shape the magnitude of the migrant mortality advantage. Our modelling strategy allows us to identify whether the migrant mortality advantage is heterogeneous by destination. Where other studies are constrained by a lack of longitudinal data that follow individuals across national borders, Familinx allows us to compare migrants to non-migrants in the origin rather than the destination country. Finally, by analyzing historical data, we can measure the migrant mortality advantage across time rather than cross-sectionally.

Methods

Data and Sample

Data for the current study come from Familinx, a genealogical dataset with information on the timing and location of vital events such as birth and death for over 86 million individuals (Kaplanis et al., 2018). Familinx was scraped from Geni.com, a website where amateur genealogists upload their family trees. User uploads are then matched with historical records such as burials and baptisms to create more complete genealogies that span centuries, across multiple world regions.

Genealogical data have been used to investigate a variety of demographic outcomes (e.g., Chong et al., 2022; Corti et al., 2024; Cozzani et al., 2023; Gavrilov et al., 2002; Gay et al., 2023; Minardi et al., 2024; Piraino et al., 2014), and represent an opportunity for demographers to answer long-discussed questions about the nature of inter- and intragenerational demographic processes (Alburez-Gutierrez et al., 2019). Individuals born in the United Kingdom and Ireland are well-represented in Familinx, and the data's genealogical structure is better suited for studying migration than similar family reconstructions using parish records which suffer from migration censorship (Ruggles, 1992). This genealogical structure also allows us to identify transnational ties between kin and thereby account for unobserved similarities between siblings.

These data consist of user-generated profiles scraped from the website Geni.com. As such, there were several instances of reporting errors and other issues which necessitated extensive data cleaning and treatment before analysis. For example, some profiles in the dataset have missing information for country of birth or death; some records have implausible values for year of birth and death leading to an impossible age at death; and the names of some birth and death locations varied by language and changed over time. In general, we use birth and death information from baptism and burial records when exact birth and death dates are missing, and discard profiles with incomplete information on year and location of birth and death. Despite the biases inherent in crowdsourced genealogies (Calderón Bernal et al., 2025; Colasurdo & Omenti, 2024; Stelter & Alburez-Gutierrez, 2022), and the messiness of the raw data, we take steps, such as documenting and publishing the data preparation procedure to ensure transparency and reliability of our data and results. We describe in detail our data preparation procedure, including our treatment of missingness in key variables, in Appendix A of the Online Supplement. Replication materials for

the entire data cleaning process and analysis are available at https://osf.io/b87t6/?view_only=73ca840636424c9c9f78687103c3d90f .

The initial sample consisted of over 86 million individuals, with parent-child ties for 43 million individuals. We reduced our analytic sample to 98,057 profiles to capture individuals who were born in the United Kingdom and Ireland between 1735-1895, representing less than one percent of the sample (see Figure A1 in the Online Supplement for a visualization of the complete data cutting process). We identify all individuals with complete information (year of birth and death, location of birth and death, and gender). Due to data limitations, we consider return migrants as non-migrants; however, given the period and destinations we analyze, there are few such cases. We limit the maximum age at death to 110 to avoid including profiles with implausibly long lifespans whose information was likely entered erroneously and set the minimum age at death to 15. The sample consists of individuals who died either within the United Kingdom and Ireland (N = 62,076) or in Canada (N = 5,068), the United States (N = 18,381), South Africa (N = 1,185), Australia (N = 8,600), and New Zealand (N = 2,747). Our final sample consists of N = 98,057 individuals (N = 35,981 migrants).

Measures

We measure the outcome of age at death by subtracting one's birth year from their death year. While the data does not allow us to measure migration directly, individuals whose birth is recorded in the United Kingdom or Ireland and death is recorded in Australia, New Zealand, South Africa, the United States, or Canada are considered migrants for the purposes of this study, as we infer that they would have migrated to their eventual death location. Individuals whose birth and death are recorded in the United Kingdom or Ireland are considered non-migrants.

To account for potential confounding between age at death and migration status, we control for gender and birth cohort, the latter measured as a categorical variable in 10-year intervals. We also include controls for the number of siblings, coded as a categorical variable (0, 1, 2, 3-5, 6+). To account for sibling effects, we construct a family ID for each profile using the unique ID of one's mother, or father's ID if the mother's is missing. Siblings are defined as individuals who share one or more parent ties, and we link individuals with their siblings. Individuals without recorded siblings are given an individual ID. Table 2 shows descriptive statistics of selected variables.

[INSERT TABLE 2 ABOUT HERE]

The migration and mortality literatures recognize the influence of socioeconomic status on one's propensity to migrate and age at death, respectively (Clouston & Link, 2021; Lindstrom & Lauster, 2001; Link & Phelan, 1995). However, as our data consists of only demographic variables such as date and location of birth, date and location of death we cannot explicitly control for socioeconomic factors. This is a limitation of our study that we discuss in further detail below.

Analytic Strategy

To examine the migrant mortality advantage for those born in the United Kingdom and Ireland between 1735-1895, we estimate age at death using mixed effect models that incorporate random effects to capture unobserved similarities between siblings. We first examine mortality differences between migrants and non-migrants (Model A; migrant effects), and then estimate differences by one's country of death (Model B; destination country effects), as we infer this to be their migration destination. We interact these variables (migrant and country of death, separately) with birth cohort to account for changing mortality conditions across time. The family random effects account for

unobserved factors within families that are shared by siblings (Abramitzky et al., 2012). Family random effects have been used similarly in other studies of the migrant mortality advantage (Abramitzky et al., 2012; Cozzani et al., 2023; Mourits & Puschmann, 2023). While survival models are common in mortality research, we did not present them in our main analysis due to concerns about how well our data fit the assumptions of hazard models. In addition, since hazard models are designed for incomplete data with right censoring issues, they were not necessary for our main analyses as all observations within our dataset are deceased individuals who have completed the life course. Instead, we include results from hazard models in the appendix (see Table C1). Findings from these models indicate that the survival time is longer for migrants compared to non-migrants and that this effect is statistically significant for all destinations except South Africa.

Results

Migration flows in the 18th and 19th centuries

Figure 1 shows the migration flows identified in the sample. The US is the largest destination country for migrants across all cohorts followed by Australia. South Africa has the fewest migrants, with migration beginning in earnest for the cohort born between 1765-1775. Across all destinations, flows are relatively low for earlier cohorts then begins to surge, particularly in cohorts born between the 1770s and 1850s, before declining in later birth cohorts. A large majority of migrants across destinations are male, though the gender composition shifts between cohorts. Female migrants make up a larger share of the flow between the 1770s and 1850s, but when the overall migration flows are smaller, the share of women decreases.

[INSERT FIGURE 1 ABOUT HERE]

A Migrant Effect?

We estimate the relationship between migration and age at death using mixed effect regression models. Results are presented as average marginal effects (AMEs), allowing us to interpret our findings as the effect of being a migrant (versus a non-migrant) on average in our sample, while holding control variables at their observed values (Mize, 2019). The AME can be interpreted as the difference between the expected age at death of migrants and non-migrants. Positive values indicate that migrants have a mortality advantage while negative values indicate a mortality disadvantage.

Figure 2 demonstrates the AME of being a migrant versus a non-migrant (Model A1) across birth cohort. We include the full regression estimates in Table B1 of the Online Supplement. While there is fluctuation in the size of the AME, it remains positive and statistically significant for every cohort in our sample, indicating that migrants had a mortality advantage over their non-migrant peers across all birth cohorts. The AME is highest for the 1840 birth cohort, which suggests a roughly 8.5-year, 95% CI [7.8, 9.3] mortality advantage of migrants over non-migrants.

[INSERT FIGURE 2 ABOUT HERE]

Next, we turn to the predicted contrasts by destination country (Model B1; see Figure 3). Here, the AMEs can be interpreted as the expected increase in age at death for migrants to a specific country. We include the total AME from Model A1 as a dashed green line and 95% confidence intervals to demonstrate the contrasts between the overall migrant advantage and individual countries' advantages. When comparing migrants to non-migrants overall in Model A1, we find an AME of 5.9 years, 95% CI [5.7, 6.2] for individuals who migrate, suggesting a 5.9-year mortality advantage of migrants over non-migrants. However, this model masks heterogeneity by destination country. When disaggregating by destination country in Model B1, we find that the

AME of migrating ranges from 2.6 years, 95% CI [1.1, 4.0] in South Africa to 8.7 years, CI [6.3, 11.2] in New Zealand. The AMEs for all destinations are positive and highly statistically significant ($p < 0.001$), suggesting a clear mortality advantage for migrants. Comparing models A1 and B1 shows that the general migrant mortality advantage is similar to the mortality advantage among migrants to Australia and the United States, while migrants to Canada and New Zealand have a greater advantage, and migrants to South Africa have less of an advantage.

[INSERT FIGURE 3 ABOUT HERE]

Supplementary Models

We estimate several supplementary models to be sure of the robustness of our results. We first limit our sample to only individuals with at least one sibling in the data (Models A2 and B2; N = 49,263). The results are substantively similar to the previous models, both for the migrant model and by destination country (see Table B2 in the Online Supplement for regression estimates). We interpret the similarities between the set of models as an indication of model robustness, as well as that the migrant mortality advantage is generally stronger among those with siblings, including sibling groups with mixed migration statuses.

In the next set of supplementary models (Models A3 and B3) we repeat our original analyses, conditioning on survival to alternative ages, i.e., 0, 5, 10, 20, 25, …, 50. In doing so, we attempt to correct for the fact that the migrant mortality advantage may vary across age (Guillot et al., 2018), or may be sensitive to alternative cut-off points. Figure 4 shows the AMEs from Model A3, which show a consistent migrant mortality advantage when conditioning on survival through age 15. As the minimum age increases, the extent of the mortality advantage decreases, though it remains positive and statistically significant ($p < 0.001$). The size of the AME decreases with

higher age cut-offs because samples with higher cutoffs have fewer individuals who die young and are biased towards those who live longer. As such, with higher age cut-offs there is less variability in age at death, so the difference between migrants and non-migrants diminishes when looking exclusively at a sample that dies older.

[INSERT FIGURE 4 ABOUT HERE]

We find a similar pattern when looking at differences by destination country in model B3 (see Figure 5): the extent of the migrant mortality advantage is constant through a cut-off of age 15, though this advantage decreases as the cut-off age increases. The AMEs are positive and statistically significant for most estimates except for cut-offs over age 25 in South Africa. The former point indicates that our findings are generally robust across age cut-off, though the estimated extent of the migrant mortality advantage varies. The latter point indicates that when the age cut-off is above 25, we do not find statistically significant differences in age at death between migrants to South Africa and non-migrants.

[INSERT FIGURE 5 ABOUT HERE]

Discussion

Using the case of emigration from the United Kingdom and Ireland to the United States, Australia, New Zealand, Canada and South Africa during the 18th and 19th centuries, we test the migrant mortality advantage hypothesis in historical migration flows. Using mixed effect regression models, we find that migrants live 5.9 years, 95% CI [5.7, 6.2] longer on average than non-migrants. The gap is heterogeneous by destination country, ranging from 2.6 years, [1.1, 4.0] in South Africa to 8.7 years [6.3, 11.2] in New Zealand. These findings are robust to alternate age cut offs and sample specifications. Taken together, this suggests that the migrant mortality

advantage is robust across birth cohort, even after accounting for unobserved similarities within families and comparing migrants to non-migrants in the United Kingdom and Ireland. We also highlight the possibility for online crowdsourced genealogical data to contribute to demographic research.

Our findings build on a long tradition of literature focused on understanding the existence and extent of the migrant mortality advantage across contexts. While the literature originally focused on the paradoxical mortality advantages found among Hispanic migrants living in the United States (Abraído-Lanza et al., 1999; Pablos-Méndez, 1994; Palloni & Arias, 2004), over time it has come to describe a range of the experiences of migrant groups in various destination countries, ranging from historical flows (Mourits & Puschmann, 2023; Puschmann et al., 2016) to more recent ones (Andersson & Drefahl, 2017; Guillot et al., 2023; Helgesson et al., 2019; Mehta et al., 2016). These studies have tended to focus on non-migrants in the destination as their comparison group. We argue that it is necessary to understand the extent to which migrants hold a mortality advantage over their compatriots left behind – especially so as many of the ones left behind are family members such as siblings who provide a stronger counterfactual group for the existence and extent of the migrant mortality advantage.

Like much of the literature (Feliciano, 2020), we confirm the existence of a migrant mortality advantage for migrants from the United Kingdom and Ireland during the 18th and 19th centuries. We also find that this advantage ranges between migrants' country of destination. This variation may be evidence of positive selection *among* migrants. For example, some destination countries such as Australia and New Zealand saw more highly skilled immigration while migrants to the United States and Canada tended to be more unskilled (Haines, 1997; Hatton, 2021; Murdoch, 2004). Yet, relative to the native-born, migrants from the United Kingdom and Ireland

to the United States tended to earn more even soon after arrival (Abramitzky et al., 2014). Reasons to migrate changed across time as well: what was originally the forced migration of convicts to Australia later became a flow of migrants seeking land and wealth in the Gold Rush (Murdoch, 2004). Another interpretation is that distance played a role in shaping the extent of the mortality advantage. Migrant survival advantages are stronger the further they move (Puschmann et al., 2016, 2017), which may be because the journey from the United Kingdom and Ireland to Australia and New Zealand was much longer than the journey to the United States and Canada.

Alternatively, the discrepancy may be a result of varying conditions across destination countries that would have differential effects on mortality. Specifically, variations in the scale, type, and intensity of conflict in British colonies during the 18th and 19th Centuries may explain the observed differences in mortality advantages for migrants. British settlers in South Africa were met with large-scale, centrally coordinated military resistance by indigenous groups over more than a century, causing heavy casualties to the British (Evans *et al.*, 2018, Beckett, 2019). Similar centrally organised resistance took place in the US and New Zealand, but the scale of losses for colonial forces was smaller, particularly in New Zealand (Draper, Kerby and Baguley, 2023, Carpenter, 2015). Canada and Australia on the other hand featured decentralized or non-confrontational forms of resistance such as economic sabotage (Kerkhove, 2020; Duckett White, 2023). These differences may partially explain the varied mortality advantage across destinations, particularly the small advantage for South Africa relative to other destinations. This explanation is also consistent with Hatton's (2004; 2021) argument that mortality differences among migrants are driven by varying conditions in the destination country rather than the origin country. A third possible mechanism driving the migrant mortality advantage may be that time spent in either the origin or destination country (i.e., one's exposure to potentially unfavorable conditions in the

origin country) may contribute towards a "weathering" effect on one's health. For example, Bakhtiari (2022) notes the *absence* of a mortality advantage for many European immigrant groups living in 20th century United States. In our application, exposure to diseases such as smallpox, in combination with persistent poverty and overcrowding throughout the United Kingdom and Ireland may have shortened non-migrants' lifespans, while migrants were able to avoid such conditions. This framing runs contrary to the typical narrative of the migrant mortality advantage because it posits that mortality advantages are due to poorer conditions experienced by non-migrants which depress their mortality. In reality, the observed variations in the migrant mortality advantage are likely a combination of all the above mechanisms: selection into destinations, differing mortality conditions in each destination, and avoiding deleterious conditions in the origin country. Disentangling these effects is not possible with the present dataset but may prove to be a promising direction for future research to explore.

Limitations

Our study suffers from a few key limitations. Several biases are inherent to crowdsourced genealogical datasets (Calderón Bernal et al., 2025): recorded people may only be those who had children, family trees may be incompletely reconstructed, and there is unequal access to information about relatives. For example, these biases may lead to childless women being less likely to be recorded or the underreporting of infant and child mortality. While we attempt to ameliorate the latter bias by conditioning our sample on survival to age 15, we do identify a gender bias in the recording of women, leading to an overrepresentation of men in the sample. Minardi et al. (2023) notes however that both the median and mean age at death for men and women were

similar during much of our study period, and the female advantage in lifespan likely did not occur until after the 1850s. Thus, we assume minimal impact of this gender bias on our findings.

Additionally, we classify individuals who moved between Ireland and the United Kingdom as “non-migrants" rather than “migrants" because our interest was primarily in migration to Canada, the United States, South Africa, New Zealand, and Australia. As such, we do not investigate the possibility of a migrant mortality advantage between the United Kingdom and Ireland. We also do not capture migration between countries within the United Kingdom nor domestic rural-urban migration during this period. These limitations are in part due to the ambiguous usage of “United Kingdom” to describe locations within the dataset and the difficulty in extracting more granular location data from free text entries. Future work should certainly investigate this possibility, as migration within the United Kingdom and Ireland occurred at non-trivial levels, often driven by many of the same factors that encouraged migration to the countries in this study (Nicholas & Shergold, 1987).

Finally, though Familinx is rich in terms of the overall size of the data, it suffers from a lack of detail at the individual level. As a result, we are unable to control for socioeconomic status (i.e., literacy, financial resources, etc.) directly, which may influence not only one’s propensity to migrate but also their mortality. We are also unable to determine when an individual migrated or account for multiple moves within one’s life, although future work could attempt to account for this using the birth location of children as a proxy for timing of migration. Migrant outcomes may be partially determined by one’s country of residence during early life (Alexander & Ward, 2018), but we are unable to capture one’s “exposure” to different epidemiological or socioeconomic conditions at specific points in their lifespan. Though it is possible to identify some “salmon” migrants who migrated abroad but returned to die in their home country using the birthplace of

their children, we are unable to do so for childless individuals. Given the length of the journey from the UK and Ireland to our destinations of interest and the costs involved, return migrants are likely a small portion of migrants who would not substantially change our results. However, future studies should investigate the impact of exposure to poor conditions in the origin country as a mechanism depressing non-migrants' mortality which some individuals avoid through migration.

Though our setting is historical, our findings are relevant for present-day migration flows. The migrant mortality advantage literature has focused on the paradox of immigrants outliving the native-born even when of a lower socioeconomic status (Abraído-Lanza et al., 1999; Boen & Hummer, 2019; Palloni & Arias, 2004). Future work on the mechanisms driving this mortality advantage is necessary to help ascertain whether it is an associational or causal relationship, that is, whether migrants would have always outlived their non-migrant peers, or whether migration *causes* a longer life.

References

Abraído-Lanza, A. F., Dohrenwend, B. P., Ng-Mak, D. S., & Turner, J. B. (1999). The Latino mortality paradox: A test of the "salmon bias" and healthy migrant hypotheses. *American Journal of Public Health*, *89*(10), 1543–1548. https://doi.org/10.2105/AJPH.89.10.1543

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012). Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration. *American Economic Review*, *102*(5), 1832–1856. https://doi.org/10.1257/aer.102.5.1832

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2014). A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration. *Journal of Political Economy*, *122*(3), 467–506. https://doi.org/10.1086/675805

Alburez-Gutierrez, D., Aref, S., Gil-Clavel, B. S., Grow, A., Negraia, D. V., & Zagheni, E. (2019). Demography in the digital era: New data sources for population research. In *Smart statistics for smart applications: Book of short papers SIS2019* (pp. 23–30). Pearson. https://doi.org/10.31235/osf.io/24jp7

Aldridge, R. W., Nellums, L. B., Bartlett, S., Barr, A. L., Patel, P., Burns, R., Hargreaves, S., Miranda, J. J., Tollman, S., Friedland, J. S., & Abubakar, I. (2018). Global patterns of mortality in international migrants: A systematic review and meta-analysis. *The Lancet*, *392*(10164), 2553–2566. https://doi.org/10.1016/S0140-6736(18)32781-8

Alexander, R., & Ward, Z. (2018). Age at Arrival and Assimilation During the Age of Mass Migration. *The Journal of Economic History*, *78*(3), 904–937. https://doi.org/10.1017/S0022050718000335

Andersson, G., & Drefahl, S. (2017). Long-Distance Migration and Mortality in Sweden: Testing the Salmon Bias and Healthy Migrant Hypotheses. *Population, Space and Place*, *23*(4), e2032. https://doi.org/10.1002/psp.2032

Arel-Bundock, V., Enevoldsen, N., & Yetman, C. (2018). countrycode: An R package to convert country names and country codes. *Journal of Open Source Software*, *3*(28), 848. https://doi.org/10.21105/joss.00848

Bakhtiari, E. (2022). The Missing Mortality Advantage for European Immigrants to the United States in the Early Twentieth Century. *Demography*, *59*(4), 1517–1539. https://doi.org/10.1215/00703370-10111916

Baranowska-Rataj, A., Barclay, K., & Kolk, M. (2017). The effect of number of siblings on adult mortality: Evidence from Swedish registers for cohorts born between 1938 and 1972. *Population Studies*, *71*(1), 43–63. https://doi.org/10.1080/00324728.2016.1260755

Blanc, G., & Wacziarg, R. (2025). *Malthusian Migrations* (Working Paper 33542; NBER Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w33542

Boen, C. E., & Hummer, R. A. (2019). Longer—But Harder—Lives?: The Hispanic Health Paradox and the Social Determinants of Racial, Ethnic, and Immigrant–Native Health Disparities from Midlife through Late Life. *Journal of Health and Social Behavior*, *60*(4), 434–452. https://doi.org/10.1177/0022146519884538

Calderón Bernal, L., Alburez-Gutierrez, D., & Zagheni, E. (2025). *Analysing biases in genealogies using demographic microsimulation* (WP-2023-034; MPIDR Working Paper). Max Planck Institute for Demographic Research. https://doi.org/10.4054/MPIDR-WP-2023-034

Chong, M., Alburez-Gutierrez, D., Del Fava, E., Alexander, M., & Zagheni, E. (2022). *Identifying and correcting bias in big crowd-sourced online genealogies* (WP-2022-005). Max Planck Institute for Demographic Research. https://doi.org/10.4054/MPIDR-WP-2022-005

Clouston, S. A. P., & Link, B. G. (2021). A retrospective on fundamental cause theory: State of the literature, and goals for the future. *Annual Review of Sociology*, *47*(1), 131–156. https://doi.org/10.1146/annurev-soc-090320-094912

Colasurdo, A., & Omenti, R. (2024). Using online genealogical data for demographic research: An empirical examination of the FamiLinx database. *Demographic Research*, *51*, 1299–1350. https://doi.org/10.4054/DemRes.2024.51.41

Corti, G., Minardi, S., & Barban, N. (2024). Trends in assortative mating in the United States, 1700–1910. Evidence from FamiLinx data. *The History of the Family*, *29*(4), 461–481. https://doi.org/10.1080/1081602X.2024.2352539

Cournil, A., & Kirkwood, T. B. L. (2001). If you would live long, choose your parents well. *Trends in Genetics*, *17*(5), 233–235. https://doi.org/10.1016/S0168-9525(01)02306-X

Cozzani, M., Minardi, S., Corti, G., & Barban, N. (2023). Birth month and adult lifespan: A within-family, cohort, and spatial examination using FamiLinx data in the United States (1700–1899). *Demographic Research*, *49*, 201–218. https://doi.org/10.4054/DemRes.2023.49.9

Donrovich, R., Puschmann, P., & Matthijs, K. (2014). Rivalry, solidarity, and longevity among siblings: A life course approach to the impact of sibship composition and birth order on later life mortality risk, Antwerp (1846-1920). *Demographic Research*, *31*, 1167–1198. https://doi.org/10.4054/DemRes.2014.31.38

Feliciano, C. (2020). Immigrant Selectivity Effects on Health, Labor Market, and Educational Outcomes. *Annual Review of Sociology*, *46*(1), 315–334. https://doi.org/10.1146/annurev-soc-121919-054639

Friedlander, D., & Okun, B. S. (2022). Demographic Transition and the Industrial Revolution in England: Inverse Rural and Urban Processes. *Journal of Family History*, *47*(4), 401–412. https://doi.org/10.1177/03631990221114065

Gavrilov, L. A., Gavrilova, N. S., Olshansky, S. J., & Carnes, B. A. (2002). Genealogical data and the biodemography of human longevity. *Biodemography and Social Biology*, *49*(3–4), 160–173. https://doi.org/10.1080/19485565.2002.9989056

Gay, V., Gobbi, P. E., & Goñi, M. (2023). *Revolutionary Transition: Inheritance Change and Fertility Decline*. CEPR. https://hal.science/hal-04285818

Gudmundsson, H., Gudbjartsson, D. F., Kong, A., Gudbjartsson, H., Frigge, M., Gulcher, J. R., & Stefánsson, K. (2000). Inheritance of human longevity in Iceland. *European Journal of Human Genetics*, *8*(10), Article 10. https://doi.org/10.1038/sj.ejhg.5200527

Guillot, M., Khlat, M., Elo, I., Solignac, M., & Wallace, M. (2018). Understanding age variations in the migrant mortality advantage: An international comparative perspective. *PLOS ONE*, *13*(6), e0199669. https://doi.org/10.1371/journal.pone.0199669

Guillot, M., Khlat, M., Gansey, R., Solignac, M., & Elo, I. (2023). Return Migration Selection and Its Impact on the Migrant Mortality Advantage: New Evidence Using French Pension Data. *Demography*, 10938784. https://doi.org/10.1215/00703370-10938784

Haines, R. (1997). 'The idle and the drunken won't do there': Poverty, the new poor law and nineteenth-century government-assisted emigration to Australia from the United Kingdom∗. *Australian Historical Studies*, *27*(108), 1–21. https://doi.org/10.1080/10314619708596025

Hatton, T. J. (2004). Emigration from the UK, 1870-1913 and 1950-1998. *European Review of Economic History*, *8*(2), 149–171. https://doi.org/10.1017/S1361491604001121

Hatton, T. J. (2021). Emigration from the United Kingdom to the United States, Canada and Australia/New Zealand, 1870–1913: Quantity and quality. *Australian Economic History Review*, *61*(2), 136–158. https://doi.org/10.1111/aehr.12218

Hatton, T. J., & Martin, R. M. (2008). *The Effects on Stature of Poverty, Family Size and Birth Order: British Children in the 1930s* (3314; IZA Discussion Paper Series). IZA. https://docs.iza.org/dp3314.pdf

Hatton, T. J., & Williamson, J. G. (1994). What Drove the Mass Migrations from Europe in the Late Nineteenth Century? *Population and Development Review*, *20*(3), 533. https://doi.org/10.2307/2137600

Helgesson, M., Johansson, B., Nordquist, T., Vingård, E., & Svartengren, M. (2019). Healthy migrant effect in the Swedish context: A register-based, longitudinal cohort study. *BMJ Open*, *9*(3), e026972. https://doi.org/10.1136/bmjopen-2018-026972

Horn, J. (1998). British Diaspora: Emigration from Britain, 1680–1815. In P. J. Marshall, A. Low, & Wm. R. Louis (Eds.), *The Oxford History of the British Empire: Volume II: The Eighteenth Century* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198205630.003.0002

Kaplanis, J., Gordon, A., Shor, T., Weissbrod, O., Geiger, D., Wahl, M., Gershovits, M., Markus, B., Sheikh, M., Gymrek, M., Bhatia, G., MacArthur, D. G., Price, A. L., & Erlich, Y. (2018). Quantitative analysis of population-scale family trees with millions of relatives. *Science*, *360*(6385), 171–175. https://doi.org/10.1126/science.aam9309

Kasakoff, A. B., & Adams, J. W. (1995). The Effect of Migration on Ages at Vital Events: A Critique of Family Reconstitution in Historical Demography. *European Journal of Population*, *11*(3), 199–242.

Lam, D., & Marteleto, L. (2008). Stages of the Demographic Transition from a Child's Perspective: Family Size, Cohort Size, and Children's Resources. *Population and Development Review*, *34*(2), 225–252. https://doi.org/10.1111/j.1728-4457.2008.00218.x

Lindstrom, D. P., & Lauster, N. (2001). Local Economic Opportunity and the Competing Risks of Internal and U.S. Migration in Zacatecas, Mexico. *International Migration Review*, *35*(4), 1232–1256. https://doi.org/10.1111/j.1747-7379.2001.tb00059.x

Link, B. G., & Phelan, J. (1995). Social Conditions As Fundamental Causes of Disease. *Journal of Health and Social Behavior*, *35*, 80. https://doi.org/10.2307/2626958

Mehta, N. K., Elo, I. T., Engelman, M., Lauderdale, D. S., & Kestenbaum, B. M. (2016). Life Expectancy Among U.S.-born and Foreign-born Older Adults in the United States: Estimates From Linked Social Security and Medicare Data. *Demography*, *53*(4), 1109–1134. https://doi.org/10.1007/s13524-016-0488-4

Minardi, S., Corti, G., & Barban, N. (2023). *Historical Patterns in the Intergenerational Transmission of Lifespan and Longevity: Evidence from the United States, 1700-1900* [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/gfxcm

Minardi, S., Corti, G., & Barban, N. (2024). Historical Patterns in the Intergenerational Transmission of Lifespan and Longevity: A Research Note on U.S. Cohorts Born Between 1700 and 1900. *Demography*, *61*(4), 979–994. https://doi.org/10.1215/00703370-11458359

Mize, T. (2019). Best Practices for Estimating, Interpreting, and Presenting Nonlinear Interaction Effects. *Sociological Science*, *6*, 81–117. https://doi.org/10.15195/v6.a4

Mourits, R. J., & Puschmann, P. (2023). Exploring familial factors in the migrant mortality advantage among domestic migrants in later life: Zeeland, the Netherlands, 1812–1962. *SSM - Population Health*, *22*, 101359. https://doi.org/10.1016/j.ssmph.2023.101359

Murdoch, A. (2004). *British emigration, 1603-1914*. Palgrave Macmillan.

Nicholas, S., & Shergold, P. R. (1987). Internal migration in England, 1818–1839. *Journal of Historical Geography*, *13*(2), 155–168. https://doi.org/10.1016/S0305-7488(87)80144-5

Pablos-Méndez, A. (1994). Mortality Among Hispanics. *JAMA: The Journal of the American Medical Association*, *271*(16), 1237. https://doi.org/10.1001/jama.1994.03510400023017

Palloni, A., & Arias, E. (2004). Paradox lost: Explaining the Hispanic adult mortality advantage. *Demography*, *41*(3), 385–415. https://doi.org/10.1353/dem.2004.0024

Piraino, P., Muller, S., Cilliers, J., & Fourie, J. (2014). The transmission of longevity across generations: The case of the settler Cape Colony. *Research in Social Stratification and Mobility*, *35*, 105–119. https://doi.org/10.1016/j.rssm.2013.08.005

Puschmann, P., Donrovich, R., Grönberg, P.-O., Dekeyser, G., & Matthijs, K. (2016). Disfavored in Life, Favored in Death? Later-Life Mortality Differences (Ages 30+) between Migrants and Natives in Antwerp, Rotterdam and Stockholm, 1850-1930. *Historical Social Research*, *41*(4). https://doi.org/10.12759/HSR.41.2016.4.257-290

Puschmann, P., Donrovich, R., & Matthijs, K. (2017). Salmon Bias or Red Herring?: Comparing Adult Mortality Risks (Ages 30–90) between Natives and Internal Migrants: Stayers, Returnees and Movers in Rotterdam, the Netherlands, 1850–1940. *Human Nature*, *28*(4), 481–499. https://doi.org/10.1007/s12110-017-9303-1

Razum, O. (2008). Migrant Mortality, Healthy Migrant Effect. In W. Kirch (Ed.), *Encyclopedia of Public Health* (pp. 932–935). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5614-7_2188

Richards, E. (2018). *The genesis of international mass migration: The British case, 1750-1900*. Manchester University Press.

Ruggles, S. (1992). Migration, Marriage, and Mortality: Correcting Sources of Bias in English Family Reconstitutions. *Population Studies*, *46*(3), 507–522. https://doi.org/10.1080/0032472031000146486

Sonneveldt, E., DeCormier Plosky, W., & Stover, J. (2013). Linking high parity and maternal and child mortality: What is the impact of lower health services coverage among higher order births? *BMC Public Health*, *13*(Suppl 3), S7. https://doi.org/10.1186/1471-2458-13-S3-S7

Stark, O., & Bloom, D. E. (1985). The New Economics of Labor Migration. *The American Economic Review*, *75*(2), 173–178.

Stelter, R., & Alburez-Gutierrez, D. (2022). Representativeness is crucial for inferring demographic processes from online genealogies: Evidence from lifespan dynamics. *Proceedings of the National Academy of Sciences*, *119*(10), e2120455119. https://doi.org/10.1073/pnas.2120455119

Swerdlow, A. J. (1991). Mortality and Cancer Incidence in Vietnamese Refugees in England and Wales: A Follow-Up Study. *International Journal of Epidemiology*, *20*(1), 13–19. https://doi.org/10.1093/ije/20.1.13

Taylor, D. (1988). Public Health in Britain 1750–1900. In D. Taylor, *Mastering Economic and Social History* (pp. 301–331). Macmillan Education UK. https://doi.org/10.1007/978-1-349-19377-6_19

Thompson, R. (2009). *Mobility and Migration: East Anglian Founders of New England, 1629-1640* (Reprint edition). University of Massachusetts Press.

Tomlins, C. (2001). Reconsidering Indentured Servitude: European Migration and the Early American Labor Force, 1600–1775. *Labor History*, *42*(1), 5–43. https://doi.org/10.1080/00236560123269

Tables

Table 1 Migrant mortality comparison approaches

Comparison	Data required	Interpretation	Example studies
Migrants and origin country non-migrants	Data from both the origin country and the destination country	There is a mortality advantage for migrants over their peers in their origin country; migrants deviate from the baseline mortality in their home countries	Abramitzky et al. (2012); Mourits & Puschmann (2023)
Migrants and destination country non-migrants	Data from the destination country	There is a mortality advantage for migrants over individuals in the destination country, though it is not clear if this is due to migration or selection	Abraído-Lanza et al. (1999); Bakhtiari (2022); Palloni & Arias

Table 2 Sample descriptive statistics

Variable	Mean/Prop.	SD	Min	Max
Number of siblings	1.604	2.324	0	21
Age at death	65.870	18.272	15	109
Proportion male	0.606	–	–	–
Proportion migrant	0.367	–	–	–
Proportion destination USA	0.511	–	–	–

Note: N = 98,057

Figures

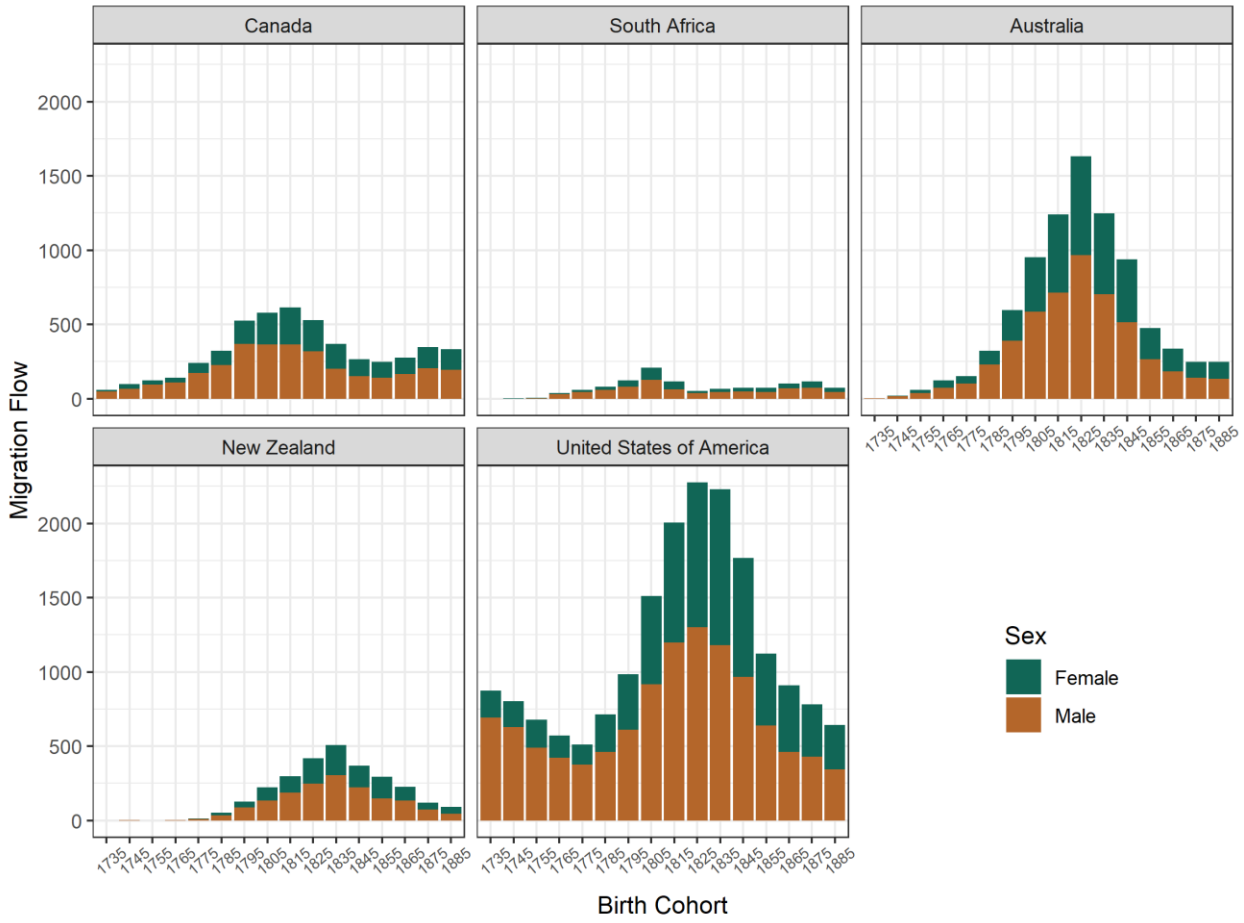

Figure 1 Observed migration flows in Familinx sample by sex, birth cohort and destination.

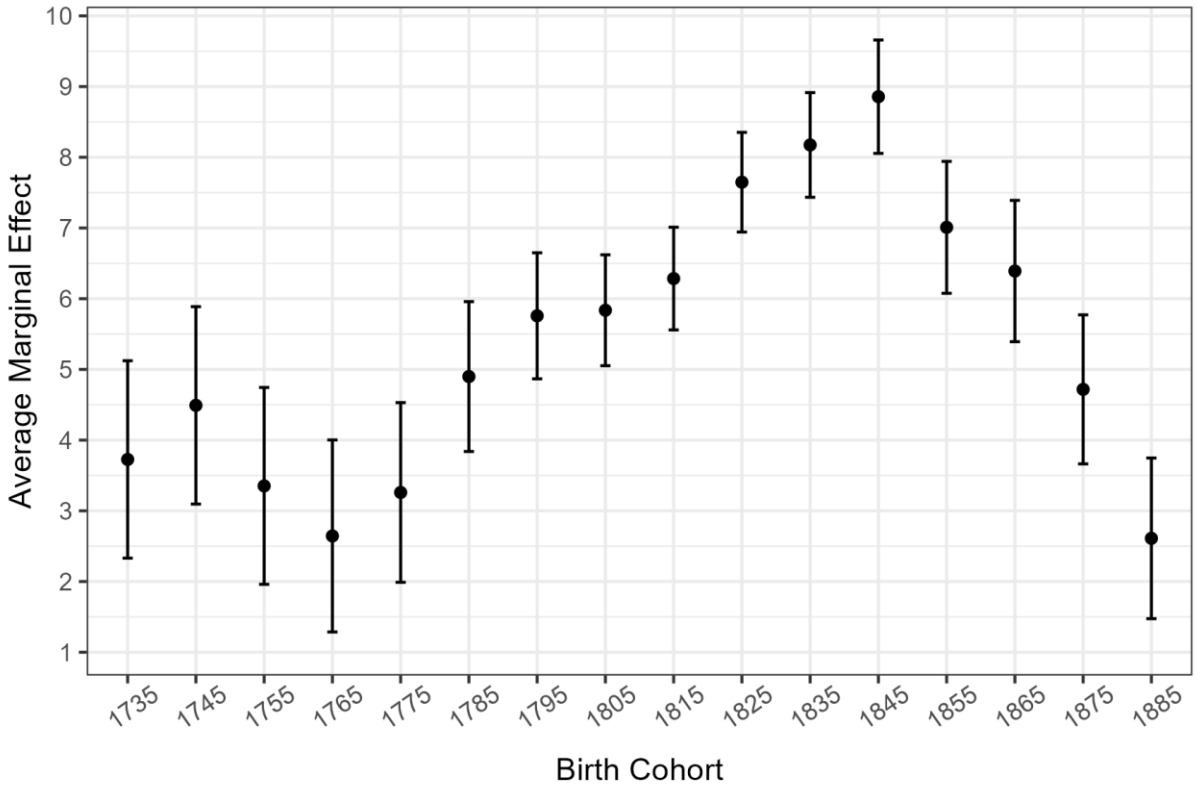

Figure 2 Average marginal effect of migration on age at death, across birth cohort (Model A1).

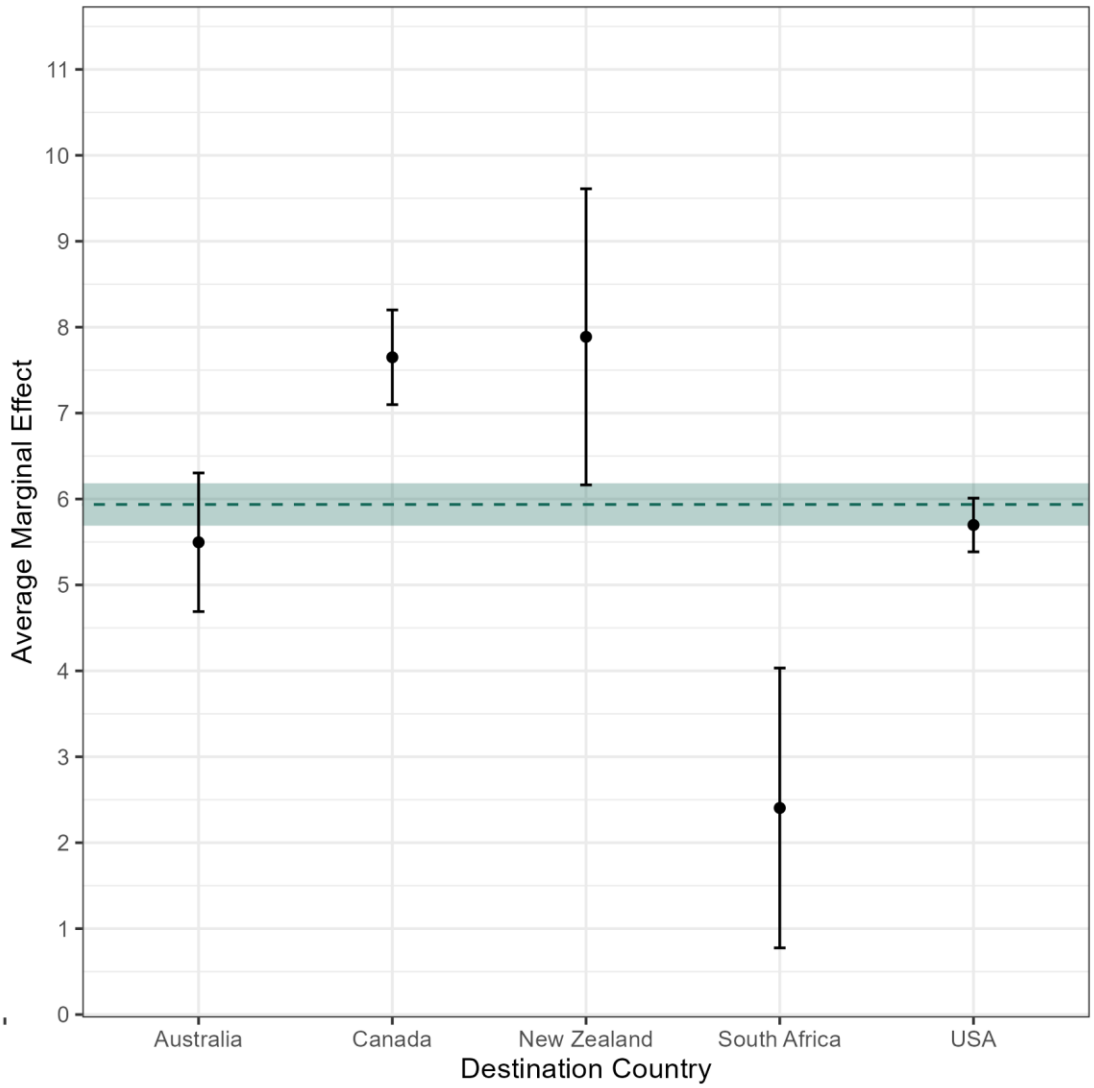

Figure 3 AME of migrating on age at death by destination country (Model B1) in black with 95% CIs. AME for Model A1 (migrant versus non-migrant) is shown by the dashed line, with 95% CIs as green shade.

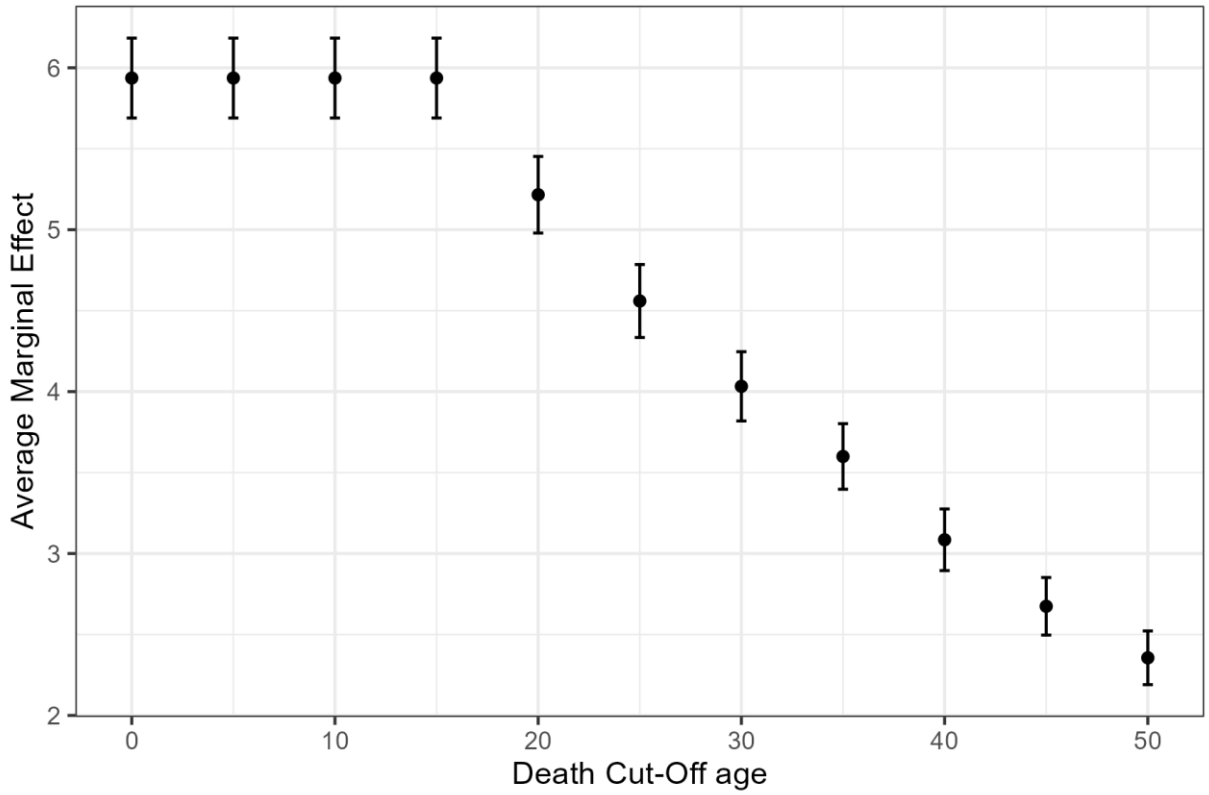

Figure 4 AMEs of migration on age at death across alternate age cutoffs (Model A3).

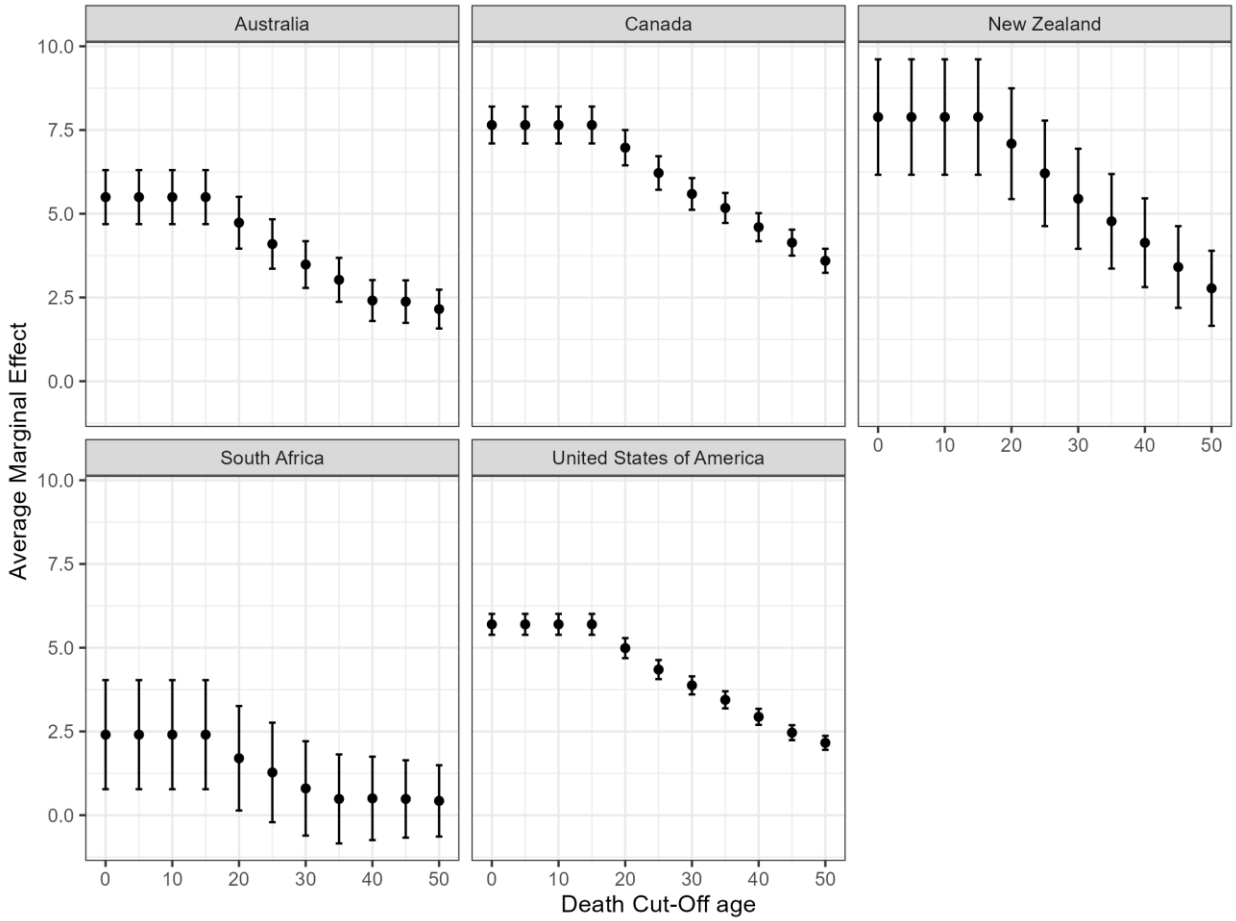

Figure 5 AMEs of migration on age at death across alternate age cutoffs and disaggregated by destination country (Model B3).

Leaving for Life: Using Online Crowd-Sourced Genealogies to Estimate the Migrant Mortality Advantage for the United Kingdom and Ireland During the 18th and 19th Centuries

Online Supplement

Table of Contents

Appendix A. Data Cleaning & Imputation

We began with a dataset of over 86 million individuals, many of whom had missing values across most variables. In preparing the data for imputations, only individuals with observed kinship ties were kept from the full data set. To address the high degree of missingness in key variables among the remaining observations, we imputed several types of missingness: (1) missingness in the birth and death year, (2) missingness in the birth and death location, and (3) missing gender. Figure A1 offers a visual representation of our data cleaning process.

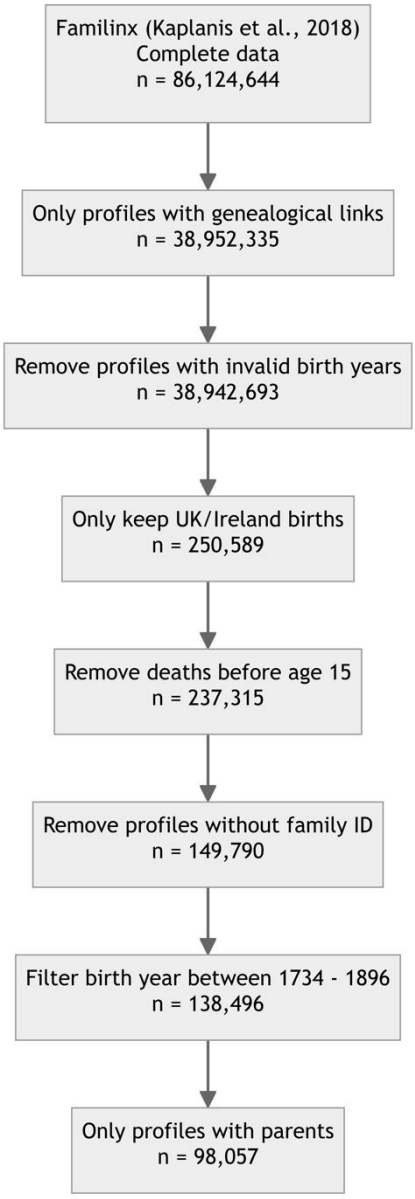

Figure A1 Flow of data cutting process

For type (1), we imputed missing birth and death years with birth and baptism years, which are included in the Familinx dataset. Similarly for type (2), we used coordinate information on birth, death, baptism, and burial to extract countries of interest for each event. Missing birth and death locations were imputed using baptism and burial locations, respectively. This included baptism and burial locations that were derived from coordinate information. Finally, entries containing abnormal birth years (negative years and years occurring after 2020) that likely contained errors were removed.

The coordinate information matching process was conducted as follows. Using the *countrycode* package (Arel-Bundock et al., 2018), columns containing country code information were then used to fill in entries that were still missing in these four columns, with imputed baptism and burial information from their respective country codes being used to fill in country level information. The *countrycode* package was also used to fill in birth and death countries with "USA" from the state-based location columns if they matched names or abbreviations of U.S. states. Finally, the free text columns were used to extract country level information through regex for rows that were still missing country information. This was repeated on the baptism and burial free text columns which were then used to impute birth and death country information that was still missing.

To fill in type (3) information for individuals with a missing gender, kinship ties were used. When two individuals were both parents of another individual in the data set and the gender information for one parent was known, this was used to fill in the gender information for the parent missing gender information.

To create the final country categories, a combination of term matching and regex approaches were used. For Ireland and Northern Ireland, regex was used to detect references to Ireland or "eire". For England, Scotland, Wales, a term matching approach was adopted to match various spellings, languages and cities referenced. Finally, for the USA, Australia, South Africa, New Zealand and Canada, terms matching in the country column included references to territories, regions, states (that had not already been picked up previously), alternate spellings and names in other languages. The full list of terms used for term matching is presented in Table A1.

Table A1 Term matching terms

Country	Term
United Kindgom	'united kingdom', 'x-england', 'england', 'x-scotland', 'x-united-kingdom', 'scotland', '(present uk)', 'x-great-britain', 'x-wales', 'gb', 'x-northern-ireland', 'uk', 'northern ireland', 'england, uk', 'u.k.', 'wales', 'england, united kingdom', 'great britain', 'england uk', 'uk:great britain', 'uk:northern ireland', 'england (present uk)', 'scotland, united kingdom', 'uk:northern ireland', 'uk.', 'scotland, uk', 'uk:isle of wight', 'ireland (present northern ireland)', 'north ireland', 'scotland, uk', 'scotland, united kingdom', 'scotland uk', 'south wales', 'britain', 'england/ uk', 'n.ireland', 'engand', 'englnd', 'northern ireland, uk', 'middlesex', 'n. ireland', 'huntingdonshire', 'lancashire', 'london', 'uk:scotland:shetland islands:mainland', 'u k', 'united kingdom of great britain and ireland', 'northern-ireland', 'nothern ireland, uk', 'ireland (northern)', 'n. ireland', 'ireland or scotland', 'scotland or ireland', 'enfland', 'storbritannia', 'bonhill, dunbartonshire, scotland', 'cambridgeshire', 'chatham', 'crickdale, wiltshire, uk', 'eicester, leicestershire, uk', 'endgland', 'england.', 'englanmd', 'englnad', 'enland', 'essex', 'fifeshire', 'gloucester', 'great briatin', 'great britai', 'herfordshire', 'lower bebington', 'newcastle

Country	Term
	upon tyne', 'north wales', 'reino unido', 'royaume uni', 'scotlans', 'swindon, wiltshire, england', 'tyrone', 'uk /england', 'uk/wales', 'uk:wales:anglesey', 'umited kingdom', 'united kinbgdom', 'united kinbgdom', 'winwick', 'woolwich, kent, uk', 'yorkshire', 'xengland', "yhdistynyt kuningaskunta", "(present u.k.)", "(present day united kingdom)", "-england", "(now united kingdom)", "present united kingdom)", ","הממלכה המאוחדת "aberdeen city", "englang", "england,uk", "england or pa", "wales or england", "verenigd koninkrijk", "unitedkingdom", "united-kingdom", "uk:scotland:barra", "uk, england", "u. k.", "heathfield, sussex, england", "united kingdrom"
Ireland	'ireland', 'ie', 'republic of ireland', 'eire', 'bydoney,tyrone ,ireland', 'ireland, uk', 'uk (ireland)', 'ireland ???', 'ireland (eire)', 'ireland.', 'or ireland', 'kilkenny', 'tipperary', 'waterford', 'ulster', 'galway', 'down', 'carlow'
United Stats of America	us", 'usa', 'united states', 'united states of america', 'america', '(present usa)', 'colonial america', 'province of new york', 'new netherland colony', 'new england colonies', 'new england', 'present united states', 'american colonies', "british america", "new netherlands", 'new york', 'american colonies [present united states]', 'british north america', 'u.s.a.', 'new netherlands (usa)', 'nieuw netherlands', 'british colonies', 'nouvelle france', "british colonies of north america", 'american colonies (present usa)', 'usa:new york:long island', 'usa:49', 'usa:massachusettes:nantucket island', 'usa:massachusettes:martha\'s vineyard', 'usa:44', 'the united states of america', 'massachusetts colony', 'british amercia', 'now usa', 'martin county, indiana, usa', 'richland county', 'u,s.a.', 'u.sa.', 'united sates', 'unitet states', 'usa.', 'usa:hawaii:kauai', "ee.uu.", "(currently) united states", "(present usa)", "(present (usa)", "(present) usa", "amerikas forente stater", "cleveland", "estados unidos", "usa:california:santa catalina island", "usa (all present day)", "us virgin islands", "kittery, york, maine", "in what will be america"
Australia	'australia', 'au', 'australien', 'australia:tasmania', 'new south wales', 'australia [green slopes hospital]', 'aust', 'western australia', 'austrailia', 'australis', 'nsw', 'port melbourne', 'portsea', 'sydney nsw', 'tasmania'
New Zealand	'new zealand', 'nz', 'new zealand:north island', 'new zealand:south island', 'new zealand.', ', new zealand', 'christchurch', 'king street, sydenham, christchurch, nz', 'new zeaand', 'new zealand of senile decay', 'new, zealand', 'tennyson street, sydenham, christchurch, nz', "king street, sydenham, christchurch, nz"

Country	Term
Canada	'canada', 'ca', 'kanadas', 'canada:27', 'united province of canada', 'british north america (present canada)', 'canada:cape breton island', 'province of canada', 'canada:11', 'canada:newfoundland', 'upper canada', '(present canada)', 'acadie', 'canad', 'can'', 'canadá', 'canada:15', 'canada:vancouver island'
South Africa	'za', 'south africa', 'suid afrika', 'cape of good hope', 'cape colony', 's africa', 'cape colony (south africa)', 'rep south africa', 'south africa', 'south afica', 'south africa.', 'union of south africa'

To produce the final data set, only rows that were complete (i.e: had no missing information regarding birth/death year, gender and birth/death location) were kept. All individuals born in the UK/Ireland and dying in the UK/Ireland, USA, Australia, Canada, New Zealand, and South Africa were kept. Figure A2 shows the counts of migrants by destination country and birth cohort.

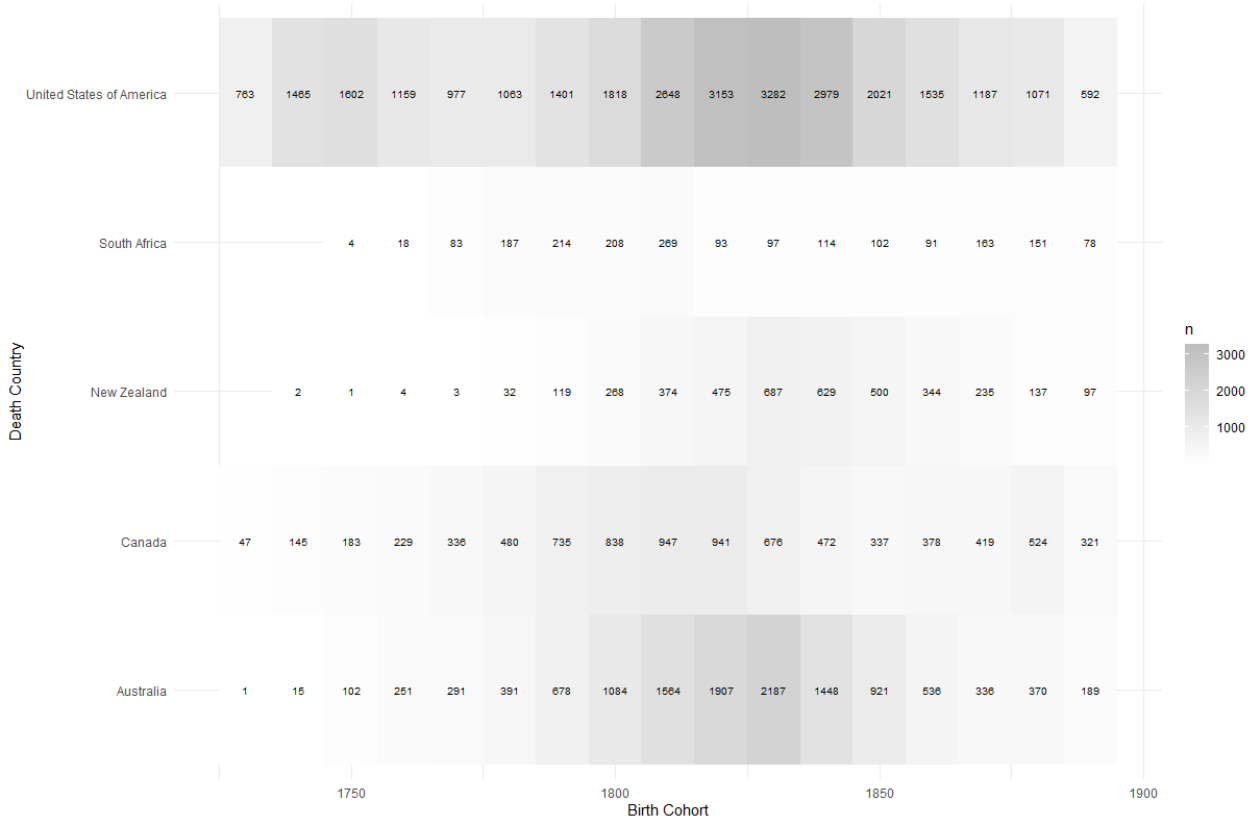

Figure A2 Counts of Migrants by Birth Cohort

Appendix B. Results and Robustness Checks

This section is organized as follows. We first present regression estimates from the main models in Table B1. Next, we present regression estimates from supplementary models that were estimated on the reduced sample of individuals with at least one sibling in Table B2. Finally, we show that our models are robust to alternate minimum age cutoffs in Figure B1 and B2.

Table B1 Results from Main Models

	Dependent variable: Age at Death	
Variable	Model A: Migrant	Model B: Country of Death
Migrant (*reference: Non-Migrant)*	3.350**	
	(1.058)	
Destination country *(reference: United Kingdom/Ireland)*		
Canada		1.242
		(4.637)
South Africa		1.992
		(2.716)
Australia		23.950
		(17.595)
New Zealand		4.728*
		(2.276)
United States of America		3.377**
		(1.071)
Birth cohort (*reference: 1730)*		
1740	-0.743	-0.744
	(0.718)	(0.718)
1750	-0.744	-0.744
	(0.710)	(0.709)
1760	0.957	0.956
	(0.698)	(0.697)
1770	2.068**	2.067**
	(0.680)	(0.679)
1780	1.413**	1.411**
	(0.669)	(0.668)
1790	0.042	0.039
	(0.660)	(0.659)
1800	-0.667	-0.666
	(0.655)	(0.654)
1810	-1.556*	-1.560*
	(0.654)	(0.653)
1820	-2.908**	-2.914**
	(0.650)	(0.649)
1830	-4.918**	-4.919**

	Dependent variable: Age at Death	
Variable	Model A: Migrant	Model B: Country of Death
	(0.653)	(0.653)
1840	-5.655**	-5.657**
	(0.656)	(0.656)
1850	-5.831**	-5.832**
	(0.660)	(0.660)
1860	-4.315**	-4.320**
	(0.665)	(0.665)
1870	-1.971**	-1.974**
	(0.668)	(0.667)
1880	1.414*	1.414*
	(0.675)	(0.674)
1890	3.220**	3.219**
	(0.733)	(0.732)
Male *(reference: female)*	0.0003	0.0002
	(0.117)	(0.117)
Number of siblings *(reference: none)*		
1	-1.860**	-1.861**
	(0.177)	(0.177)
2	-2.927**	-2.926**
	(0.213)	(0.213)
3-5	-2.894**	-2.889**
	(0.179)	(0.179)
6+	-2.468**	-2.439**
	(0.281)	(0.281)
Migrant X Birth cohort interaction *(reference: 1730)*		
1740	0.995	
	(1.286)	
1750	0.233	
	(1.266)	
1760	-0.058	
	(1.278)	
1770	-0.154	
	(1.258)	
1780	0.356	
	(1.214)	
1790	2.281†	
	(1.168)	
1800	2.201†	
	(1.143)	
1810	2.426*	

Variable	Dependent variable: Age at Death	
	Model A: Migrant	Model B: Country of Death
	(1.126)	
1820	3.442**	
	(1.119)	
1830	4.733**	
	(1.119)	
1840	5.172**	
	(1.127)	
1850	5.014**	
	(1.144)	
1860	3.095**	
	(1.165)	
1870	2.511*	
	(1.180)	
1880	-0.062	
	(1.193)	
1890	0.034	
	(1.304)	
Destination country X Birth cohort interaction *(reference: UK/Ireland & 1730)*		
Canada:1740		6.812
		(5.071)
Australia:1740		-27.468
		(18.693)
New Zealand:1740		-19.768
		(12.822)
United States of America:1740		0.717
		(1.314)
Canada:1750		3.462
		(4.989)
South Africa:1750		-9.603
		(12.907)
Australia:1750		-20.242
		(17.823)
New Zealand:1750		21.868
		(17.844)
United States of America:1750		0.051
		(1.300)
Canada:1760		4.672
		(4.918)
South Africa:1760		8.050
		(6.256)

	Dependent variable: Age at Death	
Variable	Model A: Migrant	Model B: Country of Death
Australia:1760		-20.349
		(17.718)
New Zealand:1760		3.531
		(17.984)
United States of America:1760		-0.785
		(1.340)
Canada:1770		6.418
		(4.826)
South Africa:1770		-1.007
		(3.788)
Australia:1770		-20.676
		(17.669)
New Zealand:1770		12.420
		(17.983)
United States of America:1770		-1.663
		(1.365)
Canada:1780		5.180
		(4.765)
South Africa:1780		0.523
		(3.429)
Australia:1780		-19.386
		(17.642)
New Zealand:1780		6.123
		(4.295)
United States of America:1780		-1.398
		(1.329)
Canada:1790		7.663
		(4.726)
South Africa:1790		0.373
		(3.257)
Australia:1790		-18.345
		(17.619)
New Zealand:1790		3.631
		(3.076)
United States of America:1790		0.767
		(1.267)
Canada:1800		5.398
		(4.713)
South Africa:1800		-0.017
		(3.117)

	Dependent variable: Age at Death	
Variable	Model A: Migrant	Model B: Country of Death
Australia:1800		-19.143
		(17.610)
New Zealand:1800		2.873
		(2.663)
United States of America:1800		2.302*
		(1.224)
Canada:1810		6.842
		(4.701)
South Africa:1810		-0.009
		(3.025)
Australia:1810		-19.261
		(17.606)
New Zealand:1810		3.639
		(2.545)
United States of America:1810		2.309†
		(1.183)
Canada:1820		7.709
		(4.701)
South Africa:1820		2.711
		(3.572)
Australia:1820		-18.470
		(17.604)
New Zealand:1820		3.913
		(2.494)
United States of America:1820		3.431**
		(1.165)
Canada:1830		9.520*
		(4.723)
South Africa:1830		1.410
		(3.663)
Australia:1830		-17.703
		(17.603)
New Zealand:1830		2.965
		(2.429)
United States of America:1830		5.660**
		(1.164)
Canada:1840		7.929†
		(4.752)
South Africa:1840		1.880
		(3.418)

	Dependent variable: Age at Death	
Variable	Model A: Migrant	Model B: Country of Death
Australia:1840		-16.297
		(17.606)
New Zealand:1840		4.338†
		(2.454)
United States of America:1840		5.563**
		(1.171)
Canada:1850		7.069
		(4.785)
South Africa:1850		0.987
		(3.452)
Australia:1850		-15.637
		(17.610)
New Zealand:1850		6.153*
		(2.485)
United States of America:1850		4.700**
		(1.202)
Canada:1860		4.712
		(4.774)
South Africa:1860		-2.004
		(3.473)
Australia:1860		-17.647
		(17.620)
New Zealand:1860		5.006*
		(2.550)
United States of America:1860		2.890*
		(1.240)
Canada:1870		6.444
		(4.759)
South Africa:1870		-1.099
		(3.193)
Australia:1870		-18.831
		(17.632)
New Zealand:1870		3.600
		(2.659)
United States of America:1870		2.254†
		(1.271)
Canada:1880		2.450
		(4.745)
South Africa:1880		2.230
		(3.247)

	Dependent variable: Age at Death	
Variable	Model A: Migrant	Model B: Country of Death
Australia:1880		-20.162
		(17.632)
New Zealand:1880		0.615
		(2.894)
United States of America:1880		-0.864
		(1.299)
Canada:1890		2.156
		(4.831)
Australia:1890		-19.513
		(17.674)
United States of America:1890		-0.415
		(1.484)
Constant	66.785**	66.784**
	(0.608)	(0.607)
Sibling random effects	Yes	Yes
Observations	98,057	98,057

Note: † p<0.1 *p<0.05 **p<0.01

Table B2 Results from Supplementary Models, Only Individuals with Siblings

	Dependent variable: Age at Death	
Variable	Migrant	Country of Death
Migrant (*reference: Non-Migrant)*	2.881	
	(2.139)	
Destination country *(reference: United Kingdom/Ireland)*		
Canada		12.404
		(18.608)
South Africa		2.810
		(4.025)
Australia		24.674
		(18.820)
New Zealand		8.164*
		(3.182)
United States of America		2.553
		(2.153)
Male *(reference: female)*	0.221	0.231
	(0.177)	(0.177)
Number of siblings *(reference: 1)*		
2	-1.076**	-1.072**

	Dependent variable: Age at Death	
Variable	Migrant	Country of Death
	(0.271)	(0.272)
3-5	-1.046**	-1.030**
	(0.242)	(0.242)
6+	-0.699*	-0.662†
	(0.341)	(0.341)
Constant	63.838**	63.822**
	(1.029)	(1.029)
Sibling random effects	Yes	Yes
Interaction effects	Yes	Yes
Observations	49,263	49,263

Note: We suppress coefficients for all interactions for space. † p < 0.1, * p < 0.05, ** p < 0.01

Table B3 AMEs of Models A1, B1, A2, B2

	Model			
	Full Sample		Profiles with siblings	
Variable	A1	B1	A2	B2
Migrant	5.9**		7.1**	
	(0.12)		(0.2)	
Australia		5.5**		7.7**
		(0.38)		(0.65)
Canada		7.6**		8.8**
		(0.28)		(0.5)
New Zealand		8.7**		9.8**
		(1.24)		(0.87)
South Africa		2.6**		3.2**
		(0.73)		(1.1)
USA		5.7**		6.8**
		(0.16)		(0.27)
Observations	98,057	98,057	49,263	49,263

Notes: For models A2 and B2 the sample is limited to only individuals with at least one sibling. ** *p < 0.01*

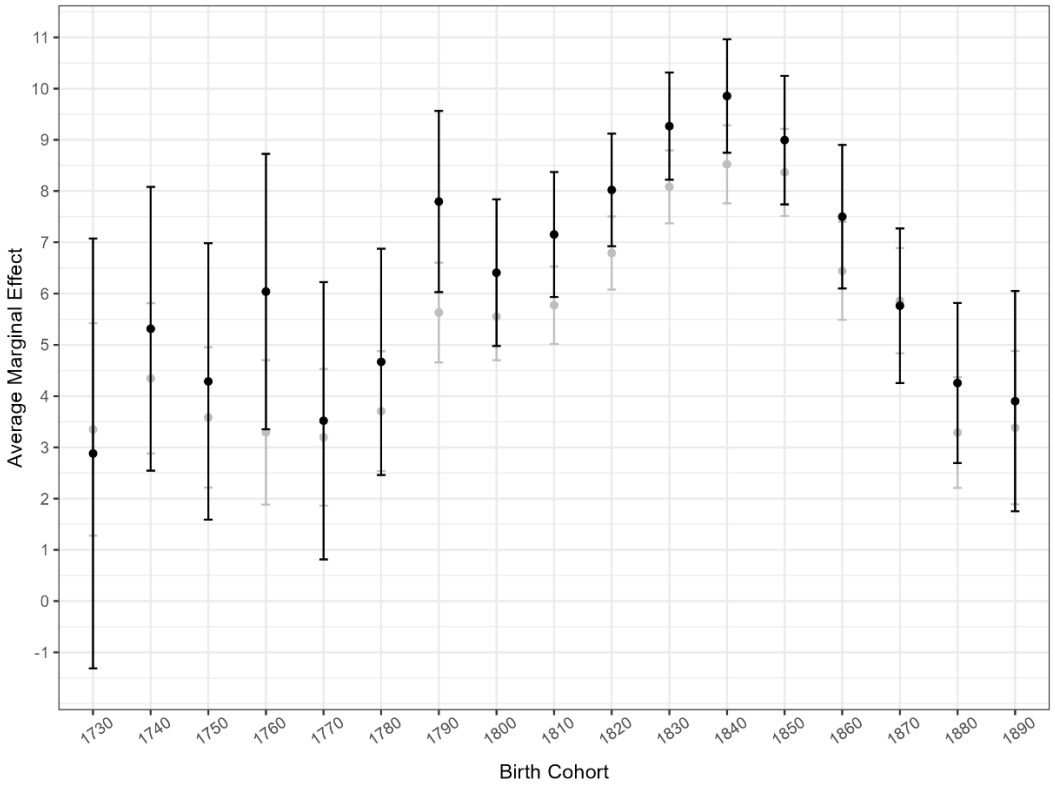

Figure B1 Average marginal effect of migration on age at death, across birth cohort (Model A2) shown in black. AMEs for full sample (Model A1) in grey.

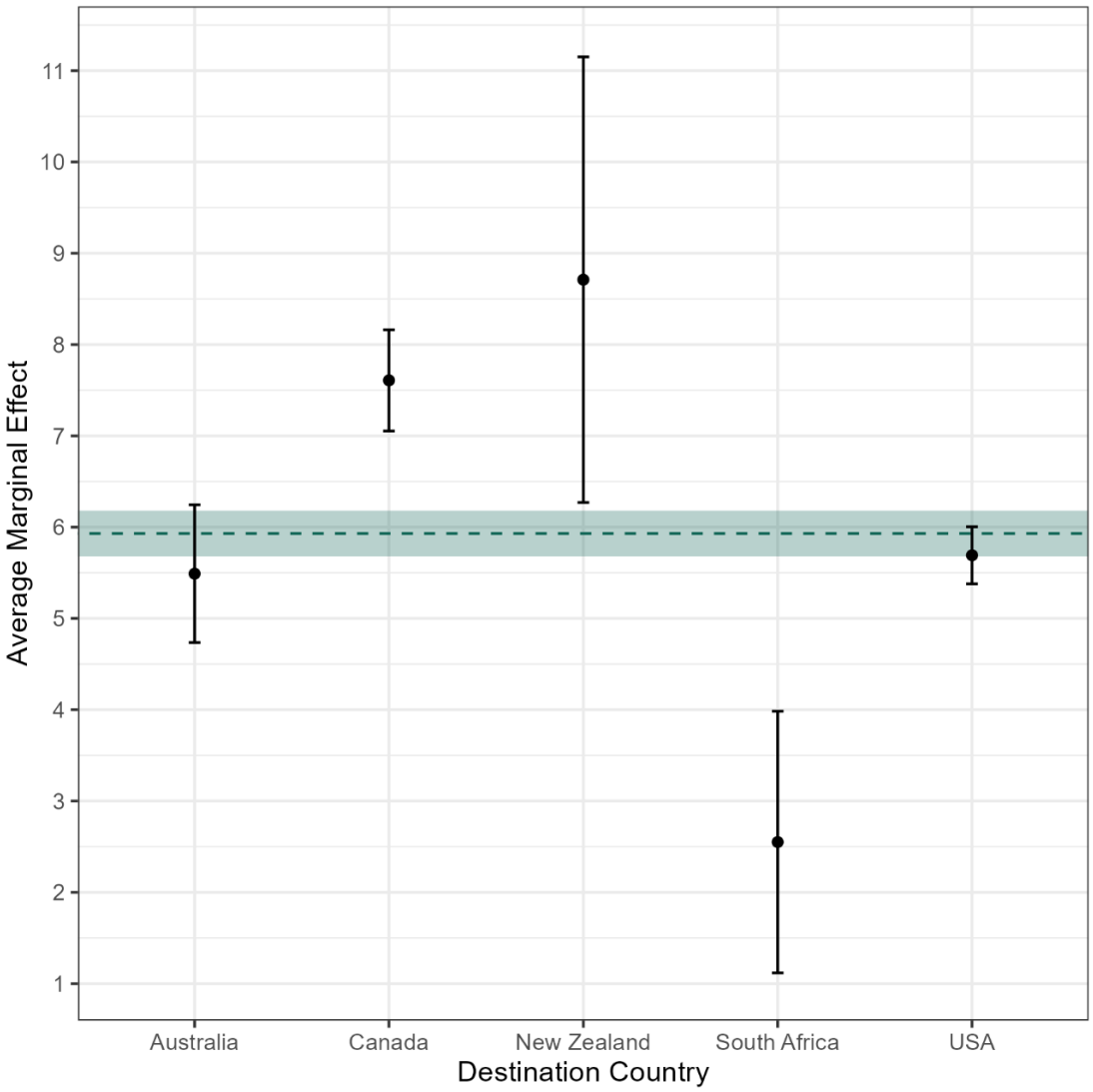

Figure B2 AME of migrating on age at death by destination country (Model B2) in black with 95% CIs. AME for Model A2 (migrant versus non-migrant) is shown by the dashed green line, with 95% CIs in green shade.

Appendix C. Results from Survival Models

Bellow we present regression estimates from the Cox mixed effects model in Table C. For the model to converge the sample is reduced to only include cohorts from 1780-1890 due to small group sizes in cohorts from 1730-1780.

Table C1 Cox mixed effects model

	Dependent variable: Age at Death	
Variable	Model A: Migrant HR (se)	Model B: Country of Death HR (se)
Migrant (*reference: Non-Migrant)*	0.752**	
	(0.041)	
Destination country *(reference: United Kingdom/Ireland)*		
Canada		0.603**
		(0.076)
South Africa		0.893
		(0.149)
Australia		0.731**
		(0.095)
New Zealand		0.731**
		(0.312)
United States of America		0.857**
		(0.054)
Birth cohort (*reference: 1780)*		
1790	1.095**	1.095**
	(0.026)	(0.026)
1800	1.121**	1.121**
	(0.025)	(0.025)
1810	1.221**	1.221**
	(0.025)	(0.025)
1820	1.302**	1.302**
	(0.025)	(0.025)
1830	1.460**	1.461**
	(0.025)	(0.025)
1840	1.509**	1.509**
	(0.025)	(0.025)
1850	1.454**	1.455**
	(0.026)	(0.025)
1860	1.330**	1.330**
	(0.026)	(0.026)

Variable	Dependent variable: Age at Death	
	Model A: Migrant HR (se)	Model B: Country of Death HR (se)
1870	1.155**	1.154**
	(0.027)	(0.027)
1880	0.983	0.983
	(0.027)	(0.027)
1890	0.784**	0.784**
	(0.028)	(0.028)
Male *(reference: female)*	1.148**	1.149**
	(0.008)	(0.008)
Number of siblings *(reference: none)*		
1	1.049**	1.049**
	(0.012)	(0.012)
2	1.066**	1.066**
	(0.014)	(0.015)
3-5	1.053**	1.054**
	(0.012)	(0.012)
6+	0.976	0.975
	(0.020)	(0.020)
Migrant X Birth cohort interaction *(reference: 1780)*		
1790	0.966	
	(0.053)	
1800	0.959	
	(0.050)	
1810	0.951	
	(0.048)	
1820	0.905*	
	(0.047)	
1830	0.811**	
	(0.047)	
1840	0.841**	
	(0.048)	
1850	0.853**	
	(0.049)	
1860	0.950	
	(0.051)	
1870	0.999	
	(0.052)	
1880	1.030	
	(0.054)	
1890	1.240**	

Variable	Dependent variable: Age at Death	
	Model A: Migrant HR (se)	Model B: Country of Death HR (se)
	(0.055)	
Destination country X Birth cohort interaction *(reference: UK/Ireland & 1780)*		
Canada:1790		0.971
		(0.099)
South Africa:1790		0.986
		(0.196)
Australia:1790		0.998
		(0.114)
New Zealand:1790		1.575
		(0.348)
United States of America:1790		0.945
		(0.070)
Canada:1800		1.012
		(0.092)
South Africa:1800		1.030
		(0.184)
Australia:1800		1.083
		(0.107)
New Zealand:1800		2.054†
		(0.107)
United States of America:1800		0.849†
		(0.067)
Canada:1810		1.021
		(0.091)
South Africa:1810		0.939
		(0.172)
Australia:1810		1.067
		(0.103)
New Zealand:1810		1.941†
		(0.321)
United States of America:1810		0.829**
		(0.063)
Canada:1820		0.951
		(0.091)
South Africa:1820		0.977
		(0.183)
Australia:1820		1.067
		(0.101)
New Zealand:1820		1.725†

Variable	Dependent variable: Age at Death	
	Model A: Migrant HR (se)	Model B: Country of Death HR (se)
		(0.319)
United States of America:1820		0.778**
		(0.062)
Canada:1830		0.887
		(0.092)
South Africa:1830		0.856
		(0.221)
Australia:1830		0.905
		(0.100)
New Zealand:1830		1.687†
		(0.100)
United States of America:1830		0.695**
		(0.061)
Canada:1840		0.983
		(0.098)
South Africa:1840		1.067
		(0.208)
Australia:1840		0.903
		(0.101)
New Zealand:1840		1.774†
		(0.317)
United States of America:1840		0.725**
		(0.061)
Canada:1850		1.088
		(0.105)
South Africa:1850		0.855
		(0.201)
Australia:1850		0.871
		(0.103)
New Zealand:1850		1.652
		(0.318)
United States of America:1850		0.761**
		(0.063)
Canada:1860		1.199†
		(0.107)
South Africa:1860		1.052
		(0.202)
Australia:1860		0.942
		(0.109)

Variable	Dependent variable: Age at Death	
	Model A: Migrant HR (se)	Model B: Country of Death HR (se)
New Zealand:1860		1.798†
		(0.319)
United States of America:1860		0.859†
		(0.066)
Canada:1870		1.109
		(0.104)
South Africa:1870		0.962
		(0.192)
Australia:1870		1.899
		(0.115)
New Zealand:1870		1.899
		(0.322)
United States of America:1870		0.922
		(0.069)
Canada:1880		1.154
		(0.099)
South Africa:1880		0.836
		(0.185)
Australia:1880		1.056
		(0.121)
New Zealand:1880		1.956*
		(0.329)
United States of America:1880		0.978
		(0.070)
Canada:1890		1.469**
		(0101)
South Africa:1890		1.260
		(0.201)
Australia:1890		1.319*
		(0.121)
New Zealand:1890		2.905**
		(0.335)
United States of America:1890		1.069
		(0.073)
Sibling random effects	Yes	Yes
Observations	84,869	84,869

Note: † p<0.1 * p<0.05 ** p<0.01