QUCEH WORKING PAPER SERIES
http://www.quceh.org.uk/working-papers

DEMOGRAPHY AND AGE HEAPING: SOLVING IRELAND'S POST-FAMINE DIGIT PREFERENCE PUZZLE

Eoin McLaughlin (University College Cork)
Christopher L. Colvin (Queen's University Belfast)
Stuart Henderson (Ulster University)

Working Paper 22-07

QUEEN'S UNIVERSITY CENTRE FOR ECONOMIC HISTORY
Queen's University Belfast
185 Stranmillis Road
Belfast BT9 5EE

May 2022

Demography and age heaping: solving Ireland's post-famine digit preference puzzle*

Eoin McLaughlin† Christopher L. Colvin‡ Stuart Henderson§

May 2022

Abstract

Age heaping in Ireland worsened in the years after the Great Irish Famine (1845–1852), even as other measures of educational attainment improved. We show how demography can account for this seemingly conflicting pattern. Specifically, we argue that a greater propensity to emigrate typified the youngest segment (23–32-year-olds) used in conventional indices of digit preference. Quantification of heaping must therefore be interpreted in light of an older underlying population which is more likely to heap. We propose how digit preference indices can adjust for such demographic change by introducing age standardisation.

Keywords: age heaping, human capital, demography.

JEL codes: N33, J10.

---

* We thank Matthias Blum, Frank Geary, Alessandro Nuvolari, Cormac Ó Gráda and Rebecca Stuart for reading an earlier version of this working paper.

† University College Cork. Email: eoin.mclaughlin@ucc.ie.

‡ Queen's University Belfast. Email: chris.colvin@qub.ac.uk.

§ Ulster University. Email: s.henderson1@ulster.ac.uk.

I. Introduction

Economic historians continue to be intrigued by patterns of digit preference in the distribution of ages. The phenomenon, now widely referred to as age heaping, first came to prominence among economic historians when Joel Mokyr drew a link between digit preference and numeracy skills in his analysis of Irish emigrants to the US.[1] Viewed as a novel solution to the lack of historical data on human capital attainment, age heaping indices have been energetically applied to a wide range of datasets to provide new illumination on human capital development across time and space.[2]

However, recent debate in the *Economic History Review* questions the efficacy of age heaping as a numeracy and human capital indicator.[3] Brian A'Hearn, Alexia Delfino and Alessandro Nuvolari argue that age heaping cannot represent numerical skills alone and posit two other explanations based on analysis of historical Italian censuses: (1) state capacity, where the accuracy of census recording reflects available state resources; and (2) culture, where digit preference mirror changes to the social admiration of youth and the elderly.[4] They conclude that changing age heaping patterns are a combined outcome of forces of modernisation rather than a narrow consequence of human capital accumulation. Joerg Baten, Giacomo Benati and Sarah Ferber reply that A'Hearn et al.'s concerns are exaggerated and reiterate the age-heaping-as-numeracy interpretation.[5] Then, a rejoinder by A'Hearn et al. argues that they are not dismissing age-heaping-as-numeracy, but rather adding nuance to the interpretation.[6]

We augment this exchange by identifying, and then subsequently solving, the "Irish heaping puzzle": a strange increase in age heaping between the 1841 and 1871 censuses (see Figures 1 and 2), despite increasing literacy, school enrolment and educational attainment. Our analysis suggests a fourth explanation for age heaping: it is partly a mechanical phenomenon,

---

[1] Mokyr, *Why Ireland starved*, which builds on Mokyr and Ó Gráda, 'Emigration and poverty'. De Bromhead, Fernihough and Ó Gráda, '175 years', now argue that Mokyr and Ó Gráda were 'almost certainly wrong'.

[2] Examples: A'Hearn, Baten and Crayen, 'Quantifying quantitative literacy', Földvári, van Leeuwen and van Leeuwen-Li, 'How did women count?'; Blum, Colvin, McAtackney and McLaughlin, 'Women of an uncertain age'; Blum and Kraus, 'Age heaping'; Cappelli and Baten, 'Numeracy development'; and, most recently, Gómez-i-Aznar, '*Ad maiorem*'.

[3] Demographers identified problems with measuring individual ages long before economic historians discovered age heaping: Shryock, Siegel, and associates, *The methods and materials*; Siegel and Swanson, *The methods and materials*. For more examples of age heaping in demography, see: Mukherjee and Mukhopadhyay, 'A study of digit preference'; Bailey and Makannah, 'Patterns of digit preference'; Fayehun, Ajayi, Onuegbu and Egerson, 'Age heaping among adults'; and Singh, Kashyap and Bango, 'Age heaping among individuals'.

[4] A'Hearn, Delfino and Nuvolari, 'Rethinking age heaping'.

[5] Baten, Benati and Ferber, 'Rethinking age heaping again'. They present technical issues with A'Hearn et al.'s methodology and cite evidence from cross-country regressions that age heaping is not correlated with measures of state capacity – although the cultural component of age heaping is left unaddressed.

[6] A'Hearn, Delfino and Nuvolari, 'Age heaping and its discontents'. Another sceptical look at the heaping-numeracy link is Beltran Tapia, Diez-Minguela, Martinez-Galarraga and Tirado-Fabregat, 'Two stories'.

driven by demographic forces. When we adjust for the underlying demographic characteristics of the population, our heaping puzzle is solved.

We show that younger populations are less inclined to report rounded ages and argue that Ireland's population was automatically more likely to heap in 1871 because it was significantly older.[7] Without taking this demographic change into consideration, scholars could incorrectly assume that Ireland had experienced a decline in numeracy, a reduction in state capacity or a change in culture. Instead, the reality is a famine-induced "premature aging" effect, precipitated by Ireland's Famine-era mortality and migration experiences.[8] Ireland's post-Famine emigrants were in large measure young female and male domestic servants and young male agricultural labourers. Only 25 per cent of the cohort aged 10–19 in 1841 remains in the 1871 census, whereas 32 per cent of the cohort aged 20–29 in 1841 remains in 1871.[9]

A reversal of age heaping in famine-affected areas is not a novel observation in the existing literature. Manzel observes a famine-related reversal for Spain; Baten, Ma, Morgan and Wang for China; and Baten, Crayen and Voth for England.[10] And for Ireland, Baten et al. first notice the pattern, while Blum, Colvin, McAtackney and McLaughlin speculated on its causes.[11] However, to our knowledge, we are the first to systematically test the hypothesis that famine-induced *demographic* change is a chief driver of heaping reversals, and the first to design a modified index to diagnose this issue.[12]

The implication of our findings is that demographic correction should be made by other researchers too; yes, demographic change may be particularly pronounced in the Irish case, but all populations experience year-on-year fluctuations due to births, aging and increased life expectancy, migration and deaths. In particular, the type of rural-urban migration typical of nineteenth-century industrialisation distorts population distributions and makes direct rural-urban comparisons fraught with difficulty. More locally to Ireland, our findings also suggest a

[7] Crayen and Baten, 'Global trends', acknowledge a higher propensity to age heap among older populations. Their simple solution is to focus only on 23–73-year-olds. However, by constructing unweighted mean Whipple indices when there were 'frequent cases of several age groups per birth decade', Crayen and Baten introduced selection bias into these measurements due to selective migration and mortality.

[8] This demographic change is described in Kennedy and Clarkson, 'Birth, death, and exile'. See also Gilleard, 'The other Victorians', where the aging population is discussed.

[9] Relatedly, Gomellini and Ó Gráda, 'Migrations', explore whether Italy's experience of selective emigration explain longitudinal age heaping patterns there.

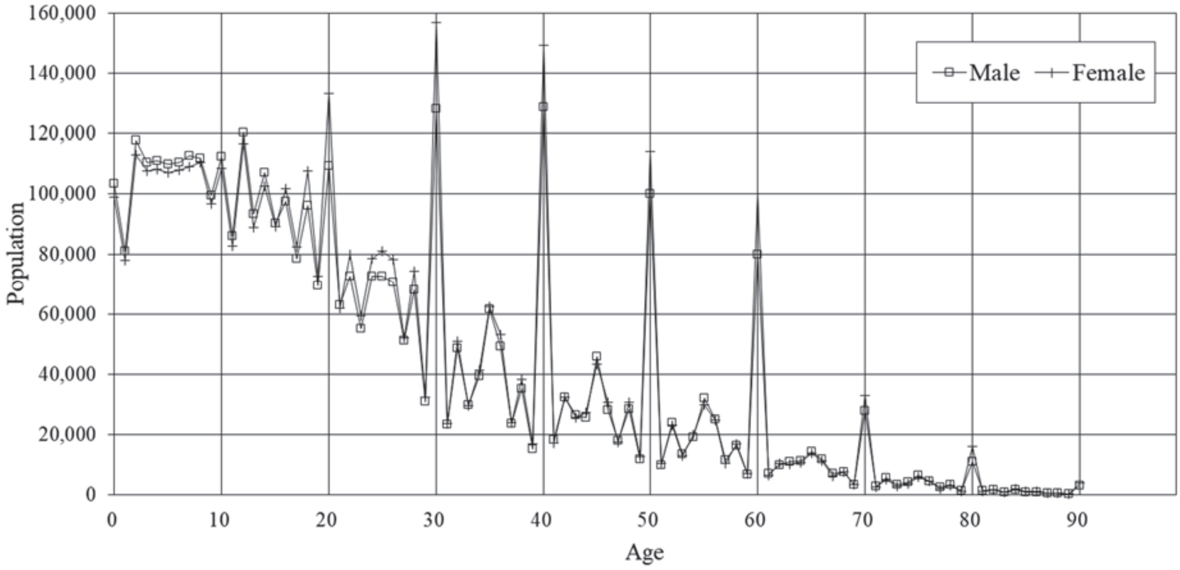
[10] Manzel, 'Essays on human capital'; Baten, Ma, Morgan and Wang, 'Evolution of living standards'; Baten, Crayen and Voth, 'Numeracy'.

[11] Baten et al., 'Evolution of living standards', p. 356; Blum et al., 'Women of an uncertain age'.

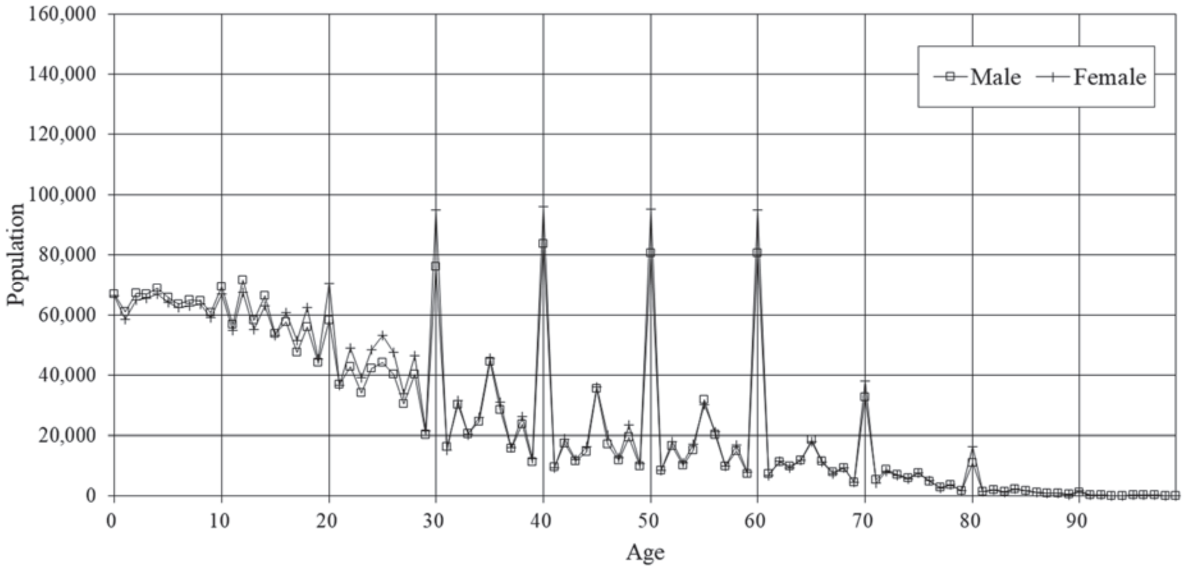
[12] Ó Gráda, 'Dublin Jewish demography', was also aware of the problem of demography in his study of Jewish migrants from Tsarist Russia in the 1911 census. Rather than constructing a single age heaping index, he simply reports separate age heaping indices for each age band.

Figure 1. *Population distribution by age*

Panel A. *1841 census*



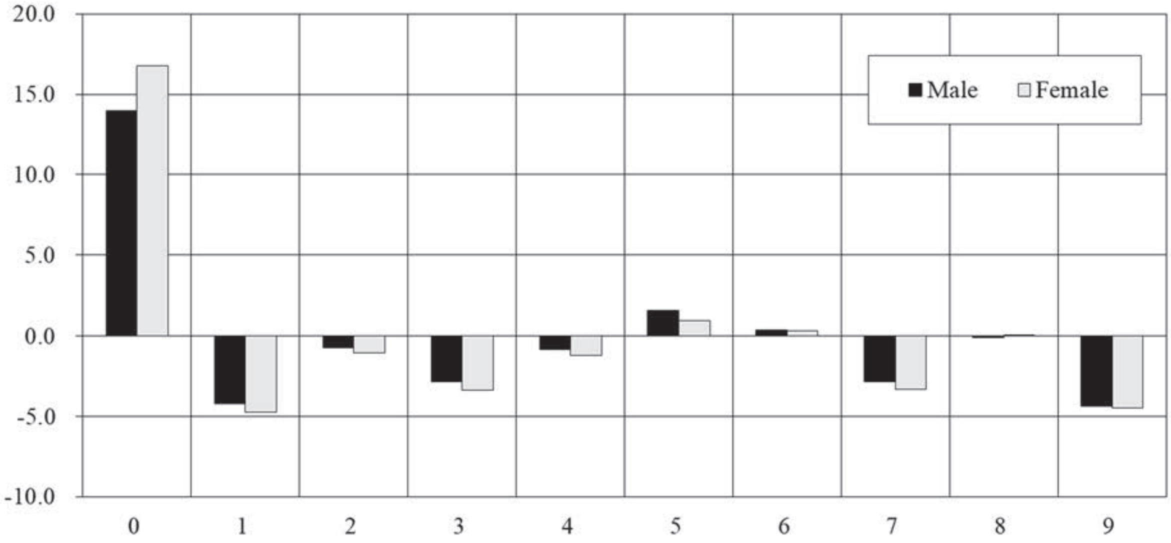
Panel B. *1871 census*



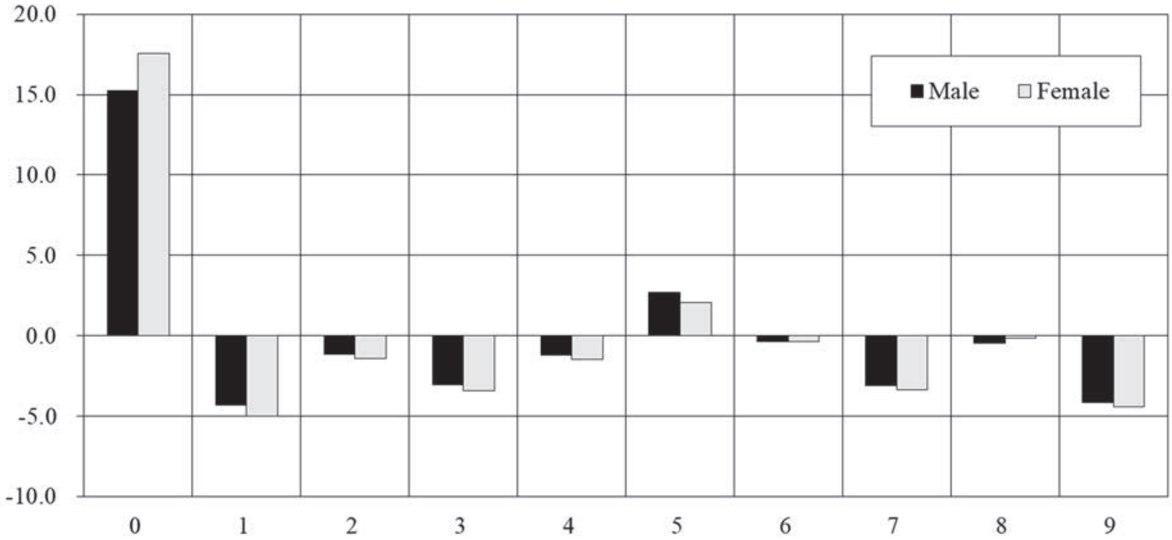
Note: The last point for 1841 is 90-years-and-over.
Source: *Report of the commissioners* (P.P. 1843, XXIV); *Census of Ireland* (P.P. 1876, LXXXI); using US Census Bureau Population Analysis System (https://www.census.gov/data/software/pas.html).

Figure 2. *Digit preference, Myers method*

Panel A. *1841 census*



Panel B. *1871 census*



Note: Robert J. Myers's blended population is a weighted sum of the number of persons reporting ages ending in each of the ten terminal digits (Myers, 'Errors and bias'; Myers, 'Accuracy of age reporting'). For 1841, only ages up to 90 are used.
Source: See Figure 1.

more explicit role for demographic aging in debates about the island's post-Famine economic performance.[17] Our inquiry proceeds by setting out the problem as we see it, and then adapting off-the-shelf ideas from the field of demography to propose diagnostic tools that others can also adopt in their work.

II. Ireland's digit preference

There are a range of alternative indices available with which to evaluate the extent of age heaping in a single metric. The most popular among economic historians has become the ABCC index, a modification of the Whipple index. The original Whipple index is a ratio of the share of people reporting an age ending in zero or five to all age statements, where the population is restricted to those aged between 23 and 62:[18]

$$WI = \left(\frac{n_{25}+n_{30}+\cdots+n_{55}+n_{60}}{1/5\times(n_{23}+n_{24}+n_{25}+\cdots+n_{62})}\right)\times 100\ if\ WI \geq 100;\ else\ WI = 100 \qquad (1)$$

where n is the sum of individuals with that specific age. As calculated, for a traditional "pyramid" shaped population distribution the WI would give greater weight to the bottom of the age distribution; because there are more younger people this means that younger age groups affect WI more than older people. A WI value of 500 indicates that all age statements end in zero or five; a value of 100 indicates no heaping.

A'Hearn, Baten and Crayen – on the suggestion of Gregory Clark – modify WI to range between zero and 100, where zero indicates that everyone reports an age terminating in 0 or 5, and 100 that there is no heaping on ages terminating in 0 and 5.[19] This modified index is argued to be interpretable as the percentage of the population that is numerate, and is given as follows:

$$ABCC = \left(1-\frac{(WI-100)}{400}\right)\times 100, if\ WI \geq 100;\ else\ ABCC = 100 \qquad (2)$$

In Table 1, we report age heaping indices for Ireland across four decennial census points, from 1841 to 1871. This suggests greater heaping in 1871 (as compared to 1841) was not an exception; heaping was more pronounced in all three census years after 1841.

[17] For recent perspectives on post-Famine economic performance, see: Begley, Geary and Stark, 'Convergence'; Henderson, 'Religion and development'; and Kenny, Lennard and O'Rourke, 'An annual index'.
[18] Or as Whipple referred to it, 'method of adjusting data troubled with these concentrations on the round numbers', Whipple, *Vital statistics*, p. 180.
[19] A'Hearn, Baten and Crayen, 'Quantifying quantitative literacy'.

Table 1. *Estimates of heaping (Whipple and ABCC), illiteracy and industrial production, 1841–1871 censuses*

Year	*Whipple (23–62)*			*ABCC (23–62)*			*Illit. % (5 years & over)*			*Production (1841=100)*		
	Male	*Female*	*Both*	*Male*	*Female*	*Both*	*Male*	*Female*	*Both*	*Output*	*Output per cap.*	*Productivity*
1841	203	213	209	74	72	73	46	59	53	100	100	100
1851	221	231	226	70	67	68	42	51	47	150	189	207
1861	217	225	221	71	69	70	35	42	39	144	204	240
1871	219	225	222	70	69	69	31	36	33	186	282	353

Note: A higher Whipple or lower ABCC implies more heaping. Illiteracy is the percentage of persons who can neither read nor write of persons 5 years and upwards. Output is measured as industrial output. Productivity is labour productivity, measured as industrial output per worker.
Source: *Report of the commissioners* (P.P. 1843, XXIV); *Census of Ireland* (P.P. 1851, XXXI); *Census of Ireland* (P.P. 1863, LVII); *Census of Ireland* (P.P. 1863, LXI); *Census of Ireland* (P.P. 1876, LXXXI). Output data from Kenny et al. 'An annual index'.

By contrast, reported literacy in Ireland increased from roughly 1-in-2 in 1841 to 7-in-10 in 1871. This runs counter to the correlation we would expect to see if heaping is an indicator of human capital.

Ireland was the first polity of the UK to receive state-funded primary education, from 1831.[20] National schools taught the three R's – reading, writing, and arithmetic – with book-keeping an especially sought-after subject.[21] It is unsurprising to see this education policy bearing fruit in terms of the reduction in illiteracy 40 years after the policy was implemented. But it is quite surprising that there was an increase in heaping over the same period, if heaping is to be interpreted narrowly as numeracy skills. It is this puzzle that we now attempt to solve.

III. Explaining the Irish puzzle

To explore the curious heaping reversal, we use tabulated age data from the 1841 and 1871 censuses of Ireland. Unlike other censuses, these have the advantage of reporting ages by county (32 counties, equivalent to NUTS-3) and therefore permit spatial comparisons across the island.[22] Figure 3 maps ABCC values by county for 1841 and 1871 and shows a striking decline in ABCC across most of the polity.

Our focus is the changing demographic composition of the population: the Irish population became older after the Famine. Among the population used in calculating the ABCC index, the main difference was a 4.35-percentage point decrease in the share of the population aged 23–32 and a 5.05-percentage point increase in the share of the population aged 53–62. This was primarily a consequence of increased emigration flows to Great Britain and North America. Irish migrants were overwhelmingly young single individuals (male and female); they tended not to migrate in family groups.[23]

We model the change in ABCC at the county level from 1841 to 1871 using a simple OLS regression. We focus on the change in the shares of the population that are used in the Whipple calculation. We also add Famine-era excess mortality and migration as controls. Famine-era excess mortality is the average annual excess death rate between 1846 and 1851.[24]

[20] Blum et al., 'Women of an uncertain age'.
[21] Coolahan, *Irish education*; Clarke, 'The teaching of bookkeeping'.
[22] Ages are only reported at the national (NUTS-1) and provincial (four provinces, NUTS-2) level in the censuses of 1851 and 1861.
[23] Ó Gráda and O'Rourke, 'Migration as disaster relief'.
[24] Mokyr, *Why Ireland starved,* Table 9.2. Mokyr provides both lower and upper bound estimates. We use the lower bound, but using the upper bound yields similar regression results.

Table 2. *Descriptive statistics*

Variable	*Mean*	*St. dev.*	*Min*	*Max*
ABCC 1841	73.42	4.02	65.75	81.76
ABCC 1871	68.55	5.40	57.28	81.19
Δ ABCC 1841–71	–4.87	2.81	–10.23	4.32
Illiteracy 1841	51.24	13.96	22.99	79.01
Illiteracy 1871	33.44	9.98	15.63	57.37
Δ Illiteracy 1841–71	–17.80	5.51	–26.09	–7.01
Δ 23–32 share in ABCC 1841–71	–4.36	2.64	–10.57	0.17
Δ 33–42 share in ABCC 1841–71	–1.91	1.35	–4.15	0.56
Δ 43–52 share in ABCC 1841–71	0.88	1.39	–1.69	3.72
Δ Female:male (23–62) 1841–71	0.94	4.99	–13.33	9.07
Famine-era excess mortality	21.73	15.74	–2.10	58.40
Famine-era migration	144.09	76.90	–177.21	259.44

Note: Illiteracy is the percentage of persons who can neither read nor write of persons 5 years and upwards.
Source: Age, illiteracy and sex: *Report of the commissioners* (P.P. 1843, XXIV); *Census of Ireland* (P.P. 1876, LXXXI). Famine-era excess mortality: Mokyr, *Why Ireland starved*, Table 9.2, lower bound. Famine-era migration: based on Ó Gráda and O'Rourke, 'Migration as disaster relief'.

Famine-era migration is based on Ó Gráda and O'Rourke's estimation of emigration between the 1841 and 1851 censuses.[25] We include the initial ABCC level in 1841 to control for level effects. Table 2 reports the descriptive statistics.

The regression results, reported in Table 3, highlight the importance of the changing composition of the population in explaining the change in ABCC over time. In specification 1, changing age structure and ABCC in 1841 explain 34 per cent of the variation in the dependent variable. With the inclusion of Famine-era migration and excess mortality in specification 3, the magnitude of the age structure effects is smaller. Notably, Famine-era migration is more statistically important than Famine-era mortality. This is to be expected; it was the oldest and the youngest that were most prone to perish during the Famine.[26] However, Famine-era migration was the main driver of the compositional change of the population – specifically, those age groups included in the most widely-used age heaping indices.

[25] Ó Gráda and O'Rourke, 'Migration as disaster relief'.
[26] Mokyr and Ó Gráda, 'What do people die of'.

Table 3. *OLS regressions of the change in ABCC and illiteracy between 1841 and 1871*

Variable	*Δ ABCC 1841–71*			*Δ Illiteracy 1841–71*		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Δ 23–32 share in ABCC 1841–71	1.390***	1.431***	0.597**	0.791	0.679	0.327
	(0.479)	(0.484)	(0.253)	(0.670)	(0.599)	(0.478)
Δ 33–42 share in ABCC 1841–71	1.040*	1.144**	0.644**	1.274**	0.942	0.819
	(0.521)	(0.534)	(0.287)	(0.573)	(0.568)	(0.505)
Δ 43–52 share in ABCC 1841–71	2.527**	2.815**	1.252**	1.401	0.558	–0.211
	(1.029)	(1.092)	(0.477)	(1.299)	(1.201)	(1.033)
Δ Female:male (23–62) 1841–71		–0.101	–0.138*		0.303***	0.286***
		(0.093)	(0.072)		(0.104)	(0.089)
Famine-era excess mortality			-0.045			–0.085*
			(0.034)			(0.049)
Famine-era migration			–0.021***			–0.012
			(0.004)			(0.008)
ABCC 1841	0.023	0.034	–0.100			
	(0.134)	(0.138)	(0.136)			
Illiteracy 1841				–0.263***	–0.256***	–0.162**
				(0.062)	(0.055)	(0.065)
Constant	–0.751	–1.331	9.281	0.323	–0.713	–2.949
	(10.770)	(10.904)	(10.647)	(1.935)	(2.153)	(2.374)
Observations	32	32	32	32	32	32
R-squared	0.338	0.362	0.715	0.710	0.766	0.834

Note: Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Illiteracy is the percentage of persons who can neither read nor write of persons 5 years and upwards.
Source: See Table 2.

Demography also matters when looking at illiteracy in specifications 4 to 6. Younger generations increasingly benefitted from the rise of the publicly funded National System of Education, and this invariably fuelled declining illiteracy over time. But of crucial importance was the increased literacy of the female population; the educational attainment of women in particular was emphasised by Fitzpatrick and Blum et al.[27] Here also Famine-era migration and excess mortality account for decreased illiteracy.

[27] Fitzpatrick, ‘‘A share of the honeycomb’’; Blum et al., ‘Women of an uncertain age’.

IV. Introducing the age-standardised age heaping index

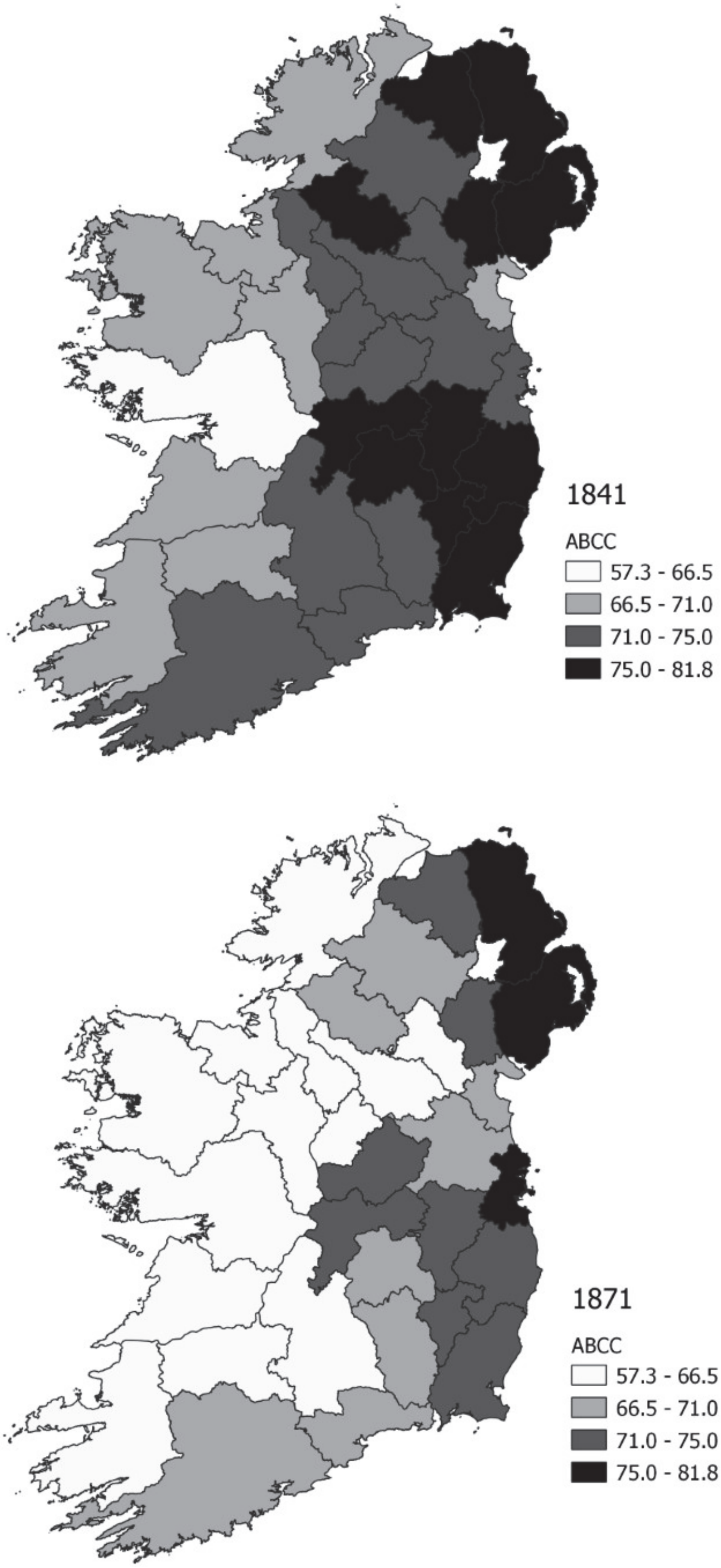
Baten et al. argue that age heaping is best analysed by birth cohort rather than from a population average. However, comparing different birth cohorts derived from one census source introduces selection bias.[28] For example, using the 1871 census and comparing those born in the 1820s with those born in the 1810s introduces bias from selective mortality and selective migration. The resultant demographic change can be significant: as noted already in the introduction, only 25 per cent of the cohort aged 10–19 in 1841 remains in the 1871 census, whereas 32 per cent of the cohort aged 20–29 in 1841 remains in the 1871 census. This is further highlighted in Figure 4, which plots the ABCC values of the cohorts born in the 1810s from both the 1841 and 1871 censuses. The decline in male and female ABCC values for the cohort born in the 1810s is driven by selective mortality and migration.

An alternative is to compare the same age group, not cohort, across census years. This allows us to compare youth with youth, middle-aged with middle-aged, and elderly with elderly across the different census years. Figure 5 illustrates the change in ABCC values between the two census points by sex and age bin. There is very little change in the 23–32 age group for men and women, while the largest change is in the 43–52 age group. Women see negligible decline in the 53–62 age group, whereas for men there is still a substantial decline. However, the main change between both years is the relative weight of each group in wider population.

Our technical solution to the demographic problem inherent in extant age heaping indices is to adjust the conventional Whipple index by standardising it by age. Our modified Whipple index functions as a demographic diagnostic tool. Adjusting for demographic composition is commonly carried out when comparing mortality rates. A classic example is the comparison of mortality rates in Florida, which is famously a destination for retirees, and Alaska, which has a younger population. Without adjusting for age composition, the Florida-Alaska comparison gives a misleading impression of mortality.[29] The same principle applies to age heaping.

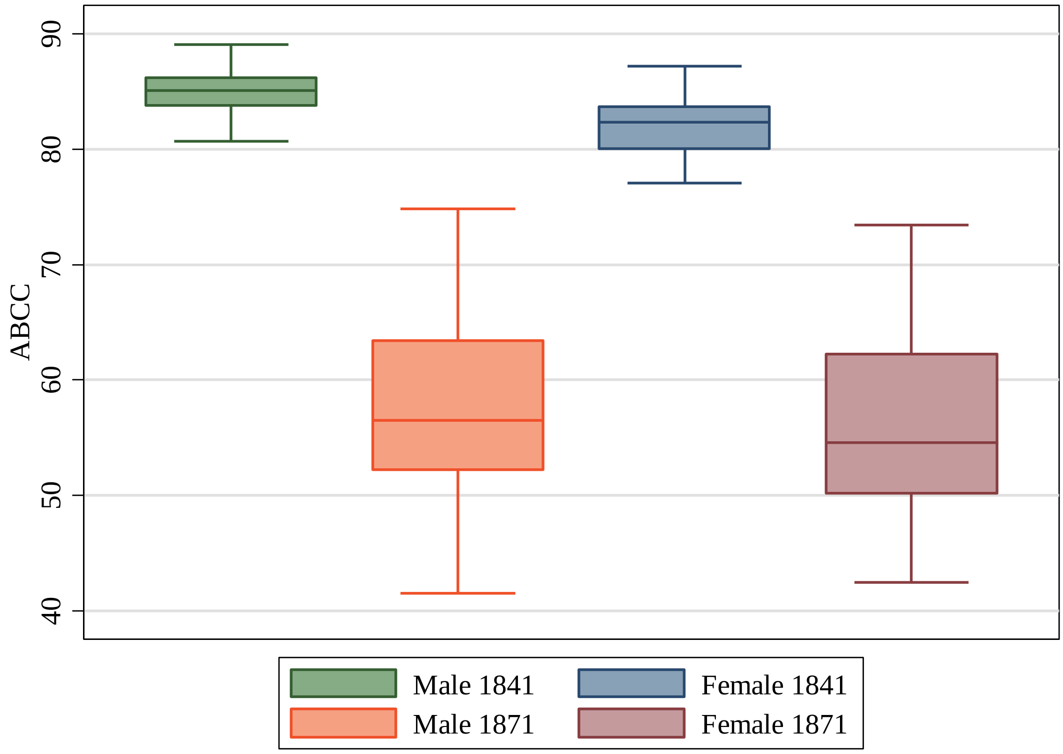
[28] Bodenhorn, Guinnane and Mroz, 'Sample-selection biases', have a similar argument in relation to anthropometrics.
[29] See Colvin and McLaughlin, 'Death, demography and the denominator', for an application of this in measuring the demographic impact of the 1918 influenza pandemic.

Figure 3. *ABCC index by county in the 1841 and 1871 censuses*



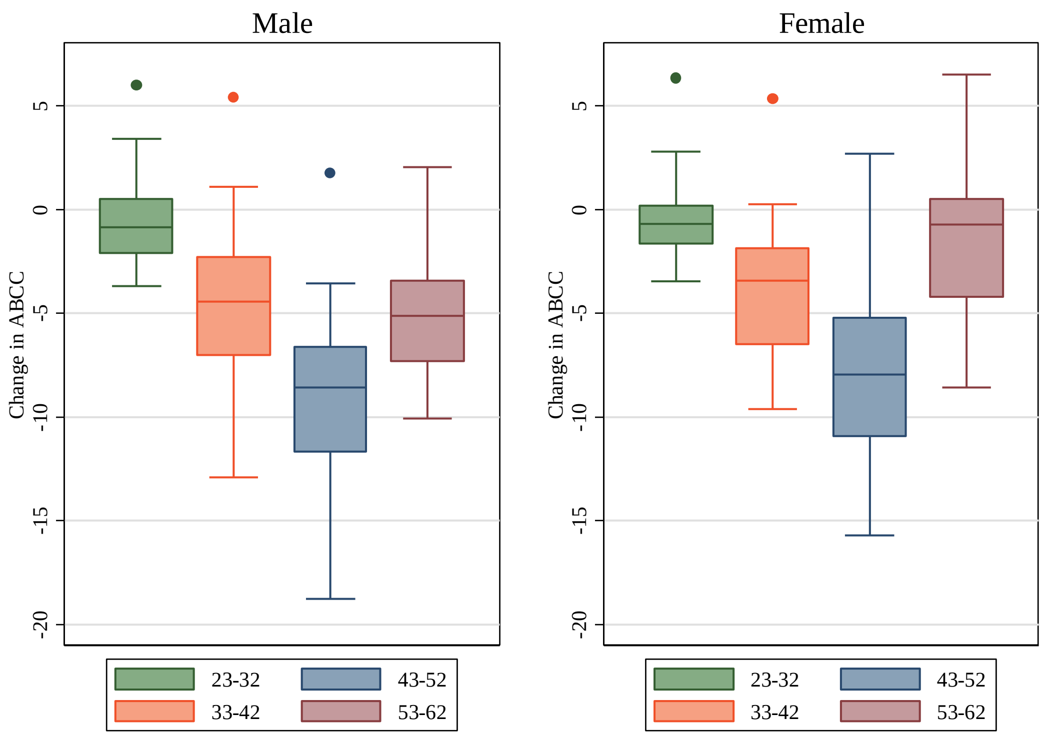
Note: Bins are nested means derived from the mean of all for both years, the mean between the minimum and the mean and the mean between the mean and the maximum for both years.
Source: See Table 2.

Figure 4. *ABCC values of cohorts born in the 1810s reported in the 1841 and 1871 censuses*



Note: Boxplot demonstrating the locality, spread and skewness of cohorts through their quartiles. We use the 23–32 cohort from 1841 and the 53–62 cohort from 1871.
Source: See Table 2.

Figure 5. *Change in ABCC values by age-category between the 1841 and 1871 censuses*



Note: Outliers plotted as individual points.
Source: See Table 2.

Whipple made a distinction between '*secessive*' and '*accessive*' populations, the former being a population which has excess emigration and the latter one with excess immigration.[30] Yet the Whipple index itself compares the relative frequency of age statements terminating in 0 and 5 for the population aged between 23 and 62 without allowance for the age composition of the population. This is because Whipple's original purpose was to adjust for 'errors in age' rather than for making comparisons of populations across time and space. Simply put, this index is not being applied in the way that it was originally intended to be.

We can account for age composition differences by constructing a weighted average Whipple index, using standardised population weights to make our adjustment.[31] As the magnitude of error is in individual ages, the weights are derived from the underlying data source by grouping the ages:

$$AgeWI = \sum\left(\frac{n_{25}+n_{30}}{1/5\times(n_{23}\ldots+n_{32})}\right)(m_{23-32})+\ldots+\left(\frac{n_{55}+n_{60}}{1/5\times(n_{53}\ldots+n_{62})}\right)(m_{53-62}) \quad (3)$$

where *m* are age-standardised weights for each age group.[32] Standardised estimates then adjust for the differences in population structure and enable us to see whether the changing composition of the population explains the differences in heaping observed.[33] Essentially, our age-standardised Whipple index allows us to compare like-with-like.

The age-standardised Whipple index can be incorporated into a comparable framework as the age-standardised ABCC (*AgeABCC*) by replacing *WI* with *AgeWI* in Equation 2. Our weights are derived from the 1841 census (reported in Table 4), enabling us to directly compare 1841's younger population with 1871's older population. This is akin to the use of a base year in the calculation of real GDP. The effect of this adjustment is best illustrated by comparing the unadjusted and adjusted ABCC values for 1871, shown in Figure 6. The adjustment (i.e., standardisation) using the 1841 weights increases the mean ABCC values for both men and women.

Our adjustment leads to a significant alteration of the 'change in ABCC between 1841 and 1871' variable used previously as the dependent variable in our regression analysis. Figure

---

[30] Whipple focuses on the population distribution of those between 15 and 50. Classification of secessive and accessive is based on whether this population is below or above 50%. The examples given are Sweden (50%) versus Massachusetts in 1910 (57%); Whipple, *Vital statistics*, p. 190.

[31] As Whipple noted: 'for purposes of computation and comparison it is often convenient to have some standard of age distribution which can be used as a basis of reference' (Whipple, *Vital statistics*, p. 191).

[32] These constitute the following age bins: 23–32, 33–42, 43–52, and 53–62.

[33] Demographers refer to adjustments for age and digit preference as ways to reduce it in the underlying data source (e.g., discussion in Siegel and Swanson, *The methods and materials*).

Table 4. *Population weights and sex ratios by age bin*

	Age bin			
	23–32	*33–42*	*43–52*	*53–62*
1841 census weights	0.39	0.27	0.20	0.14
1871 census weights	0.35	0.25	0.21	0.19
1841 female population share	0.53	0.52	0.51	0.52
1871 female population share	0.54	0.52	0.53	0.52

Source: See Table 2.

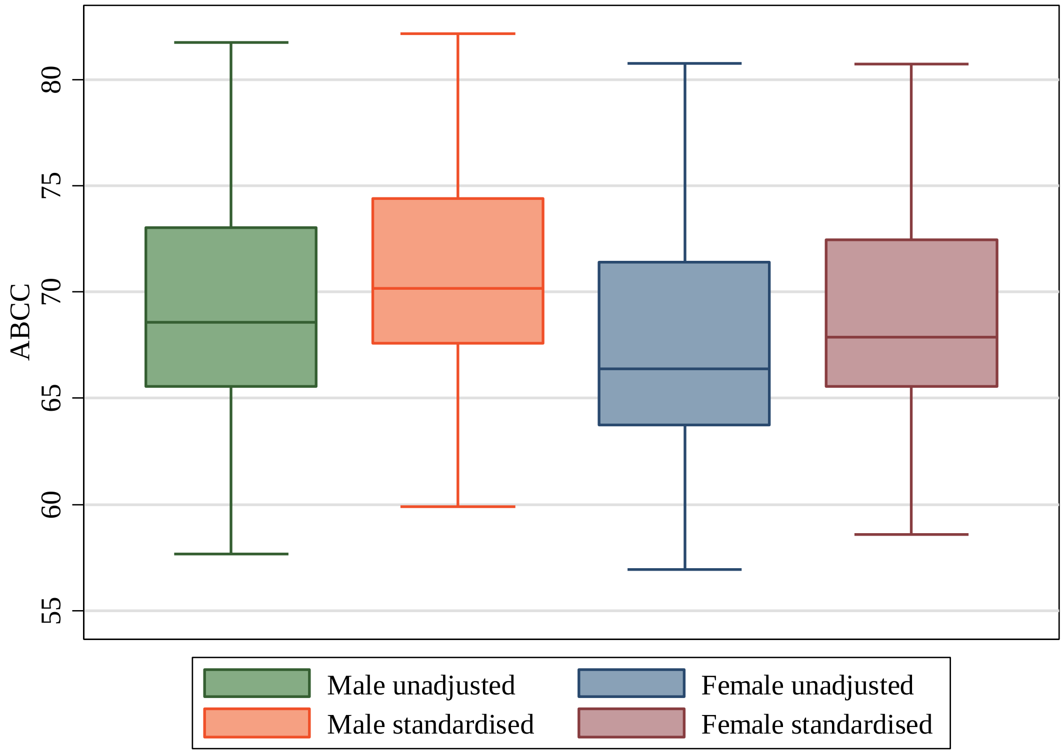
7 shows the percentage difference of the change (1841–1871) between unadjusted and standardised figures in a boxplot. After standardisation, the change in ABCC values was 30 per cent lower for males and 33 per cent lower for females (relative to the change when unadjusted figures were used). In extreme cases, such as in County Westmeath, the difference between 1841 and 1871 was 66 per cent lower for males and 177 per cent lower for females; the latter reflecting a decrease in heaping when standardised figures are used (as compared to an increase in heaping when unadjusted figures are used). County Dublin, which is unusual due to the reduction in heaping it experienced from 1841 to 1871, has larger changes in ABCC when standardised figures are used. This underscores the sizeable implications of standardisation. As a final exercise, in Appendix Tables A1 and A2, we report the unadjusted and age-standardised ABCC indices by sex and county.

V. Conclusion

We have solved the Irish digit preference puzzle. Heaping increased because there was a decreasing share of those less likely to heap in the total population, a consequence of emigration patterns unleashed by the Famine. Our findings have implications for the relevance and interpretation of age heaping as a human capital indicator. As we illustrate here, age-standardised estimates of ABCC are required to identify whether changing demography is driving changes in age heaping, and then make more meaningful comparisons of heaping across time and space.[34] Populations change over time. This need not be as dramatic as the case of Ireland, but could be due to industrialisation and the push-pull factors influencing rural-urban migration flows.

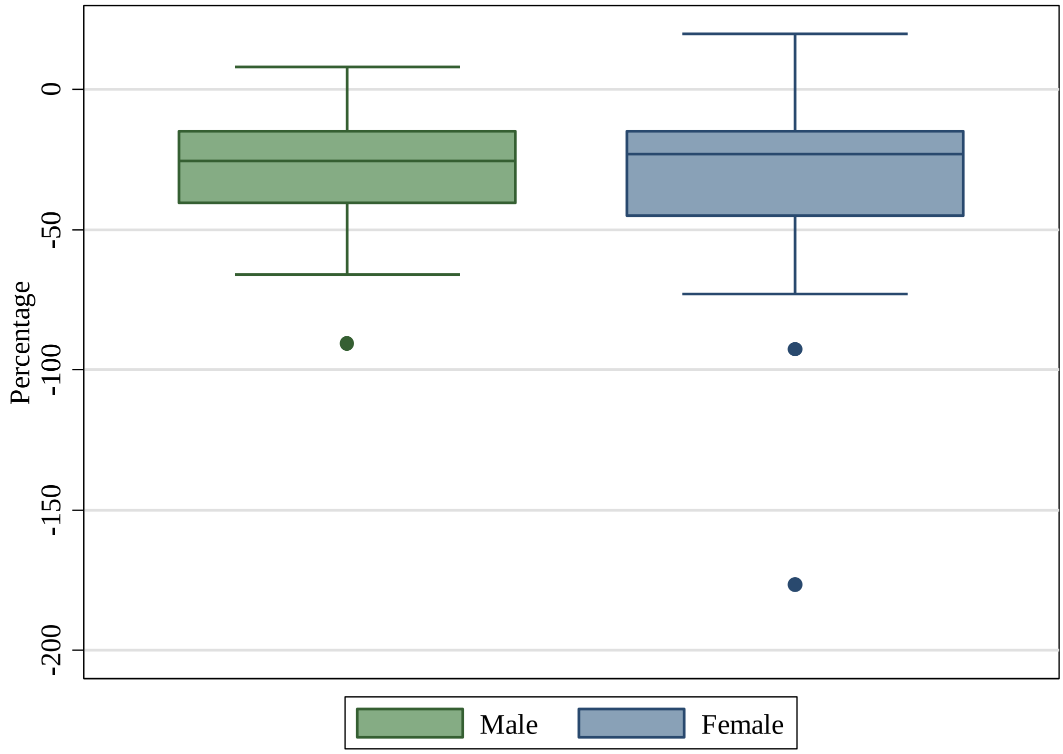
[34] For example, in Crayen and Baten, 'Global trends'.

Figure 6. *Unadjusted and age standardised ABCCs in 1871*



Source: See Table 2.

Figure 7. *Percentage difference of ΔABCC (1841–1871) between unadjusted and standardised*



Source: See Table 2.

In more modern settings, measures of digit preference are still recorded to control for data quality. For example, Pardeshi highlights how recorded ages in survey data in India are very poor quality given the rounding of ages and suggests more innovative ways to collect such data.[35] Economic historians have now firmly adopted age heaping as a proxy for numeracy, but, as we demonstrate, any such interpretation is fraught with difficulty. Instead, we suggest heaping should initially be used more narrowly as a check on the quality of historical age statements. We suggest the economic history profession now (temporarily) put aside the heaping-as-numeracy interpretation alongside all its various rivals. What drives heaping changes over time should in the first instance be explored in the context of wider population dynamics, such as demographic and epidemiological transitions.

Footnote references

A'Hearn, B., Baten, J., and Crayen, D., 'Quantifying quantitative literacy: age heaping and the history of human capital', *Journal of Economic History*, 69 (2009), pp. 783–808.

A'Hearn, B., Delfino, A., and Nuvolari, A., 'Rethinking age heaping: a cautionary tale from nineteenth-century Italy', *Economic History Review*, 75 (2021), pp. 111–137.

A'Hearn, B., Delfino, A., and Nuvolari, A., 'Age heaping and its discontents: a response to Baten, Benati, and Ferber', *Economic History Review* (2022), URL: https://doi.org/10.1111/ehr.13140.

Anbinder, T., Connor, D., Ó Gráda, C., and Wegge, S., 'The problem of false positives in automated census linking: evidence from nineteenth-century New York's Irish immigrants', UCD Centre for Economic Research Working Paper Series, Paper No. WP21/14 (June 2021), URL: https://www.ucd.ie/economics/t4media/WP21_14.pdf.

Bachi, R., 'The tendency to round off age returns: measurement and correction', *Bulletin of the International Statistical Institute*, 33 (1951), pp. 195–221.

Bailey, M., and Mukannah, T. J., 'Patterns of digit preference and avoidance in the age statistics of some recent African censuses', *Journal of Official Statistics*, 9 (1993), pp. 705–715.

Baten, J., Ma, D., Morgan, S. and Wang, Q., 'Evolution of living standards and human capital in China in the 18–20th centuries: evidences from real wages, age-heaping, and anthropometrics', *Explorations in Economic History*, 47 (2010), pp. 347–359

Baten, J., Benati, G., and Ferber, S., 'Rethinking age heaping again for understanding its possibilities and limitations', *Economic History Review* (2022), URL: https://doi.org/10.1111/ehr.13139.

Baten, J., Crayen, D., and Voth, H.-J., 'Numeracy and the impact of high food prices in industrializing Britain, 1780-1850', *The Review of Economics and Statistics*, 96 (2014), pp. 418–430.

Begley, J., Geary, F., and Stark, T., 'Convergence in the pre-1911 Atlantic economy: what really happened to wages in Ireland between 1881 and 1911?', *Cambridge Journal of Economics*, 40 (2016), pp. 43–67.

Beltrán Tapia, F., Diez-Minguela, A., Martinez-Galarraga, J., and Tirado-Fabregat, D., 'Two stories, one fate: age-heaping and literacy in Spain, 1877–1930', *Revista de Historia Economica* (2021), URL: https://doi.org/10.1017/S0212610921000033.

Bodenhorn, H., Guinnane, T. W., and Mroz, T. A., 'Sample-selection biases and the "industrialization puzzle"', *Journal of Economic History*, 77 (2017), pp. 171–207.

Blum, M., Colvin, C. L., McAtackney, L., and McLaughlin, E., 'Women of an uncertain age: quantifying human capital accumulation in rural Ireland in the nineteenth century', *Economic History Review*, 70 (2017), pp. 187–223.

Blum, M., and Krauss, K.-P., 'Age heaping and numeracy: looking behind the curtain', *Economic History Review*, 71(2018), pp. 464–79.

Budd, J. W., and Guinnane, T., 'Intentional age-misreporting, age-heaping, and the 1908 Old Age Pensions Act in Ireland', *Population Studies*, 45 (1991), pp. 497–518.

Cappelli, G., and Baten, J., 'Numeracy development in Africa: new evidence from a long-term perspective (1730–1970)', *Journal of Development Economics*, 150 (2021), article no. 102630.

[35] Pardeshi, 'Age heaping and accuracy'.

Clarke, P., ‘The teaching of bookkeeping in nineteenth-century Ireland’, *Accounting, Business and Financial History*, 18 (2008), pp. 21–33.
Coolahan, J., *Irish education: its history and structure 1800–1980* (Dublin, 1981).
Crayen, D., and Baten, J., ‘Global trends in numeracy 1820–1949 and its implications for long-term growth’, *Explorations in Economic History*, 47 (2010), pp. 82–99.
de Bromhead, A., Fernihough, A., and Ó Gráda, C., ‘175 years of mismeasuring in Ireland? Some traps set by historical data’, Statistical and Social Inquiry Society of Ireland, Dublin (May 2022).
Fayehun, O., Ajayi, A. I., Onuegbu, C., and Egerson, D., ‘Age heaping among adults in Nigeria: evidence from the Nigeria Demographic and Health Surveys 2003-2013’, *Journal of Biosocial Science*, 52 (2020), pp. 132–139.
Fitzpatrick, D., ‘‘A share of the honeycomb’: Education, emigration and Irishwomen’, *Continuity and Change*, 1 (1986), pp. 217–234.
Földvári, P., van Leeuwen, B., and van Leeuwen-Li, J., ‘How did women count? A note on gender-specific age heaping differences in the sixteenth to nineteenth centuries’, *Economic History Review*, 65 (2012), pp. 304–313.
Gilleard, C., ‘The other Victorians: age, sickness and poverty in 19th-century Ireland’, *Ageing & Society*, 36 (2016), pp. 1157–1184.
Gomellini, M., and Ó Gráda, C., ‘Migrations’, in G. Toniolo, ed., *The Oxford handbook of the Italian economy since Unification* (Oxford, 2013), pp. 271–302.
Gómez-i-Aznar, È., ‘*Ad maiorem Dei gloriam*: numeracy levels in the Guarani Jesuit missions’, *Economic History Review* (2022), URL: https://doi.org/10.1111/ehr.13169.
Henderson, S., ‘Religion and development in post-Famine Ireland’, *Economic History Review*, 72 (2019), pp. 1251–1285.
Hippe, R., and Baten, J. ‘Regional inequality in human capital formation in Europe, 1790–1880’, *Scandinavian Economic History Review*, 60 (2012), pp. 254–289.
Kennedy, L., and Clarkson, A., ‘Birth, death, and exile: Irish population history, 1700–1921’, in: Graham, B. J., and Proudfoot, L. J. (eds.), *An historical geography of Ireland* (London, 1993).
Kenny, S., Lennard, J., and O’Rourke, K. H., ‘An annual index of Irish industrial production, 1840–1913’, London School of Economics and Political Science, Economic History Working Series, Paper No. 312 (July 2021), URL: http://eprints.lse.ac.uk/107427/2/WP312_002_.pdf.
Manzel, K., ‘Essays on human capital development in Latin America and Spain’, unpublished PhD thesis (University of Tubingen, 2009), URL: https://publikationen.uni-tuebingen.de/xmlui/bitstream/handle/10900/47891/pdf/Version30_01_2012.pdf.
Mokyr, J., *Why Ireland starved: a quantitative and analytical history of the Irish economy, 1800–1850* (London, 1983).
Mokyr, J., and Ó Gráda, C., ‘Emigration and poverty in prefamine Ireland’, *Explorations in Economic History*, 19 (1982), pp. 360–84.
Mokyr, J., and Ó Gráda, C., ‘What do people die of during famines: the Great Irish Famine in comparative perspective’, *European Review of Economic History*, 6 (2002), pp. 339–363.
Mukherjee, B. N., and Mukhopadhyay, B. K., ‘A study of digit preference and quality of age data in Turkish censuses’, *Genus*, 44 (1988), pp. 201–227.
Myers, R. J., ‘Errors and bias in the reporting of ages in census data’, *Transactions of the Actuarial Society of America*, XLI (1940), reprinted in: *Handbook of statistical methods for demographers* (US Bureau of the Census, 1960).
Myers, R. J., ‘Accuracy of age reporting in the 1950 United States census’, *Journal of the American Statistical Association*, 49 (1954), pp. 826–831.
Ó Gráda, C., and O’Rourke, K., ‘Migration as disaster relief: lessons from the Great Irish Famine’, *European Review of Economic History*, 1 (1997), pp. 3–25.
Ó Gráda, C., ‘Dublin Jewish demography a century ago’, *The Economic and Social Review*, 37 (2006), pp. 123–147.
Pardeshi, G. S., ‘Age heaping and accuracy of age data collected during a community survey in the Yavatmal District, Maharashtra’, *Indian Journal of Community Medicine*, 35 (2010), pp. 391–395.
Shrycock, H. S., Siegel, J. S., and associates, *The methods and materials of demography* (Amsterdam, 1976).
Siegel, J. S., and Swanson, D. A., *The methods and materials of demography* (Amsterdam, 2004).
Singh, M., Kashyap, G. C., and Bango, M., ‘Age heaping among individuals in selected South Asian countries: evidence from Demographic and Health Survey’, *Journal of Biosocial Science* (2021), URL: https://doi.org/10.1017/S0021932021000249.
Whipple, G. C., *Vital statistics: an introduction to the science of demography* (2nd Edition, New York, 1923).

Official publications

Census of Ireland, 1851, pt. VI, 'General report' (P.P. 1851, XXXI).
Census of Ireland, 1861, pt. II, 'Report and tables on ages and education, volume II' (P.P. 1863, LVII).
Census of Ireland, 1861, pt. V, 'General report, appendix, county tables, summary, index' (P.P. 1863, LXI).
Census of Ireland, 1871, pt. III, 'General report, with illustrative maps and diagrams, summary tables, and appendix' (P.P. 1876, LXXXI).
Report of the commissioners appointed to take the Census of Ireland for the year 1841 (P.P. 1843, XXIV).
Thirteenth census of the United States taken in the year 1910, volume 1 Population (United States Census Office, 1913).
Twelfth census of the United States taken in the year 1900. Population, part 1 (United States Census Office, 1901).

Table A1. *Unadjusted and standardised ABCC indices for men, by county, 1841 and 1871*

Province	*Unadjusted indices*				*Standardised indices*				*Difference unadjusted and standardised*	
County	*ABCC*		*Difference with national average*		*Age-standardised ABCC*		*Difference with national average*			
	1841	*1871*	*1841*	*1871*	*1841*	*1871*	*1841*	*1871*	*1841*	*1871*
	/100		*%*		*/100*		*%*		*p. p.*	
Leinster										
Carlow	78.85	72.36	5.50	4.29	78.78	74.00	5.21	4.32	–0.07	1.63
Dublin	76.95	81.06	2.95	16.82	76.80	81.24	2.57	14.52	–0.15	0.18
Kildare	79.14	76.46	5.89	10.19	79.14	76.59	5.69	7.98	0.00	0.14
Kilkenny	75.89	72.25	1.53	4.13	75.62	73.96	0.99	4.27	–0.27	1.71
King's*	76.82	73.02	2.78	5.23	76.96	74.38	2.79	4.85	0.14	1.36
Longford	74.34	67.05	–0.54	–3.37	74.95	68.53	0.10	–3.39	0.61	1.48
Louth	70.56	68.81	–5.60	–0.83	71.19	70.29	–4.92	–0.90	0.63	1.48
Meath	73.53	69.09	–1.62	–0.42	73.79	70.45	–1.45	–0.68	0.26	1.36
Queen's**	76.78	69.00	2.72	–0.56	76.71	70.83	2.45	–0.15	–0.07	1.83
Westmeath	74.59	73.02	–0.20	5.24	74.92	74.38	0.06	4.86	0.33	1.36
Wexford	82.58	74.70	10.48	7.65	82.79	76.10	10.57	7.28	0.21	1.40
Wicklow	80.73	73.28	8.01	5.61	80.72	74.59	7.81	5.15	–0.01	1.31
Munster										
Clare	72.94	64.83	–2.41	–6.57	72.40	67.20	–3.31	–5.26	–0.54	2.37
Cork	73.01	68.35	–2.32	–1.49	72.75	70.05	–2.85	–1.24	–0.27	1.70
Kerry	70.66	64.97	–5.47	–6.37	70.48	66.78	–5.87	–5.86	–0.17	1.81
Limerick	72.93	67.86	–2.43	–2.19	72.63	69.55	–3.00	–1.95	–0.30	1.69
Tipperary	73.16	68.21	–2.12	–1.70	72.65	69.88	–2.97	–1.49	–0.51	1.67
Waterford	72.73	67.35	–2.70	–2.93	72.13	69.52	–3.67	–1.99	–0.60	2.17
Ulster										
Antrim	82.11	81.73	9.86	17.79	82.19	82.16	9.77	15.82	0.08	0.43
Armagh	75.86	73.06	1.50	5.30	76.25	74.44	1.83	4.95	0.39	1.38
Cavan	73.05	64.68	–2.27	–6.78	73.45	66.14	–1.90	–6.76	0.40	1.46
Donegal	69.31	63.30	–7.27	–8.77	70.07	65.29	–6.42	–7.96	0.76	1.99
Down	80.24	76.99	7.35	10.96	80.55	78.35	7.58	10.46	0.31	1.37
Fermanagh	77.60	69.03	3.82	–0.51	77.99	70.45	4.16	–0.68	0.39	1.41
Londonderry	76.43	71.54	2.25	3.11	77.07	72.66	2.93	2.43	0.64	1.12
Monaghan	71.94	65.66	–3.75	–5.38	72.56	67.49	–3.10	–4.85	0.62	1.84
Tyrone	74.42	67.70	–0.43	–2.43	75.19	69.08	0.42	–2.61	0.77	1.38
Connacht										
Galway	67.19	62.00	–10.11	–10.64	67.21	64.57	–10.24	–8.97	0.02	2.57
Leitrim	76.64	66.08	2.54	–4.77	76.76	67.74	2.52	–4.50	0.12	1.67
Mayo	68.49	57.66	–8.37	–16.90	68.50	59.90	–8.52	–15.56	0.01	2.24
Roscommon	70.97	65.49	–5.05	–5.62	71.18	67.61	–4.93	–4.68	0.21	2.13
Sligo	71.37	63.76	–4.51	–8.11	71.65	65.66	–4.31	–7.43	0.28	1.90
National										
Mean	74.74	69.39			74.88	70.93			0.13	1.55
Weighted***	74.21	70.31			74.27	71.77				

Note: * King's County is now known as County Offaly. ** Queen's County is now known as County Laois. Underlined constitute the six counties of Northern Ireland from 1921. *** Weighted by county population size (23–62-year-olds).
Source: See Table 2.

Table A2. *Unadjusted and standardised ABCC indices for women, by county, 1841 and 1871*

Province	*Unadjusted indices*				*Standardised indices*				*Difference unadjusted and standardised*	
County	*ABCC*		*Difference with national average*		*Age-standardised ABCC*		*Difference with national average*			
	1841	*1871*	*1841*	*1871*	*1841*	*1871*	*1841*	*1871*	*1841*	*1871*
	/100		*%*		*/100*		*%*		*p. p.*	
Leinster										
Carlow	75.80	70.37	5.01	3.85	75.71	71.63	4.93	3.72	–0.09	1.26
Dublin	72.32	76.80	0.18	13.34	71.75	77.12	–0.57	11.67	–0.57	0.32
Kildare	75.59	72.82	4.71	7.47	75.64	73.62	4.83	6.61	0.06	0.80
Kilkenny	72.14	69.40	–0.06	2.41	72.04	71.30	–0.17	3.24	–0.11	1.90
King's*	73.33	70.59	1.59	4.18	73.36	72.06	1.67	4.35	0.03	1.47
Longford	71.91	65.06	–0.38	–3.98	72.52	65.97	0.50	–4.48	0.61	0.91
Louth	67.34	66.22	–6.70	–2.27	67.26	67.17	–6.79	–2.73	–0.09	0.95
Meath	70.94	66.52	–1.73	–1.82	70.97	67.60	–1.64	–2.11	0.03	1.08
Queen's**	73.36	67.62	1.63	–0.21	73.21	69.06	1.46	0.00	–0.16	1.44
Westmeath	70.80	70.10	–1.92	3.46	71.09	71.61	–1.48	3.70	0.29	1.51
Wexford	80.26	74.92	11.18	10.57	80.49	76.05	11.55	10.11	0.23	1.12
Wicklow	79.86	74.43	10.64	9.84	79.78	75.27	10.56	8.99	–0.08	0.84
Munster										
Clare	68.73	59.72	–4.78	–11.87	68.22	63.06	–5.46	–8.70	–0.52	3.34
Cork	69.73	65.11	–3.40	–3.91	69.66	67.06	–3.46	–2.90	–0.07	1.95
Kerry	67.28	61.73	–6.79	–8.90	67.12	63.92	–6.98	–7.44	–0.16	2.19
Limerick	68.78	63.72	–4.71	–5.96	68.39	65.71	–5.22	–4.85	–0.39	1.99
Tipperary	69.03	63.77	–4.36	–5.90	68.48	65.68	–5.09	–4.90	–0.55	1.91
Waterford	69.85	66.24	–3.23	–2.25	69.63	68.18	–3.50	–1.28	–0.22	1.94
Ulster										
Antrim	81.46	80.76	12.85	19.19	81.44	80.73	12.86	16.90	–0.02	–0.03
Armagh	75.43	72.22	4.50	6.59	75.51	72.84	4.64	5.47	0.08	0.62
Cavan	69.82	63.66	–3.28	–6.05	69.92	64.67	–3.11	–6.36	0.10	1.00
Donegal	70.12	64.34	–2.86	–5.04	70.46	65.45	–2.36	–5.24	0.34	1.10
Down	79.31	77.21	9.87	13.95	79.45	77.80	10.10	12.65	0.14	0.59
Fermanagh	75.76	69.78	4.96	2.98	76.00	70.58	5.33	2.19	0.24	0.80
Londonderry	76.22	73.35	5.59	8.25	76.55	73.89	6.09	6.99	0.33	0.54
Monaghan	70.69	66.00	–2.07	–2.60	70.81	66.82	–1.87	–3.25	0.12	0.82
Tyrone	74.17	69.61	2.76	2.73	74.55	70.27	3.32	1.75	0.38	0.66
Connacht										
Galway	64.35	59.69	–10.85	–11.92	64.12	62.29	–11.14	–9.81	–0.23	2.60
Leitrim	73.27	65.06	1.50	–3.99	73.09	66.38	1.29	–3.88	–0.18	1.33
Mayo	66.56	56.93	–7.80	–15.99	66.36	58.58	–8.04	–15.18	–0.20	1.65
Roscommon	67.87	62.60	–5.98	–7.61	67.63	64.32	–6.27	–6.87	–0.23	1.72
Sligo	67.81	61.98	–6.06	–8.53	67.81	63.29	–6.02	–8.36	0.00	1.31
National										
Mean	72.18	67.76			72.16	69.06			–0.03	1.30
Weighted***	71.60	68.54			71.51	69.78				

Note: * King's County is now known as County Offaly. ** Queen's County is now known as County Laois. Underlined constitute the six counties of Northern Ireland from 1921. *** Weighted by county population size (23–62-year-olds).
Source: See Table 2.