DOI: 10.1111/ehr.13140

THE ECONOMIC HISTORY REVIEW

ORIGINAL ARTICLE

Age heaping and its discontents: A response to Baten, Benati, and Ferber

Brian A'Hearn[1] | **Alexia Delfino**[2] | **Alessandro Nuvolari**[3]

[1] Pembroke College, University of Oxford

[2] Department of Economics, Bocconi University

[3] Sant'Anna School of Advanced Studies, Institute of Economics

Correspondence
Alessandro Nuvolari
Email: alessandro.nuvolari@santannapisa.it

Funding information
Italian Ministry of University and Research, Grant/Award Number: 2017YLBYZE

Abstract

This note is a rejoinder to Baten, Benati, and Ferber. We reiterate that, on close inspection, nineteenth-century Italian census data contain a number of anomalies that sit uncomfortably with a straightforward age-heaping-as-numeracy interpretation. In particular, we respond to Baten, Benati, and Ferber on the technical matters they have raised; then we show that our findings are robust to such criticism. Finally, we conclude with some general reflections on age heaping as a numeracy indicator.

KEYWORDS
age heaping, human capital, Italy, numeracy

In our article 'Rethinking age heaping: a cautionary tale from nineteenth-century Italy', we advocated a careful approach to age heaping in economic history.[1] In particular, we argued against assuming that age heaping is an accurate measure of numeracy skills unless this can be justified from a careful assessment of the historical context and the data-generating process. In their comments on our article, Baten, Benati, and Ferber (hereafter Baten et al.) are in substantial agreement with our main conclusion.[2] Still, they believe that our concerns are exaggerated. Their comments provide us with the opportunity to clarify our position further and to flag some important remaining issues in age-heaping research.

Baten et al. worry that our argument may be misperceived as entirely negative, a blanket attack on all age-heaping research. To dispel any such impression, it may be worth recounting the origins of our project. Our original research question was whether age heaping could reveal meaningful differences in human capital endowments below the surface of a sea of illiteracy that

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

[2] Baten, Benati, and Ferber, 'Rethinking age heaping again'.

characterized early post-unification Italy. Our original approach was therefore firmly rooted in the newly emerging age-heaping-as-numeracy interpretation, to which one of us was an early contributor. Only as the census data began to reveal anomalies that sat uncomfortably with a straightforward numeracy interpretation did we embark on a broader reconsideration of the merits and the limitations of the current use of age heaping as an unambiguous measure of individual cognitive ability. Our goal remains improving our understanding of age heaping, not discarding it from the economic historian's toolkit.

The comments of Baten et al. entwine positive and negative strands. On the positive side they make a case for a reliable correlation between age heaping and literacy and numeracy, reanalysing our data as well as presenting new evidence from other times and places. We find little to quibble with here and welcome the intriguing new evidence from Spanish Inquisition records and contemporary Africa. For their part, Baten et al. accept—or at least do not contest—what we say about the acquisition of quantitative reasoning outside the context of school education and the salient influence of culture and state capacity on age awareness and accurate reporting. The negative elements include criticisms of our techniques and our interpretation of the Italian census data. In what follows we will first respond, briefly, on the technical matters (section I), and then show that our results on literacy and marital status are robust to such criticism (section II), before concluding with some reflections on age heaping as a measure of numeracy (section III).

I

Baten et al. argue that our estimates of age heaping are unconventional, and constitute a new variant. Two criticisms relate to the data we study. The first is that we analyse age data that are not uniformly distributed, and hence inappropriate for the Whipple index (W).[3] In fact, every researcher—including Whipple himself—who has studied the census data of a country with a classic 'population pyramid' demographic structure has adopted this approach. Furthermore, census-based Whipple values are a key part of what we think we know about age heaping around the world in the late nineteenth century.[4] To be sure, our New York passenger data show a somewhat sharper decrease in frequencies at higher ages than the census, but the difference is one of degree, not of kind. We reported W for the emigrants for the sake of simplicity and comparability with other figures in the article; the results are identical with alternatives better suited to this sort of distribution, such as the Myers blended index. There is no doubt that emigrant ages were dramatically less heaped than census ages.[5]

The second data criticism is that in our article there is limited attention to the requirements of self-reporting without cross-checking in samples used for estimating numeracy using age heaping.[6] We are puzzled by this remark because this point is very much in line with our insistence on studying the details of the data-generating process. Despite raising this issue in their survey of the age-heaping literature, Tollnek and Baten play down such concerns for a wide range of historical

[3] Ibid., p. 2.

[4] Crayen and Baten, 'Global trends.'

[5] The Myers index, which captures any form of digit preference in the data rather than heaping on multiples of five specifically, when calculated over ages 23 to 72, yields values for the emigrants of 1.90 (men) and 3.55 (women), compared to 9.06 and 12.04, respectively, in the census. These values can be interpreted as the percentage of observations that are misreported. For details of the Myers index, see Shyrock, Siegel, and Stockwell, *Methods and materials of demography*.

[6] Baten et al., 'Rethinking age heaping again', p. 2.

sources. Concerning census data specifically, they write: 'In censuses executed by governmental authorities and in times in which obligatory identification did not exist we can assume that ages are not counterchecked'.[7] We think that is probably right for the Italian census of 1871 but perhaps not for that of 1901, in which we see a sharp drop in age heaping in the south. The general lesson is that it is not generally obvious whether cross-checks have been made. Superficially, the 1871 census data appear to have been recorded as declared by respondents; after all, they are clearly heaped. On that basis the Italian censuses would pass the acid test proposed by Tollnek and Baten, and they have in fact been used in global age-heaping databases. One layer deeper, a bit of probing into the details of census procedure reveals that enumerators were urged to double-check ages. In fact, however, such guidelines were widely ignored, and the repeated exhortations of central officials reflect their growing frustration rather than their control of the process.

Two further criticisms pertain to our measurements. The first regards our *T*-index, developed for use with age frequencies aggregated into five-year bins in the later Italian censuses. Baten et al. argue that this is something quite different from the Whipple index and should not be called a measure of age heaping.[8] In fact the intuition underlying the two indices is similar, the difference being that the *T*-index captures only heaping on multiples of 10, whereas *W* is also sensitive to heaping on fives (while *both* miss heaping on even numbers). In any case the censuses of 1861 and 1871, which cover a wide range of demographic structures and age-heaping intensities, allow us to calculate both *W* and *T* values; they are well correlated, as discussed in the online appendix to our article.[9] We remain confident in our claim that heaping on 10s decreased sharply in the census of 1901 and make no broader claims about the *T*-index. Finally, Baten et al. claim that we fail to organize the data by birth decade or consider age groups consistently.[10] We reject the claim about birth cohorts. This issue only arises with longitudinal data, for in a cross-section birth cohort and age group are indistinguishable. When we have longitudinal data, we do analyse it explicitly by birth cohort, as in figure 4 of our article, which traces individual cohorts through successive censuses over 50 years. As for age groups, it is true that our figure 2 and tables 1 and 3 reported *W* values averaged across age groups for 1871, but this was entirely for simplicity of presentation; in what follows we control explicitly for age group and show that none of our results are affected.

II

In 'Rethinking' we documented surprising variation in age heaping within the groups of the literate and illiterate, and highlighted the paradox of literate men in southern cities reporting their ages no more accurately than illiterate peasant women in the north.[11] Baten et al. criticize us for assuming that census-reported literacy is a good measure of education.[12] This was not an assumption, but the outcome of our analysis of several alternative measures of literacy and schooling, briefly discussed in footnote 40 of the article.[13] We continue to trust census literacy as a measure of education more than census age heaping as a measure of numeracy.

[7] Tollnek and Baten, 'Age heaping', p. 362.

[8] Baten et al., 'Rethinking age heaping again', p. 2.

[9] A'Hearn et al., 'Rethinking age heaping', online app. S1.

[10] Baten et al., 'Rethinking age heaping again', p. 2.

[11] A'Hearn et al., 'Rethinking age heaping', p. 121.

[12] Baten et al., 'Rethinking age heaping again', p. 7.

[13] A'Hearn et al., 'Rethinking age heaping', n. 40.

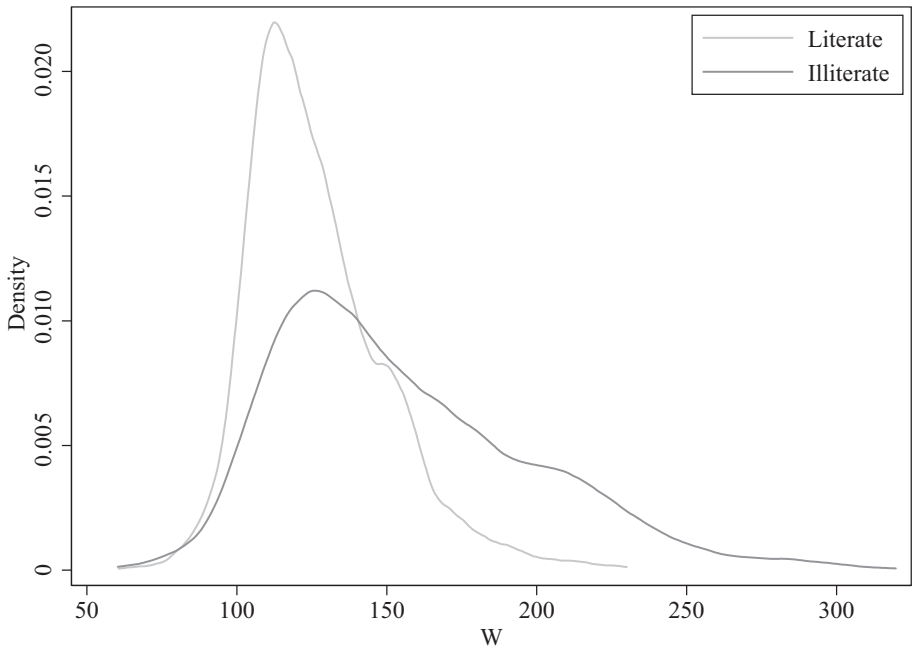

FIGURE 1 Age heaping by literacy status. Epanechnikov kernel density estimates of the distribution of Whipple index values for each (province–age group–gender–residence) observation, separately by literacy status. *Source*: Our calculations; MAIC, *Censimento*.

Baten et al. go on to restate the case for literacy and age heaping being correlated.[14] We agree about this, but reiterate that literacy has surprisingly little explanatory power. Figure 1 presents the distributions of W in the 2760 observations of our dataset.[15] Were education the main determinant of numeracy, and numeracy in turn of age heaping, we would expect two distinct distributions; instead, there is a very substantial overlap. If we regress W on literacy, region, gender, urban residence, and age group, only 15 per cent of the variation can be attributed to literacy, compared to 31 per cent for region or 12 per cent for age group (table 1). Education has more explanatory power within certain subgroups, but does not account well for differences between such groups—the difference between north and south in particular. Of course, numeracy could be acquired outside school, and this would attenuate neat contrasts between literates and illiterates. But non-academic sources of, and incentives for, numeracy were present even where age heaping was high according to our evidence, which Baten et al. do not contest. Differences in local state capacity and culture offer a more plausible explanation for the north–south gradient of age-heaping patterns and its evolution over time. Table 2 shows that none of these conclusions is sensitive to our treatment of age groups.

Another anomaly documented in our article was the significant variation in age heaping by marital status.[16] Baten et al. worry that this undermines the credibility of women's age heaping as a

[14] Baten et al., 'Rethinking age heaping again', p. 7.

[15] 69 provinces × 5 age groups × 2 genders × 2 residence categories (urban and rural) × 2 literacy status possibilities yields 2760 observations on W. When studying marital status, not cross-tabulated with literacy in the census, we have three marital status categories instead of two literacy categories, and the number of observations rises to 4140.

[16] A'Hearn et al., 'Rethinking age heaping', tab. 3.

TABLE 1 Analysis of variance of age heaping

	Partial SS	(%)	Df	F	*p*-value
Model	2 311 112	61	24	176.2	0.00
Region	1 185 277	31	17	127.6	0.00
Literacy	569 369	15	1	1 041.6	0.00
Age group	470 831	12	4	215.3	0.00
Gender	78 019	2	1	142.7	0.00
Urban	7 616	0	1	13.9	0.00
Residual	1 495 044	39	2 735		

Note: Analysis of variance based on a regression of *W* on dummies for region (18), literacy, age group (5), gender, and urban residence. Partial sums of squares are invariant to ordering. Percentage figures are relative to the total sum of squares.
Sources: Our calculations; MAIC, *Censimento.*

TABLE 2 Whipple index values with three treatments of age group

	Urban				Rural			
	Literate		Illiterate		Literate		Illiterate	
	Male	Female	Male	Female	Male	Female	Male	Female
Panel A: Simple average, all age groups								
North	110	127	119	137	109	119	119	130
Centre	117	132	145	154	119	126	146	153
South	129	138	173	185	141	149	191	198
Panel B: Age group 33–42 only								
North	110	127	116	134	108	118	117	128
Centre	121	134	141	150	120	127	143	149
South	125	132	163	172	134	142	181	185
Panel C: Regression-standardized 33–42 basis								
North	112	128	117	132	109	120	116	126
Centre	120	132	141	152	120	127	143	149
South	122	132	164	174	136	141	181	185

Notes: Macro-areas are defined as follows: north: Piedmont, Liguria, Lombardy, Veneto; centre: Emilia, Tuscany, Marche, Umbria, Lazio; South-Islands: Abruzzi, Puglia, Campania, Basilicata, Calabria, Sicily, Sardinia. Panel A: simple averages of *W* across all age-groups (identical to tab. 1 of A'Hearn et al., 'Rethinking age heaping'). Panel B: *W* for the 33–42 age group. Panel C: *W* standardized to an age-group 33–42 basis, based on a regression of *W* on macro-area, age group, gender, residence, literacy, and their two-way interactions.
Sources: Our calculations; MAIC, *Censimento*.

measure of women's human capital, and present evidence of correlations between female literacy and female age heaping in their tables 1 and 2 and figure 1. Here again we do not disagree about the correlations, but reiterate that the differences in age heaping by marital status are surprising and systematic. This is just as true for men as women; we focused on women only because the issue had arisen previously in the literature, which offers a plausible conjecture about women's ages being adjusted to that of their husbands.

Figure 2 plots *W* values standardized to an age 33–42 basis for each macro-area, gender, and marital status, while figure 3 zooms in on the south and shows values for each age group

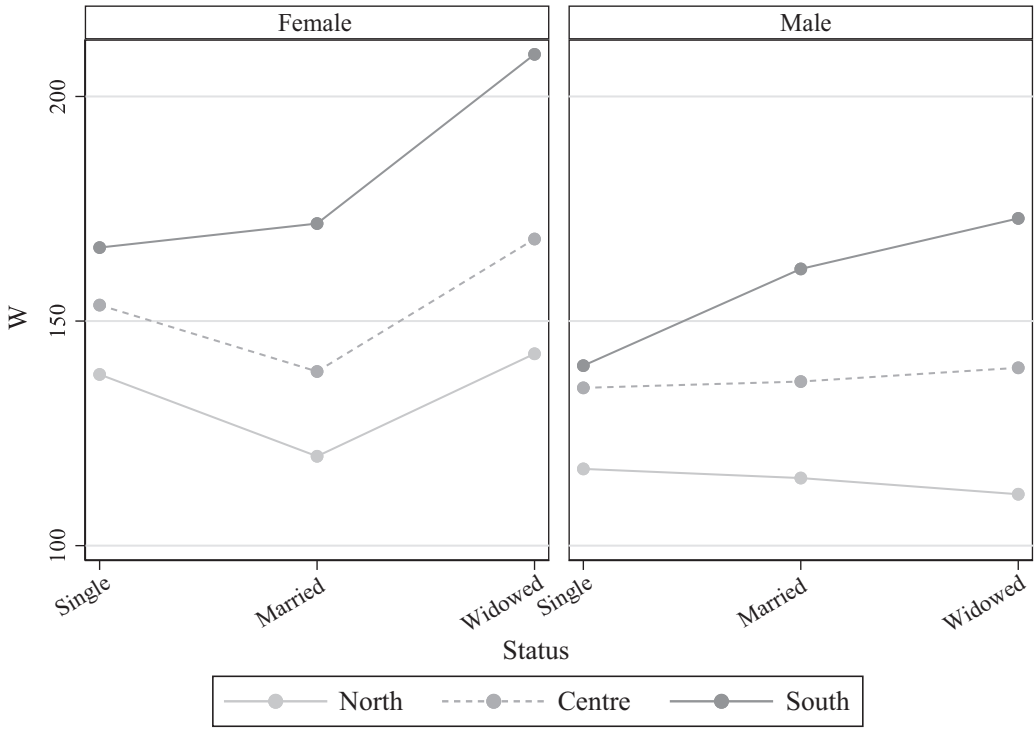

FIGURE 2 Age heaping by marital status. Average *W* values standardized to age group 33–42 based on a regression of *W* on macro-area, age group, residence, gender, marital status, and their two-way interactions. *Sources*: As for fig. 3.

separately.[17] Significant differences by marital status are apparent among women in all parts of the country, and, in the south, for men as well. Consider this last case in more detail. Among men aged 23–32 in the south, *W* is 37 points higher for married than for single men. If age heaping accurately reflects human capital, this implies substantial negative selection into (early) marriage. This is quite surprising given that, as Baten et al. note, 'during the mid-nineteenth century, marriage typically required a sufficient income'.[18] More plausible is that marriage caused men to report their ages differently.

III

Age heaping's correlation with literacy—both of them associated with the modernization process on our reading—justifies some confidence in its ability to proxy for education. As figure 4

[17] Figs. 3 and 4 plot predicted values from a regression of *W* on macro-area, age group, gender, residence, marital status, and their two-way interactions. Fig. 3 presents the figures for the 33–42 age group, which is equivalent to subtracting the relevant estimated age-group-related effects from every observation and averaging the resulting age-corrected figures; in that sense we call these estimates 'standardized to age 33–42'. The same procedure underlies panel C of tab. 2. Fig. 4 plots predicted values for each age group separately.

[18] Baten et al., 'Rethinking age heaping again', p. 5.

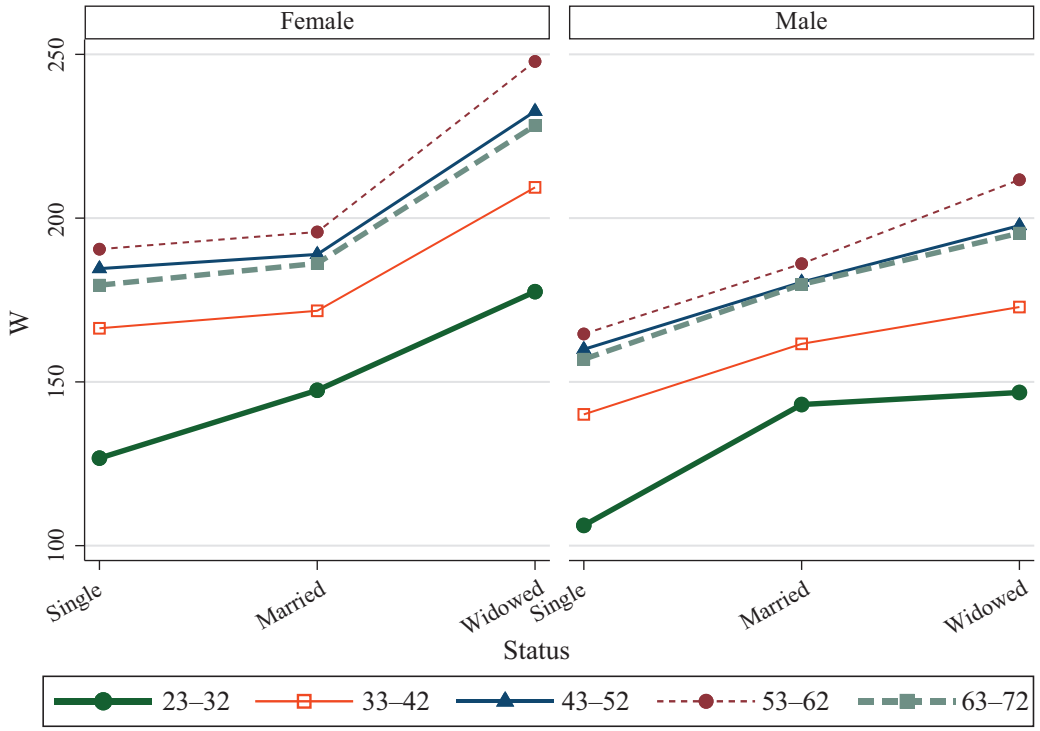

FIGURE 3 Age heaping by marital status and age group, South-Islands macro-area. Predicted values for each age group based on the model described for fig. 2. *Sources*: As for fig. 1 [Colour figure can be viewed at wileyonlinelibrary.com]

illustrates, this relationship is both non-linear and inconstant across groups in the Italian census, which complicates drawing precise quantitative inferences about literacy from age heaping.[19] Still, age heaping should capture the direction of meaningful variations between groups.

More problematic is the interpretation of age heaping as an indicator of numerical skills. Setting aside our concerns about culture and state capacity impinging on age heaping, let us assume that it is driven by numeracy. What exactly does that mean? We can think of numeracy, or *in*numeracy, varying on an extensive and an intensive margin. The extensive margin is the size of the innumerate group in the population, the intensive margin the severity of cognitive impairment relative to some standard of proficiency. The Whipple index is most naturally interpreted as measuring the *extensive* margin. A W value of 200 results when multiples of five are twice their expected share of ages, that is, 40 per cent instead of 20 per cent. If innumerates all report a heaped age while numerates report their age accurately, $W = 200$ implies a 25 per cent rate of innumeracy; more generally a linear relationship between innumeracy and W is implied by these assumptions.

The *intensive* margin of innumeracy has not been clearly theorized in relation to age heaping. What degree of numerical reasoning ability is the watershed between accurate and inaccurate knowledge of age? Does innumeracy mean individuals cannot count and have no clear sense of magnitudes or differences between numbers? That extreme would be sufficient for them not to

[19] Nonlinear relationships like those plotted in fig. 4, in which all age groups are pooled, are also found within each age group. Baten et al. similarly adopt non-linear specifications in exploring the correlation of literacy and age heaping, relating the logarithm of W to the level of literacy.

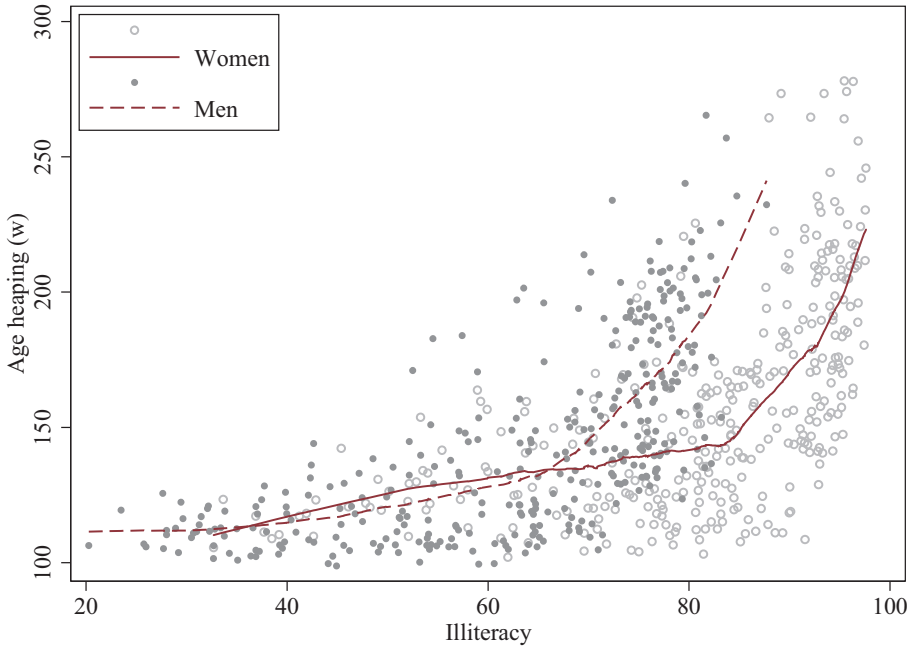

FIGURE 4 Illiteracy and age heaping. Each point represents one (province × age group × gender) observation. (Rural and urban figures are aggregated for clarity of presentation.) The fitted lines are estimated by locally weighted regression with a bandwidth equal to 20% of the sample. *Sources*: As for fig. 3 [Colour figure can be viewed at wileyonlinelibrary.com]

understand age in a quantitative way. Such a degree of ignorance would have major economic consequences. Or does innumeracy mean, at the other extreme, that although people can count and do the basic arithmetical operations, they are not quite fluent enough in mental mathematics to find a problem like (age = 2021–1964) effortless, and report an approximate age to save themselves the trouble? People with this degree of numeracy might struggle with compound interest or expected value but be functionally numerate in everyday life. Clarifying the relationship between age heaping and innumeracy is imperative if age heaping is to realize its potential as an independent indicator of specific cognitive abilities.

ACKNOWLEDGEMENTS

This research has been supported by the project 'Lost highway. Skills, technology and trade in Italian economic growth, 1815–2018', funded by the Italian Ministry of University and Research, Progetti di Rilevante Interesse Nazionale (PRIN) Bando 2017, n. 2017YLBYZE.

Open Access Funding provided by Scuola Superiore Sant'Anna within the CRUI-CARE Agreement.

REFERENCES

A'Hearn, B., Delfino, A., and Nuvolari, A., 'Rethinking age heaping: a cautionary tale from nineteenth-century Italy', *Economic History Review*, 75, 1 (2022), pp. 111–37.

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

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.

Shryock, H. S., and Siegel, J. S., *The methods and materials of demography*, ed. E. G. Stockwell (San Diego, Calif., 1976).

Tollnek, F. and Baten, J., 'Age-heaping-based human capital estimates', in C. Diebolt and M. Haupert, eds., *Handbook of cliometrics* (2nd edn., Cham, 2019), pp. 357–80.

OFFICIAL PUBLICATION

MAIC (Ministero di Agricoltura, Industria e Commercio), *Censimento generale al 31 dicembre 1871 della popolazione del Regno d'Italia*, II: *Popolazione classificata per età sesso, stato civile e istruzione elementare* (Rome, 1875).

How to cite this article: A'Hearn, B., Delfino, A., and Nuvolari, A., 'Age heaping and its discontents: A response to Baten, Benati, and Ferber', *Economic History Review*, 00 (2022), pp. 1–9. https://doi.org/10.1111/ehr.13140