Human Lifetime Entropy in a Historical Perspective (1750-2014)*

Patrick Meyer† Gregory Ponthiere‡

February 25, 2019

Abstract

This paper uses Shannon's entropy index to the base 2 to quantify the risk relative to the age at death in terms of bits (i.e. the amount of information revealed by tossing a fair coin). We first provide a simple decomposition of Shannon's lifetime entropy index that allows us to analyze the determinants of lifetime entropy (in particular its relation with Wiener's entropy of the event "death at a particular age conditional on survival to that age") and to study how the risk about the duration of life is resolved as the individual becomes older. Then, using data on 37 countries from the Human Mortality Database, we show that, over the last two centuries, (period) lifetime entropy at birth has exhibited, in all countries, an inverted-U shape pattern with a maximum in the first half of the 20th century (at 6 bits), and reaches, in the early 21st century, 5.6 bits for men and 5.5 bits for women. It is also shown that the entropy age profile shifted from a non-monotonic profile (in the 18th and 19th centuries) to a strictly decreasing profile (in the 20th and 21st centuries).

Keywords: mortality risk, longevity, age at death, entropy, measurement.

JEL codes: D80, J11, N30.

*We authors would like to thank Bertrand Villeneuve, Joel Sobel, as well as two anonymous referees for their helpful comments and suggestions on this paper.

†IMT Atlantique, Lab-STICC, Univ. Bretagne Loire, F-29238 Brest, France, E-mail: patrick.meyer@imt-atlantique.fr

‡University Paris East (ERUDITE), Paris School of Economics, and Institut universitaire de France. [corresponding author] Address: ENS, bd Jourdan 48, 75014 Paris. E-mail:gregory.ponthiere@ens.fr Telephone: 0033 01 43 13 63 71. Fax 0033 01 43 13 63 10.

1 Introduction

Mors certa, hora incerta. As Jankélévitch emphasized in his treatise *La Mort* (1977), uncertainty about the duration of life constitutes a major dimension of human condition. Everyone knows that death is inevitable, and, hence, that one will necessarily die one day. But no one knows when precisely (which year, which day, which hour) he or she will die. Life tables, which have been computed since the 17th century, allowed humans to shift from uncertainty to risk about the duration of life, but this shift does not solve the problem: with or without life tables, it is still the case that individuals ignore when death will occur.

Although there exist various indicators of life riskiness, such as indicators of verticalization and horizontalization of survival curves (see Wilmoth and Horiuchi, 1999), empirical studies show that individuals have still major difficulties in evaluating how risky their life is. In a seminal study using questionnaires aimed at eliciting subjective survival probabilities, Hamermesh (1985) shows that, although respondents are well aware of the levels of improvements within current life tables, the subjective distribution of the age at death is flatter than the actuarial distribution, implying that the subjective variance of lifespan is larger than the variance of lifespan in actual life tables. Mirowski (1999) and Brouwer and van Exel (2005) reveal significant biases in the subjective assessment of survival prospects, including excessive optimism. Dormont et al (2018) show that subjective uncertainty about length of life is quite large, and can hardly be explained by standard socio-demographic variables (except the age).

How could one improve individual knowledge on risk about the duration of life? That question raises general concerns about attitudes and cognitive bias when assessing risky situations (see Tversky and Kahneman 1975, Gigerenzer et al 2005). One possibility could be to develop an indicator of life riskiness that would make the riskiness of life somewhat *commensurable* with the riskiness involved in other, more common, situations. Thus, in order to make individuals have a more precise idea of life riskiness, it can be useful to make it commensurable with other risks, through the use of a common standard.

The goal of this paper is to develop an indicator of risk about the duration of life whose metric has a concrete counterpart for the layman, and makes that risk commensurable with the risk involved in more common situations. We propose to measure risk about the duration of life by means of Shannon's lifetime entropy index defined to the base 2 (Shannon 1948, Pierce, 1980):[1]

$$H_k = -\sum_{i=k}^{115} p_{i,k} \log_2 \left(p_{i,k}\right) \tag{1}$$

where $p_{i,k}$ is the probability of a life of duration i for an individual of age k.

The index H_k measures the mathematical expectation, along the life cycle, of the amount of information that is learnt from the event "death at a particular age $i \geq k$". That index quantifies the risk relative to the duration of life (or,

[1] Throughout this paper, the maximal age is fixed to 115 year (at that age, the life table is closed, i.e. $d_{115} = 1$).

similarly, the risk about the age at death) in terms of *bits*, i.e. the amount of information revealed by tossing a fair coin. The choice of that measurement unit - the bit - amounts to take coin flipping as the common standard on the basis of which life riskiness is assessed. Relying on that standard can be justified on the grounds that humans are familiar with coin flipping since the Roman empire (see Lanciani 1892).

This paper has a twofold goal. First, on the theoretical side, we propose to study the determinants of Shannon's lifetime entropy index. In particular, we analyze the link between, on the one hand, the measurement of risk about the duration of life, and, on the other hand, the measurement of risk of death at a particular age (conditionally on survival to that age). Understanding this link allows us also to study how the risk about the duration of life is (progressively) resolved as the individual becomes older. Second, on the empirical side, we compute Shannon's lifetime entropy index using period and cohort life tables for 37 countries from the Human Mortality Database, to study the size and evolution of the risk about the duration of life over the last two centuries. This allows us to measure the amount of information revealed by the event "death at a particular age" in terms of bits, and, hence, to make life riskiness commensurable with the risk involved in tossing a fair coin. By computing lifetime entropy indexes by age, we are also able to quantify how the amount of risk about the duration of life that remains *unresolved* varies with the age.

Anticipating our results, our theoretical analysis of Shannon's lifetime entropy index allows us, by means of a simple decomposition of the index, to rewrite H_k as the mathematical expectation, across the remaining lifecycle, of the amount of information learnt by death at a particular age $i \geq k$, which is the sum of two components: on the one hand, the amount of information revealed by the event "death at age i conditionally on survival to that age" (measured by Wiener's entropy of the single event "death at age i conditionally on survival to age i"); on the other hand, the amount of information revealed by the event "surviving up to age i" (measured by Wiener's entropy of the single event "surviving up to age i"). Shannon's lifetime entropy aggregates all those quantities of information, while weighting these by the probability of occurrence of death at age $i \geq k$. That decomposition of Shannon's lifetime entropy index allows us also to show that lifetime entropy does not necessarily decrease in case of survival from age k to age $k+1$, and can increase, if the probability of death at age k is larger than the probabilities to die at higher ages ($k+1$, $k+2$, etc.).

On the empirical side, we identify 6 stylized facts: (1) over the last two centuries, (period) lifetime entropy at birth exhibits an inverted-U shape with a maximum in the first half of the 20th century (at about 6 bits); (2) curves of (period) lifetime entropy at birth for men and women crossed during the 20th century; (3) over the last 150 years, Western countries have been converging in terms of (period) lifetime entropy at birth towards 5.6 bits for men and 5.5 bits for women; (4) the lifetime entropy age profile has shifted from a non-monotonic profile (in the 18th and 19th centuries) to a strictly decreasing profile (in the 20th and 21st centuries); (5) the lifetime entropy age profile becomes steeper at higher ages, coinciding with a rise in the speed of learning about one's duration of life; (6) men exhibit a higher lifetime entropy than women below ages 50-55, and a lower one after ages 50-55.

This paper complements the literature on the measurement of risk about the duration of life (see Wilmoth and Horiuchi 1999, Van Raalte and Caswell 2013). Existing measures of life riskiness include measures of the horizontalization of survival curves, which capture how long a cohort can live before aging-related deaths reduce the proportion of survivors (Lan Karen Cheung et al 2005). Other measures of life riskiness include indicators of the verticalization of survival curves, such as the coefficient C_{50} (Kannisto 2000), which measures the shortest age interval necessary to concentrate 50 % of life durations, and indicators of the standard deviation in the age at death (Lan Karen Cheung and Robine 2007, Edwards and Tuljapurkar 2005). Indicators of verticalization include also measures of interquantile range, Gini indices of the length of life (Smits and Monden 2009) and measures of life disparity, i.e. the average remaining life expectancy at the ages when death occurs (Vaupel et al 2011).

In particular, this paper brings a contribution to the literature on the measurement of the variance of the age at death using entropy indicators, such as Demetrius (1976), Nagnur (1986), Hill (1993) and Noymer and Coleman (2014).[2] Our contribution with respect to those studies is twofold. First, we use, unlike those studies, Shannon's lifetime entropy index defined to the base 2. This particular base allows us not only to make our measure of lifetime entropy rely on the intuitive bit metric (which makes the risk about the duration of life commensurable with the risk involved in the most widespread randomization device, i.e. tossing a fair coin), but, also, to analyze how the risk about the duration of life can be related to Wiener's entropy of the single event "death at a particular age conditionally on survival to that age" (Wiener's entropy being also defined to the base 2). Second, we apply Shannon's lifetime entropy index to the Human Mortality Database, and consider various period/cohort measures of lifetime entropy, in order to quantify how large the risk about the duration of life is, and how that risk varies across ages, genders, countries and epochs.

The rest of the paper is organized as follows. Section 2 studies, from a theoretical perspective, the determinants of Shannon's lifetime entropy index, and examines how that index is related to Wiener's entropy index of a single event (death at a particular age conditionally on surviving to that age). Section 2 also examines how Shannon's lifetime entropy index at a particular age varies with the age, in order to study how the risk about the duration of life is resolved as the individual becomes older. Then, Section 3 studies, on the basis of the Human Mortality Database (period life tables), the dynamics of Shannon's lifetime entropy at birth over the last 2 centuries. Still using period life tables of the Human Mortality Database, Section 4 studies how Shannon's lifetime entropy varies along the life cycle. Then, Section 5 computes Shannon's lifetime entropy index based on cohort life tables. Conclusions are drawn in Section 6.

[2] See the Appendix on the relationship between Shannon's lifetime entropy index defined to the base 2 and other entropy indicators.

2 Theory

This section aims at analyzing, from a theoretical perspective, the determinants of Shannon's lifetime entropy index defined to the base 2. Our analysis will also allow us to explain the reasons why this index constitutes an intuitive measure of the quantity of risk associated to a human life.

For that purpose, we will proceed in three steps. In a first step, we use Wiener's entropy index in order to study the amount of information revealed by death at a particular age k (conditionally on having survived to that age). Then, in a second stage, we will present Shannon's lifetime entropy index as a measure of the expected amount of information revealed by a death at an age k (unconditionally) and will relate this to Wiener's entropy index. Finally, we will study the dynamics of Shannon's lifetime entropy index with the age.

2.1 Learning about death at age k: Wiener's entropy

Let us first consider how one can quantify the amount of information revealed by the death of an individual at a particular age. For that purpose, this section will rely on Wiener's entropy index (see Wiener 1965), which measures the amount of information revealed by the occurrence of a single event whose probability of occurrence is known.[3]

When a person of age k dies, while the probability of death at age k (conditionally on reaching age k) is d_k, the amount of information learnt by the event "death at age k" is given by Wiener's entropy:

$$W(d_k) = -\log_2(d_k) \tag{2}$$

That formula provides the amount of information revealed by the death at age k (conditionally on survival to age k) in terms of bits, i.e. the amount of information revealed from tossing a single fair coin.

It is easy to see that $W(d_k)$ equals 0 bits when the death was certain (i.e. when $d_k = 1$). Indeed, in that case, nothing is learnt from the event of a death. However, when $d_k < 1$, Wiener's entropy index is strictly positive and decreasing in d_k: the less likely was the death at age k, and the larger is the quantity of information that we learn from the death of an individual at age k.

At this stage, an illustration may be useful, to better understand the meaning of the measurement unit (i.e. the bit) on which Wiener's entropy index relies. If, for instance, the probability to die at age k is equal to 1/128 (about 0.78 %), then the death of an individual at age k reveals a quantity of information equal to 7 bits, that is, the amount of information revealed by tossing 7 fair coins. How large is that amount of information? To see this, one should remind that, when 7 fair coins are drawn, there exist $2^7 = 128$ possible outcomes in terms of "heads" (H) and "tails" (T) (permutations being regarded as yielding different

[3] An axiomatization of Wiener's entropy index can be found in Aczél and Daroczy (1975). Those authors show that Wiener's entropy $W(p) = -\log_2(p)$ is the unique function $F(p)$ satisfying 3 axioms: (1) non-negativity $F(p) \geq 0$; (2) additivity: $F(pq) = F(p) + F(q)$; (3) normalization: $F(1/2) = 1$.

outcomes). Thus, once the particular combination of "heads" and "tails" that emerges from the 7 draws is known, one learns quite a lot, since many other outcomes were also possible but did not occur. In the light of this, when we say that the death of an individual who had a probability of 1/128 to die reveals 7 bits of information, we make the occurrence of this death event commensurable with the emergence of one particular combination of "heads" and "tails" out of the 128 different possible combinations from tossing 7 fair coins. When we say that the death of an individual at age k reveals 7 bits of information, we mean that that event "death at age k" was exactly as unlikely as any combination of "heads" and "tails" resulting from 7 draws of a fair coin.

Wiener's entropy of the event "death at age k" can be regarded as a simple way to quantify the extent to which the event of a death brings an important piece of information. As such, Wiener's entropy allows us to quantify how surprising or astonishing the event of a death is. By doing so, Wiener's entropy indicator allows us to put numbers on some important aspect of the mental attitude of individuals towards death.[4]

Quantifying how astonishing the event of a death is can cast some light on the history of human's attitudes towards death. Actually, as shown by Ariès (1975), humans' attitude towards death has strongly evolved over time. Ariès (1975) emphasized that, during the longest part of History, humans used to be familiar with death, since death was something that occurred quite often. However, with the decline of mortality, human's attitude towards death has changed: in the 20th century, death has become something strange, a kind of taboo, about which individuals hardly talk.[5] Wiener's entropy of the event "death at age k" is a simple way to quantify that, as the strength of mortality went down, the degree of astonishment due to the event of a death went up, and reached levels that are so large that death has become something with which individuals are no longer familiar.

Although Wiener's entropy index provides an intuitive measure of the quantity of information revealed by the death of an individual at a particular age, and, as such, helps us quantifying the change in the attitude of humans towards death, it suffers nonetheless from an important limitation. Wiener's entropy index of the event "death at age k" provides only a local measure of life riskiness. Actually, Wiener's entropy gives us the amount of information revealed by the death at a particular age *provided* the individual reached that age. One may want to have a more global measure of the risk about the duration of life, which would quantify the amount of information that we learn from the death of someone at a particular age, unconditionally on surviving to that age. A more global measure of life riskiness is proposed in the next section.

[4] Note that this is only one aspect of mental attitudes towards death. See Godelier (2018) on the representation of death in 14 religious systems of thought.

[5] In his history of mental attitudes towards death, Ariès (1975) distinguishes between 4 steps: (1) death as a familiar event (*mort familière*); (2) emergence of individualism of death (*mort de soi*); (3) emergence of a concern for joint survival (*mort de toi*); (4) death as a taboo (*mort interdite*). In the above paragraph, we are concerned with the progressive shift from (1) to (4), and we deliberately leave dimensions (2) and (3) aside.

2.2 Learning about death at age k: Shannon's lifetime entropy

In order to measure the quantity of information revealed by a death at a particular age, unconditionally on surviving to that age, this paper will rely on Shannon's lifetime entropy index defined to the base 2. Shannon's lifetime entropy index at age k can be written as:[6]

$$H_k = -\sum_{i=k}^{115} p_{i,k} \log_2 (p_{i,k}) \tag{3}$$

where $p_{i,k}$ is the probability of a life of length i for an individual of age k.

Shannon's lifetime entropy index at age k provides a measure, in terms of bits, of the quantity of risk associated to the duration of life beyond age k. It measures the mathematical expectation, along the (remaining) life cycle, of the amount of information revealed by a death at age $i \geq k$.[7] Shannon's lifetime entropy index at age k is also a measure of the expected amount of risk about the age at death that remains unresolved at age k.

In order to interpret that formula, it is useful to relate Shannon's lifetime entropy index with Wiener's entropy index of a single event. Actually, Shannon's lifetime entropy index can be decomposed as follows:

$$\begin{aligned} H_k &= -\sum_{i=k}^{115} (s_{i,k} d_i) \log_2 (s_{i,k} d_i) \\ &= \sum_{i=k}^{115} (s_{i,k} d_i) \left[-\log_2 (s_{i,k}) + (-\log_2 (d_i)) \right] \\ &= \sum_{i=k}^{115} (s_{i,k} d_i) \left[W(s_{i,k}) + W(d_i) \right] \end{aligned} \tag{4}$$

where $s_{i,k} = \prod_{j=k}^{i-1} (1 - d_j)$ is the probability of survival up to age i for a person of current age k. The product $s_{i,k} d_i$ is the probability of a life of exact duration i for an individual of current age k.

When rewritten in that way, we see immediately that when an individual dies at age $i \geq k$, there are two pieces of information that are revealed from that

[6] Note that Shannon's entropy index is defined for all kinds of distributions, including these for which outcomes are not equally likely. The distribution with equally-likely outcomes is just one case among many others. An interesting feature of that particular case is that, for a distribution with n possible outcomes, it is when those outcomes are equally likely, and thus arise with a probability $1/n$, that entropy is maximal. That property was part of Khinchin's (1957) early axiomatization of the entropy index (for all bases).

[7] Simple axiomatizations of Shannon's entropy index defined to the base 2 can be found in in Aczél and Daroczy (1975). A previous axiomatization, for more general entropy indexes (any base) can be found in Khinchin (1957). Khinchin (1975) shows that the general entropy index can be derived from 3 conditions on the (continuous) function $G(p_1, ..., p_n)$: (1) $G(\cdot)$ takes its maximum at $\left(\frac{1}{n}, ..., \frac{1}{n}\right)$; (2) $G(AB) = G(A) + G(B)$; (3) $G(p_1, ..., p_n, 0) = G(p_1, ..., p_n)$.

event. First, we learn that the individual has survived to age i, which gives an amount of information $-\log_2(s_{i,k})$, which coincides with Wiener's entropy of the event "survival to the age i"; second, we also learn that the individual died at age i after having survived to that age, which gives us an extra amount of information equal to $-\log_2(d_i)$, which coincides with Wiener's entropy of the event "death at age i" (conditionally on survival to that age).

Thus, from a lifecycle perspective, the event "death at age i" reveals an amount of information that is larger than $W(d_i)$, since the mere fact of having survived to age i already provides a quantity of information $W(s_{i,k})$. Indeed, when a life has exact duration i, we discover both that the individual survived to age i, and died at age i.

In the light of the above formula, Shannon's lifetime entropy index at age k can be interpreted as a mathematical expectation, across the entire lifecycle, of the total amount of information revealed by the event of a death at a particular age $i \geq k$. Indeed, it aggregates, across all ages $i \geq k$, the amounts of information revealed by the event "death at age i" (unconditionally on having survived to that age), i.e. $W(s_{i,k}) + W(d_i)$, and weights these amounts of information with the probability of occurrence of the associated events.

Shannon's lifetime entropy index at age k can be used as a measure of the risk about the duration of life that remains to be resolved at age k. Clearly, the larger (resp. lower) Shannon's lifetime entropy index at age k is, and the larger (resp. lower) the degree of astonishment associated to a death at a future age $i \geq k$ is, implying that the risk about the duration of life that is unresolved at age k is larger.

An interesting feature of lifetime entropy is that it takes its maximal value when all possible durations of life are equiprobable, that is, when $p_{i,k} = p_k \forall i \geq k$, which implies that $p_k = \frac{1}{115-k}$.[8] Hence, whether lifetime entropy is higher or lower at a particular period of History depends on how distant the distribution of the possible durations of life is with respect to the equiprobable case. From that perspective, one can make the conjecture that, in preindustrial times, the high infant and child mortality used to make deaths concentrated at low ages, leading us far from the equiprobable case. Regarding the modern demographic regime, the rectangularization of the survival curve is associated to the concentration of deaths at high ages, leading us also far from the equiprobable case. We can thus expect that Shannon's lifetime entropy was low in preindustrial times, and also low in modern times, and that it may have taken higher values during the transition from one regime to the other. Section 3 will explore that issue further, using data from the Human Mortality Database.

[8] This property was made explicit in Khinchin's (1957) early axiomatic characterization of the general entropy index. In the case of Shannon's lifetime entropy index, it is easy to check this by derivating the values of all $p_{i,k}$ that maximize H_k subject to the constraint that the $p_{i,k}$ sum up to 1. The solution is given by each $p_{i,k} = \frac{1}{115-k}$. When $k = 0$, $p_{i,k} = \frac{1}{115}$, which yields the maximal lifetime entropy at birth, equal to 6.905 bits.

2.3 Shannon's lifetime entropy along the life cycle

Let us now examine how Shannon's lifetime entropy index evolves with the age. This section will, by comparing Shannon's lifetime entropy index at two successive ages k and $k+1$, allow us to better understand how the mere survival from age k to age $k+1$ affects the amount of risk about the duration of life that remains unresolved.

For that purpose, let us rewrite the lifetime entropy index only in terms of age-specific probabilities of death:

$$H_k = -\sum_{i=k}^{115} \left(\prod_{j=k}^{i-1} (1-d_j) d_i \right) \log_2 \left(\prod_{j=k}^{i-1} (1-d_j) d_i \right) \tag{5}$$

We have

$$\begin{aligned}
H_k &= -\sum_{i=k}^{115} \left(\prod_{j=k}^{i-1} (1-d_j) d_i \right) \log_2 \left(\prod_{j=k}^{i-1} (1-d_j) d_i \right) \\
&= -d_k \log_2 (d_k) - (1-d_k) d_{k+1} \left[\log_2 (1-d_k) + \log_2 (d_{k+1}) \right] \\
&\quad -(1-d_k)(1-d_{k+1}) d_{k+2} \left[\log_2 (1-d_k) + \log_2 (1-d_{k+1}) + \log_2 (d_{k+2}) \right] \\
&\quad -(1-d_k)(1-d_{k+1})(1-d_{k+2}) d_{k+3} \left[\begin{array}{c} \log_2 (1-d_k) + \log_2 (1-d_{k+1}) \\ + \log_2 (1-d_{k+2}) + \log_2 (d_{k+3}) \end{array} \right] \\
&\quad -...
\end{aligned}$$

Writing the Shannon's lifetime entropy index at age $k+1$, we obtain:

$$\begin{aligned}
H_{k+1} &= -d_{k+1} \log_2 (d_{k+1}) - (1-d_{k+1}) d_{k+2} \left[\log_2 (1-d_{k+1}) + \log_2 (d_{k+2}) \right] \\
&\quad -(1-d_{k+1})(1-d_{k+2}) d_{k+3} \left[\log_2 (1-d_{k+1}) + \log_2 (1-d_{k+2}) + \log_2 (d_{k+3}) \right] \\
&\quad -(1-d_{k+1})(1-d_{k+2})(1-d_{k+3}) d_{k+4} \left[\begin{array}{c} \log_2 (1-d_{k+1}) + \log_2 (1-d_{k+2}) \\ + \log_2 (1-d_{k+3}) + \log_2 (d_{k+4}) \end{array} \right] \\
&\quad -...
\end{aligned}$$

By multiplying the latter expression by $(1-d_k)$, and then substituting for H_k, one obtains:

$$\begin{aligned}
H_{k+1}(1-d_k) &= H_k + d_k \log_2 (d_k) + (1-d_k) d_{k+1} \log_2 (1-d_k) \\
&\quad +(1-d_k)(1-d_{k+1}) d_{k+2} \log_2 (1-d_k) \\
&\quad +(1-d_k)(1-d_{k+1})(1-d_{k+2}) d_{k+3} \left[\log_2 (1-d_k) \right] \\
&\quad +...
\end{aligned}$$

This expression can be rewritten as:

$$H_{k+1}(1-d_k) = H_k + d_k \log_2 (d_k) + (1-d_k) \log_2 (1-d_k) \sum_{i=k+1}^{115} d_i \prod_{j=k+1}^{i-1} (1-d_j) \tag{6}$$

This formula can be interpreted as follows:

$$\underbrace{H_{k+1}}_{\text{Shannon entropy at age } k+1}(1-d_k)$$

$$= \underbrace{H_k}_{\text{Shannon entropy at age } k} - d_k \underbrace{\left(-\log_2(d_k)\right)}_{\text{Wiener entropy of event death at age } k} - (1-d_k)\underbrace{\left(-\log_2(1-d_k)\right)}_{\text{Wiener entropy of event non-death at age } k}\underbrace{\sum_{i=k+1}^{115} d_i \prod_{j=k+1}^{i-1}(1-d_j)}_{\text{probability of a life of length } >k} \quad (7)$$

This formula admits a simple interpretation. Conditionally on being alive at age k, lifetime entropy at age $k+1$ can arise with a probability $1-d_k$. The product $H_{k+1}(1-d_k)$ is equal to lifetime entropy at the previous age, H_k, from which we subtract two things: first, the information that would have been revealed from death at age k, which would have occurred with probability d_k; second, the information revealed from non-death at age k (which arises with probability $1-d_k$), times the probability of a life of length superior to k, which is actually equal to $\sum_{i=k+1}^{115} p_i = 1 - p_k = 1 - d_k$. The last term of the above formula captures the updating of the probabilities of lives of exact lengths $i > k$ once there has been survival from age k to age $k+1$. When one had age k, the probabilities of lives of lengths $k+1$, $k+2$, etc., depended on $1-d_k$, which is no longer the case once age $k+1$ is reached.

In order to study the process of information revelation in case of survival from age k to age $k+1$, let us rewrite the above equation as:

$$H_{k+1} - H_k = \underbrace{\frac{d_k}{1-d_k}\left[H_k - W(d_k)\right]}_{?} - \underbrace{(1-d_k)W(1-d_k)}_{+} \quad (8)$$

That formula gives us the variation of remaining lifetime entropy between two successive ages. If $H_{k+1} - H_k < 0$ (resp. > 0), the passage from age k to age $k+1$ reduces (resp. increases) the (expected) amount of information revealed by the event of a death at a future age.

From that equation, it is easy to see that, when $d_k = 0$, the first and the second terms of the right-hand side are equal to 0, so that H_{k+1} is equal to H_k. When there is no risk of death at age k, the mathematical expectation of the amount of information revealed by a death at age $i \geq k$ is exactly equal to the expected amount of information revealed by a death at age $i \geq k+1$, since no information is revealed by surviving from age k to age $k+1$.[9]

[9] In line with Khinchin (1957), adding or subtracting a scenario that has no chance to occur does not affect the value of the entropy index.

Consider now the general case where $d_k > 0$. Whereas one may believe, at first glance, that lifetime entropy tends to decrease with the age, the above formula suggests that this is not necessarily the case. Actually, whether lifetime entropy decreases or increases with the age depends on the sign of the first term of the right-hand side, and on its size relative to the one of the second term, which is necessarily negative.

The general case is a decrease in lifetime entropy with the age, that is, $H_{k+1} - H_k < 0$. That general case arises when d_k takes a low value with respect to the probabilities of death at later ages. Indeed, in that case, the Wiener term $W(d_k)$ is larger than H_k (since premature death reveals then a larger quantity of information than the mathematical expectation of the quantity of information revealed from death at a higher age), implying that the first term on the right-hand side of equation (8) is negative, leading to $H_{k+1} - H_k < 0$.

However, the case where $H_{k+1} - H_k > 0$ can arise, if d_k takes a high value with respect to the probabilities of death at later ages, in such a way that the mathematical expectation of the quantity of information revealed by a death at a future age $i \geq k$ is larger than the quantity of information revealed from the event of death at age k, that is, $H_k > W(d_k)$. In that case, the first term of the right-hand side of (8) is positive. That condition is necessary, but not sufficient for $H_{k+1} - H_k > 0$. One needs also that the first term of the right-hand side exceeds, in absolute value, the second term of the right-hand side. To achieve this, it has to be the case that $H_k - W(d_k)$ is sufficiently large.

3 Lifetime entropy in a historical perspective

This section uses life tables from 37 countries in the Human Mortality Database to study the evolution of Shannon's lifetime entropy to the base 2 across countries, periods and genders.[10]

As a starting point, this section focuses on lifetime entropy *at birth*, based on period life tables. Lifetime entropy at birth H_0 measures the mathematical expectation of the amount of information that is revealed by a death at a particular age beyond age 0. This can also be interpreted as a measure of the amount of risk about the age at death that is unresolved at age 0, in the sense that a higher lifetime entropy at birth means that there is, on average, more astonishment associated to a death at age beyond age 0. Thus lifetime entropy at birth provides a synthetic measure of the degree of risk relative to the age at death when considering the entire potential lifespan.[11]

[10] Figure A2 in the Appendix summarizes the availability, in terms of year interval, of period life tables for each country within the database.

[11] Section 4 will consider the evolution of lifetime entropy by age, to examine the speed at which the risk about the duration of life is resolved as individuals become older.

3.1 Lifetime entropy at birth by country and gender

Let us first take the subsample of countries with the longest time series. Figure 1 shows, over the period 1750-2014, the evolution of lifetime entropy at birth for women and men in France, England and Wales, Iceland and Sweden.[12]

On each graph, the y axis measures life entropy at birth, expressed in bits units. Those measures can be interpreted as follows. A lifetime entropy at birth equal to 6 bits for French women in 1850 means that, based on the survival conditions prevailing in 1850 for French women, the expected amount of information learnt from the event of a death at an age $k \geq 0$ for a French women in 1850 was equal to the amount of information learnt from discovering the combination resulting from the flipping of 6 fair coins.

Figure 1 shows that lifetime entropy at birth H_0 exhibits, over the period considered, an inverted-U shape, with a maximum level during the first part of the 20th century. Over the period considered, lifetime entropy has first tended to increase over time (the event of a death becoming, on average, more astonishing), and, then, to decrease over time (the event of a death becoming, on average, less astonishing), with a turning point that varies across countries, but is generally located in the first half of the 20th century. Thus the risk about the duration of life has first increased during the 19th century and the first half of the 20th century, and, then, after 1950, has tended to decrease over time.

Note that this stylized fact is robust, in the sense that it is observed in all countries for which we have a sufficiently long time series. Moreover, as we show in the Appendix using moving averages, that stylized fact does not depend on a particular historical event (e.g. a war or an epidemic). On the contrary, the smoothed curves based on 20-year moving averages also exhibit an inverted-U pattern for all countries under study. This suggests that this pattern is a robust stylized fact, which is not driven by a particular historical event.

This stylized fact can be interpreted as follows. In the early 19th century, mortality is high. Hence, Wiener's entropy $W(d_i)$ associated to the event "death at age i (conditionally on survival to that age)" is low. Given that the event of a death is highly likely, one learns little from the event of a death. The terms $W(d_i)$ thus quantify what Ariès (1975) called the familiarity of individuals with death. It is also true that, at those times, survival to higher ages is surprising (implying higher Wiener terms $W(s_{i,0})$). Whether or not the high terms $W(s_{i,0})$ can counterbalance the low terms $W(d_i)$ within the index H_0 depends on the weights assigned to the various terms of H_0. Under high infant/childhood mortality, a large relative weight is given to the first few years of life, whereas a low weight is given to more distant ages.[13] Hence the low degree of surprise associated to the event "death at an early age" dominates in the calculus of H_0, leading to a low value for H_0. This measures the fact that, in the early

[12] It should be stressed that the stylized facts discussed in this section are observed for all countries for which we have long time series (see the Appendix for other countries).

[13] One may argue that, in preindustrial times, there was no lifecycle view, because of the high level of mortality, which shortened the time horizon. This is precisely what is captured by the index H_0, which computes the expected amount of information revealed by a death at age $i \geq 0$, while weighting each possible scenario by its probability of occurrence.

19th century, high infant/childhood mortality concentrated deaths during the childhood period.

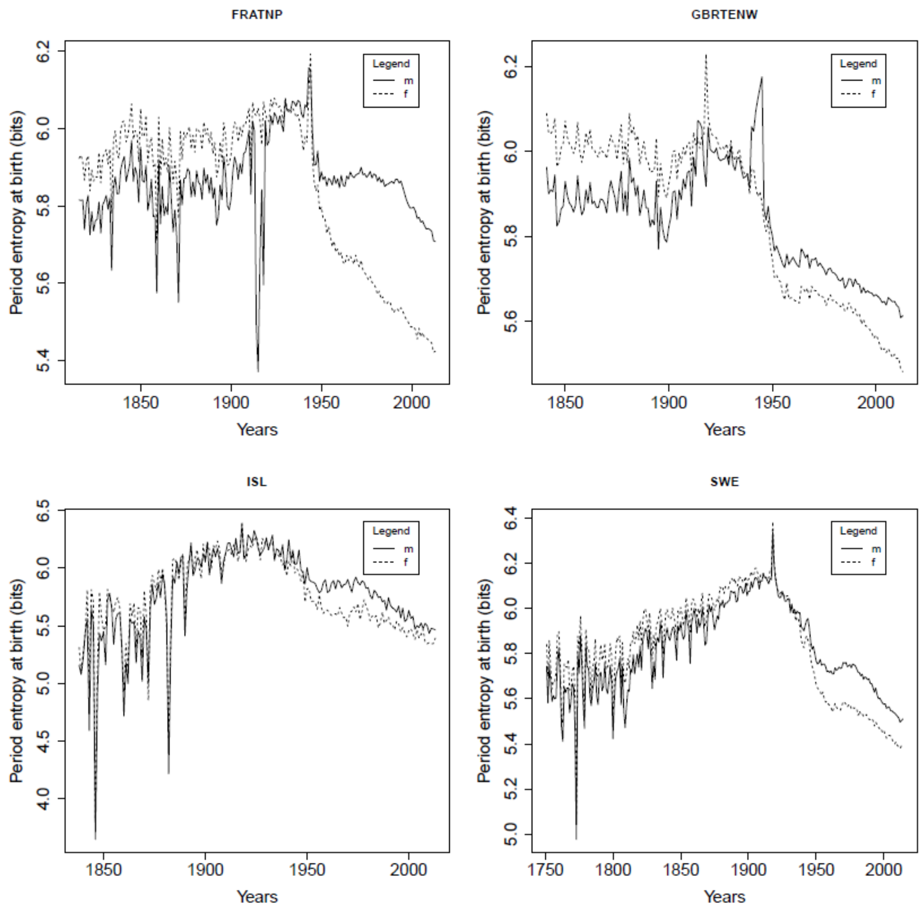

Figure 1: Lifetime entropy at birth (period) in France (FRATNP), England and Wales (GBRTENW), Iceland (ISL) and Sweden (SWE), for women (f) and men (m).

Then, as mortality goes down during the 19th century, two effects are at work. The reduction of mortality, which takes place mainly during childhood (with the exception of infant mortality) makes death more surprising, which increases the terms $W(d_i)$, but makes also survival to higher ages less surprising, which reduces the terms $W(s_{i,0})$. Given that infant mortality remains high, there is still a high weight, within the index H_0, assigned to what happens early in life. Hence the rise of the terms $W(d_i)$ for low ages i dominates, which pushes lifetime entropy up. The decrease of mortality during childhood made the age at death less concentrated, leading to a rise of lifetime entropy.[14]

During the 20th century, infant mortality has strongly fallen, leading to give more and more weight, within the index H_0, to what happens during adulthood. Since mortality reductions after 1950 concerned mainly old adulthood, there was a rise of Wiener terms $W(d_i)$ relative to ages above 60 (death at high ages becoming more astonishing), as well as a fall of Wiener terms $W(s_{i,0})$, survival to high ages becoming less astonishing. The net effect of this was a decline of lifetime entropy H_0. This decline of lifetime entropy reflects that, after 1950, the rectangularization of the survival curve has concentrated deaths around some high ages, with the consequence of making the event of a death reveal, on average, less information than before.[15]

This first stylized fact - the inverted-U shape of lifetime entropy at birth - is related to the fall of mortality that started in the early 19th century, which constitutes the first phase of the demographic transition (see Lee 2003 and Galor 2012). Regarding the underlying mechanisms behind the mortality decline (at different ages), various factors have been studied in the literature, such as improvements in the quantity and quality of food (Fogel 1994), public policies favouring hygiene and public health (Easterlin 1999), and the development and diffusion of medical research (de la Croix and Sommacal 2009). Quantifying the contributions of those factors to the mortality decline at different ages, and, hence, to the pattern of lifetime entropy, would be a quite complex task, which would go far beyond the scope of this paper.

Besides the inverted-U shape of lifetime entropy at birth, Figure 1 invites two other observations. A second observation consists of the variability of lifetime entropy at birth. Lifetime entropy exhibits very large fluctuations in the 19th century, and much smaller variability after 1900. This fact is well illustrated by the case of Iceland, which exhibits variations with magnitudes superior to 1.5 bit in the 1840s and 1880s, whereas fluctuations are of much smaller magnitudes in the 20th century. That stylized fact is due to the stronger variation of probabilities of death d_k before 1900, which make lifetime entropy fluctuate over time. The larger variability of survival conditions in more distant times (due to a larger vulnerability to bad harvests and to epidemics) has been widely studied by historians (see Wrigley 1969, Clark 2007).

A third observation concerns the comparison of lifetime entropy at birth across gender. In each country considered, lifetime entropy at birth is, at the beginning of the period under study, higher for women than for men, whereas, at the end of the period, lifetime entropy at birth is lower for women than for men. There has been an inversion of the gender rank in terms of lifetime entropy at birth. The reason why the curves of lifetime entropy for men and women diverge after 1950 lies in the fact that women experimented, in comparison to men, a

[14] Death being less concentrated at some ages, the event of each death reveals, on average, more information than before.

[15] It should be stressed, however, that the increase in the probabilities to reach higher and higher ages also affects the comparison of lifetime entropy, by adding, at ages above 90, extra terms with a relatively low $s_{i,k}$ leading to a relatively high $W(s_{i,k})$, and also a relatively high d_k, leading to low $W(d_k)$ at very high ages.

stronger rectangularization. This stronger rectangularization for women leads to a larger concentration of the age at death at some ages, with the consequence of a lower lifetime entropy.

Let us now focus on some large countries. Figure 2 shows, for women and men, lifetime entropy at birth in Japan, Russia and the U.S. during the 20th century. Two main observations can be made. First, whereas all countries under comparison exhibit quite similar levels of H_0 around 1960 (equal to about 6 bits for men and 5.9 bits for women), the pattern of lifetime entropy varies strongly across countries. While lifetime entropy at birth for women in Japan has declined by about 0.4 bit between 1960 and 2010, it has stagnated around 5.8 bits in Russia. Thus the evolution of life riskiness varies strongly across those countries. Second, although series for men and women in the same country are correlated, lifetime entropy is, over the 20th century lower for women than for men. Moreover, the decline in lifetime entropy differs across genders. In the case of Japan the decline was larger for women.

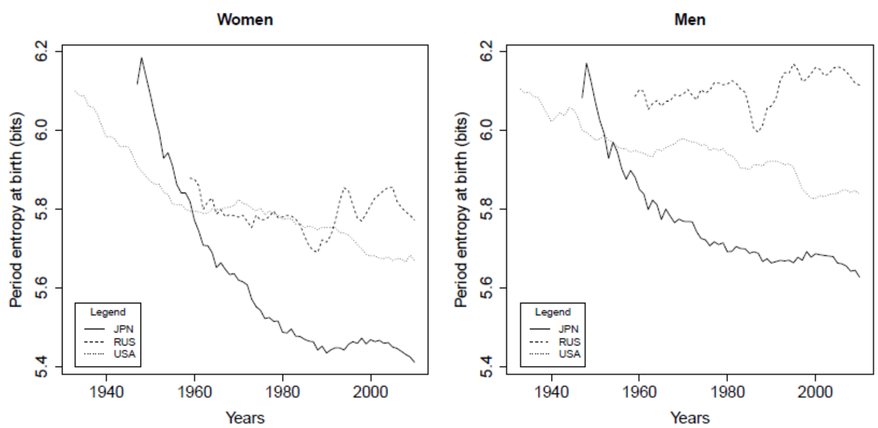

Figure 2. Lifetime entropy at birth (period) in Japan (JPN), Russia (RUS) and the U.S. (USA), for women (left) and men (right).

3.2 Lifetime entropy at birth: convergence or divergence?

Whereas the previous section highlighted a stylized fact concerning the general pattern of lifetime entropy H_0 during the last two centuries, it should be stressed that there is a large heterogeneity across countries, in terms of the magnitudes of the variations in lifetime entropy during that period. For instance, lifetime entropy in Iceland increased by about 1 bit between 1850 and 1900, whereas, over the same period, it grew by about 0.1 bit in France.

In order to examine whether countries converge or diverge in terms of lifetime entropy at birth, Figure 3 plots, for women (top) and men (bottom), the (period) lifetime entropy at birth in the countries for which we have observations starting in 1855 or before. We also show the associated standard deviation coefficient.

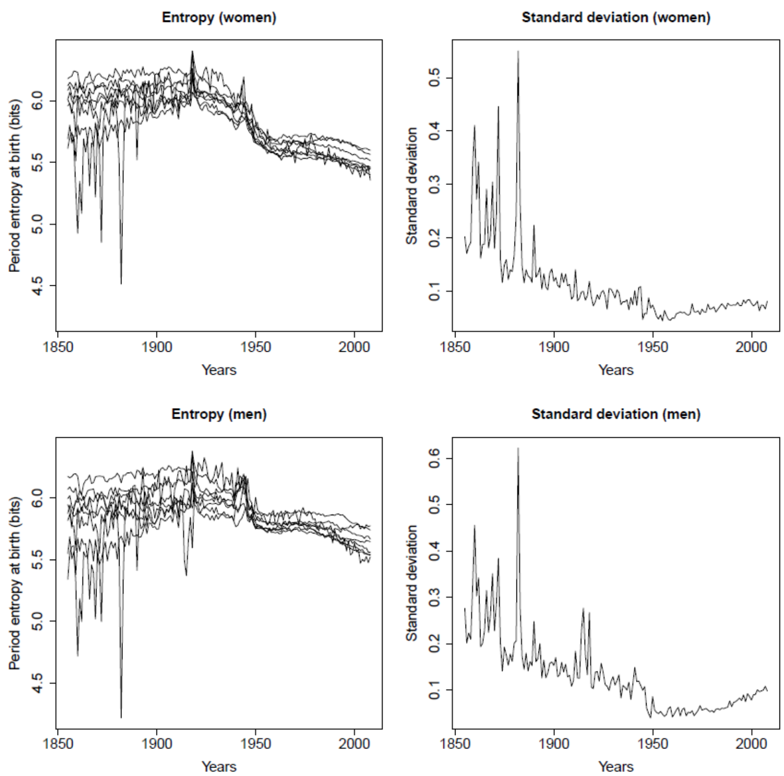

Figure 3: Lifetime entropy at birth (period) in Denmark, England and Wales, Finland, France, Iceland, Netherlands, Norway, Scotland, Sweden and Switzerland and associated standard deviation by year, for women (f) and men (m).

As shown on Figure 3, the countries under comparison exhibit quite different levels of lifetime entropy at birth H_0 in 1855, ranging between 5.3 bits and 6.3 bits for men, and between 5.6 bits and 6.3 bits for women. Moreover, there has been, over the period 1855-2008, a convergence of countries in terms of lifetime entropy at birth, with H_0 converging towards 5.6 bits for men and 5.5 bits for women. The extent of the convergence is measured by the standard deviation coefficient, which has, for men (resp. women), decreased from 0.28 (resp. 0.20) to 0.10 (resp. 0.10) between 1855 and 2008. This convergence, which arises for both men and women, constitutes another important stylized fact.

The large heterogeneity across countries in terms of lifetime entropy in the 19th century (left part of Figure 3), as well as the convergence that took place later on (right part of Figure 3), can be explained by considering the timing of the demographic transition across countries. All countries under study have gone through the first phase of the demographic transition (decline of mortality). However, not all countries entered the process of the demographic transition at the same time. The significant differences, across countries, in terms lifetime entropy in the 19th century (left part of Figure 3) reflect the different degrees of advancement of countries in the first phase of the demographic transition at that time. Then, there has been, over time, a convergence of countries in terms of survival conditions, which has contributed to a convergence in terms of lifetime entropy (right part of Figure 3). This explains why there is more cross-country heterogeneity on the increasing part of the inverted-U curve than on the decreasing part of the inverted-U curve.

Whereas the different timing of the demographic transition across countries - as well as the convergence of survival conditions during the 20th century - can explain the results shown on Figure 3, one should, however, keep in mind that there remain significant differences across countries. This is well illustrated by the case of Russia on Figure 2. Russia exhibits, due to particular historical circumstances, patterns for lifetime entropy that differ from those of other countries, both in terms of level and in terms of slope of H_0. Thus, even if some global convergence is at work, there remain sizeable differences across countries.

4 Lifetime entropy along the lifecycle

Up to now, we focused only on lifetime entropy measured at birth H_0, which measures the risk about the duration of life while taking the human lifespan as a whole. However, we can also compute life entropy for all other ages, and consider how life entropy evolves along the life cycle.

Focusing on life entropy at all ages of life allows us to examine how quickly the risk about the duration of life is resolved as a person becomes older. As studied in Section 2.3, it is not necessarily the case that survival from age k to age $k+1$ reduces the amount of risk about the duration of life that remains unresolved. This section proposes to reexamine, from a historical perspective, how the risk about the duration of life is progressively resolved with the age.

4.1 Lifetime entropy age profiles: the case of Sweden

In order to study the evolution of the lifetime entropy age profile over long periods of time, let us first focus on the case of Sweden. Figure 4 shows, for women and men, the lifetime entropy age profile for years 1751, 1839 and 2014. Each curve can be interpreted as follows: it gives, for each age k, the expected amount of information that is learnt in case of death at age $i \geq k$. The higher that amount of information is, and the higher is the degree of astonishment associated with the event of a death at age $i \geq k$.

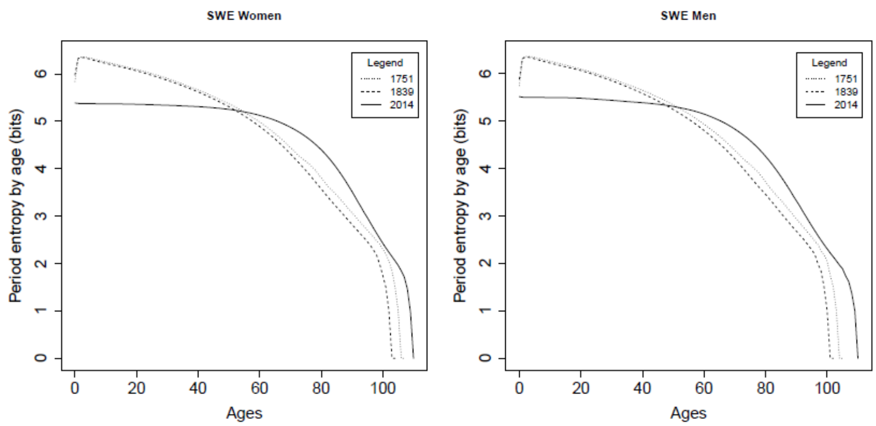

Figure 4: Period lifetime entropy by age in Sweden for women (left) and men (right).

Figure 4 can be interpreted as follows. Take, for instance, the lifetime entropy age profile for women in 2014. If one reads the graph from the x axis to the y axis, the entropy age profile tells us that, on the basis of the survival conditions prevailing for women in Sweden in 2014, the amount of risk about the duration of life is, at age 80, equal to about 4.5 bits. This means that, at age 80, the expected amount of information to be learnt from a death at ages beyond 80 years is equal to the amount of information revealed by discovering the combination of "heads" and "tails" resulting from tossing between 4 and 5 fair coins. Alternatively, one can read the graph from the y axis to the x axis, and look for critical ages at which the (expected) amount of information to be revealed about the age at death is equivalent to the amount of information associated to discovering the outcome from tossing a given number of fair coins. For instance, the age at which the (expected) amount of information to be revealed about the age at death is equivalent to 3 fair coins is 95 years.

Two main observations can be made. First, the overall shape of the lifetime entropy age profile - in particular its monotonicity - has changed over time. In the 18th and 19th centuries, lifetime entropy was, at the beginning of life, increasing with the age (i.e. $H_{k+1} - H_k > 0$), and then decreasing with the age (i.e. $H_{k+1} - H_k < 0$). On the contrary, in the 20th century, lifetime entropy decreases monotonically over the entire age interval (i.e. $H_{k+1} - H_k < 0$).

To explain that change, it should be reminded that, as discussed in Section 2.3, (remaining) lifetime entropy H_k does not necessarily go down in case of survival from age k to age $k+1$. Actually, as we showed above, it is possible that surviving from age k to age $k+1$ increases - rather than decreases - lifetime entropy, which coincides with a rise in the quantity of risk about the duration of life. Figure 4 shows that this case arises in early childhood during the 18th and 19th centuries. At that epoch, surviving the first few years of life was opening many possibilities in terms of life duration, and, hence, increased the amount of risk about the duration of life. The situation is very different nowadays, where infant mortality is much lower, so that lifetime entropy does no longer increase with the age during early childhood.[16]

A second observation concerns the slope of the lifetime entropy age profile during adulthood, that is, the speed at which risk about the duration of life is reduced as the individual becomes older. The lifetime entropy age profile becomes steeper as age goes up: the gap $\frac{|H_{k+1}-H_k|}{H_k}$ becomes increasingly large as age k increases. Thus the speed at which one learns about one's age at death goes up as the individual becomes older. At very high ages, a larger amount of risk about the duration of life is resolved as the person becomes one year older.

The intuition behind that result goes as follows. At high ages k, there is a higher risk of death d_k. This has two effects on H_k. First, the Wiener's terms $W(d_i)$ associated to the event "death at age i (conditionally on survival to age i)" go down, since death becomes, at high ages, less astonishing. Second, the Wiener's terms $W(s_{i,k})$ associated to the event "survival to age i" go up, since surviving to even higher ages becomes more surprising at very old ages. While those two effects go in opposite directions, the fact that mortality becomes larger at higher ages implies that the weights within the H_k formula are high for ages close to k, and become increasingly small for more distant ages. For ages i close to k, we have that $W(s_{i,k})$ is small (survival to close ages is not surprising), and since d_i becomes larger, $W(d_i)$ becomes smaller. This leads to a (remaining) lifetime entropy index H_k that is low, and which falls as age k goes up. At extremely high ages k, the number of different possible durations of life is small, so that deaths are concentrated at some ages, leading to low (remaining) lifetime entropy H_k.

4.2 Lifetime entropy age profile across countries

Let us now compare lifetime entropy age profiles across countries and gender. For that purpose, Figure 5 shows, for men and women, lifetime entropy profiles by age for Australia, France, England and Wales, Japan, Russia and the U.S., for the first and the last observations. Note that, here gain, when considering the first observation, the lifetime entropy profile is not monotonically decreasing with the age, but exhibits a small rise at very young ages. This coincides with the phenomenon discussed in section 4.1.

[16] That first observation, which illustrates the discussion in Section 2.3, can be made for all countries in the sample for which we have a sufficiently long time series.

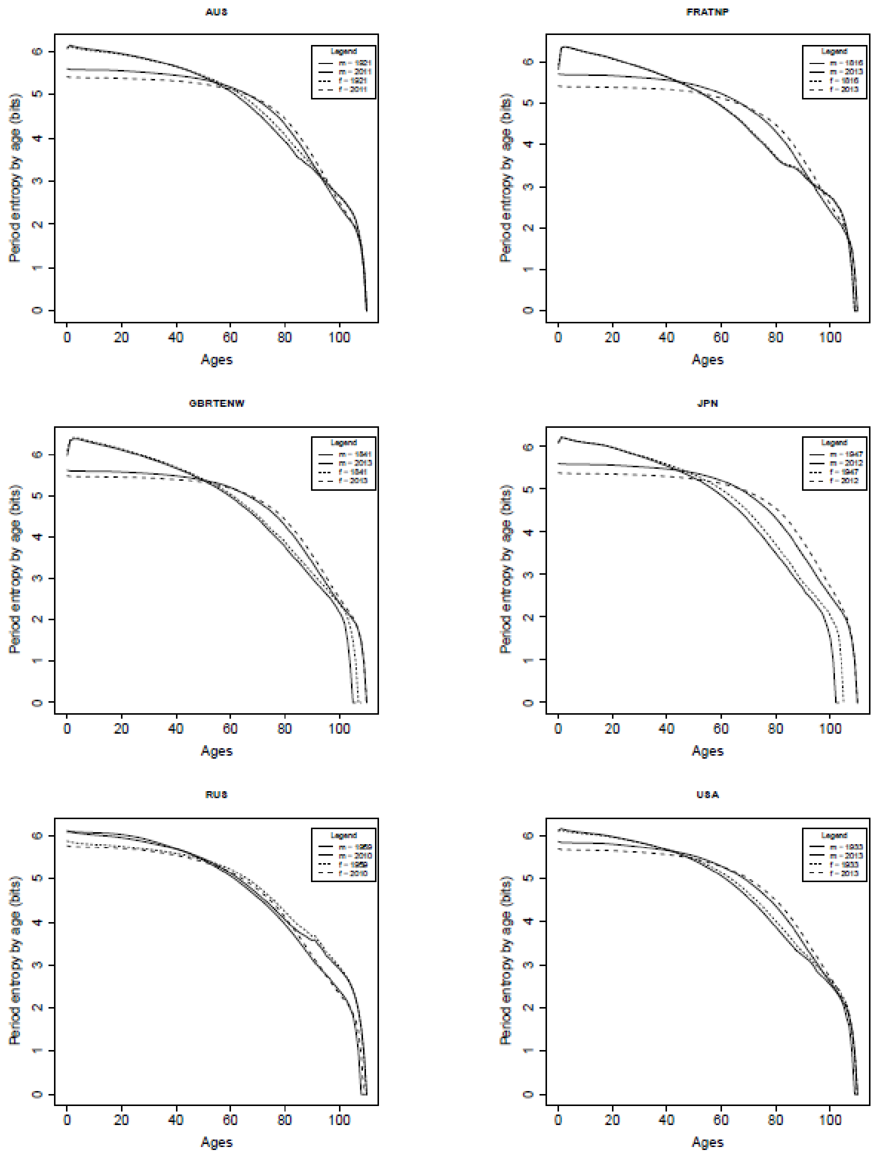

Figure 5: Period lifetime entropy by age in Australia (AUS), France (FRATNP), England and Wales (GBRTENW), Japan (JPN), Russia (RUS) and the US (USA) for women (f) and men (m), first and last available year.

Focusing on the last observation under study (early 21st century), the comparison of lifetime entropy age profiles across countries, gender and period allows us to identify two stylized facts.

First, whatever the country, the gender and the period considered (abstracting from the short increasing part of the entropy age profile mentioned above), all entropy age profiles exhibit, in the early 21st century, a slope that is, in absolute value, increasing with the age. That feature is a robust stylized fact, which can be observed for each country and gender under study (beyond the 6 countries shown here).

At the beginning of life (i.e. the young age), the entropy age profile is quasi flat, and very little information about the duration of life is learnt when the individual becomes older (i.e. $H_{k+1} \approx H_k$). Then, beyond age 55, one starts learning, year after year, more about one's duration of life. But the learning process remains quite slow: the amount of risk about one's age at death falls from about 5 bits around age 55 to 2 bits at age 100. Finally, at the very old age ($>$ 100 years), there remains little risk about the duration life (about 2 bits), and the learning about the duration of one's life is extremely rapid: the remaining 2 bits of entropy vanish in just a few years. The intuition is simple: at those very high ages, there remains a very small number of possible durations of life. Hence, the information revealed by a death at a particular age is then extremely small, and as small as the information revealed by discovering the combination of "heads" and "tails" obtained through drawing 2 fair coins.

Another robust stylized fact concerns the comparison of lifetime entropy profiles by gender. As shown in Figure 5, the lifetime entropy age profiles for men and women intersect around age 50. Clearly, men exhibit, in the first half of life, a higher life entropy than women. Then, in the second half of life, this is the opposite, and women exhibit a higher life entropy than men.

5 Cohort lifetime entropy

Up to now, our analysis of lifetime entropy has relied only on period life tables. As a consequence, our lifetime entropy index H_k for a given year t did not capture the actual risk about the duration of life, but the expected risk about the duration of life, while assuming that age-specific mortality rates of year t will keep on prevailing during the entire lifetime of the individual. That assumption is strong, especially over the period considered, during which there was a strong fall of mortality.

As a complement, this section proposes to study lifetime entropy on a cohort basis (i.e. based on cohort life tables). Cohort lifetime entropy at birth measures the average amount of information that was revealed by the death of a cohort member at age $k \geq 0$. As such, this index measures how surprising or astonishing the death of a cohort member was.

In comparison with the period lifetime cohort index, the cohort lifetime entropy index must be interpreted in a different way. Whereas the former provides a synthetic measure of the risk about the duration of life in a given year (based on the period life table of that year), the latter provides a synthetic measure of the average risk about the duration of life within a cohort born in a particular year, and whose members died at many different ages.

5.1 Cohort lifetime entropy at birth

Figure 6 shows, for men, the patterns of cohort lifetime entropy at birth in France, England and Wales, Iceland and Sweden, and compares these patterns with the ones of period lifetime entropy at birth studied above.[17]

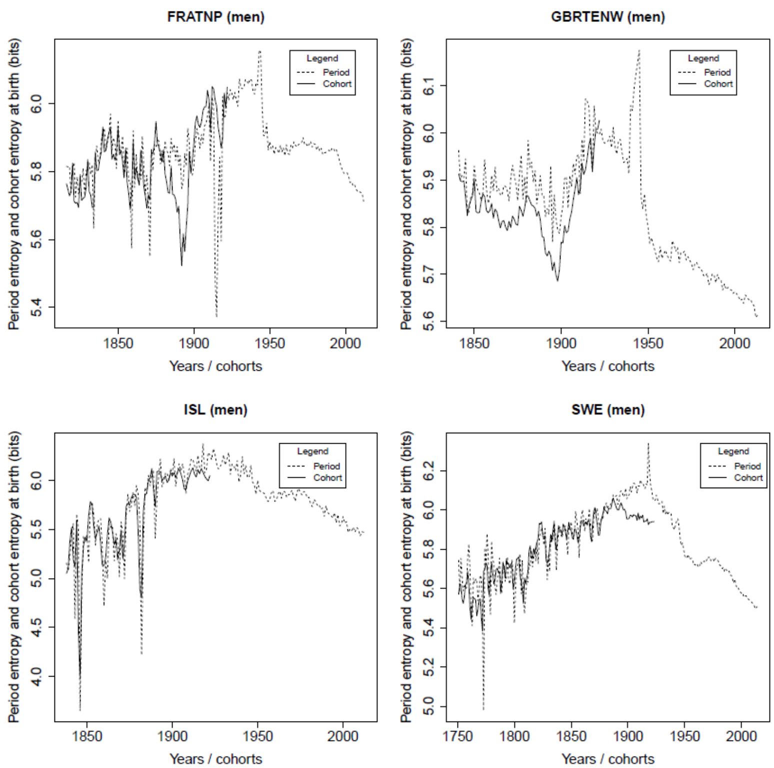

Figure 6: Period and cohort lifetime entropy at birth in France (FRANTNP), England and Wales (GBRTENW), Iceland (ISL) and Sweden (SWE), for men.

A first observation is that the relation between period and cohort lifetime entropies at birth varies strongly across periods and countries, reflecting the different timing of shifts in survival curves across time and space. For instance, we can see that, for the cohorts of men born in England and Wales between 1845 and 1900, period lifetime entropy at birth has systematically exceeded cohort lifetime entropy at birth. The members of those cohorts faced, during their life, a lower lifetime entropy than the one that could have been expected based

[17]A similar comparison is carried out for women in the Appendix. See also the Appendix on cohort lifetime entropy at birth in other countries.

on the yearly age-specific probabilities of death. On the contrary, for cohorts of men born in Iceland between 1845 and 1900, cohort lifetime entropy is not systematically below period lifetime entropy.

The comparison of period and cohort lifetime entropies can also be used to quantify the impact of World War One on the riskiness of life. Period lifetime entropy at birth for French males fell strongly at the beginning of the War, making lifetime entropy at birth drop from about 6 bits to 5.4 bits. Turning now to cohort analysis, Figure 6 shows that the impact of the War is not concentrated on just a few cohorts, but, rather, is spread on many cohorts, born between 1875 and 1895. Cohorts of French males born in the 4th quarter of the 19th century are characterized by lifetime entropy levels that fall strongly across cohorts, from 5.9 bits for men born in 1875 to about 5.5 bits for men born in 1895. Thus Figure 6 shows how different were the lives of successive cohorts born in just two decades, depending on their age at the beginning of the War.

Regarding the issue of convergence or divergence of countries in terms of cohort lifetime entropy at birth, it is shown in the Appendix that there has been an unambiguous convergence in terms of cohort lifetime entropy at birth for cohorts born between 1855 and 1920. Note that this convergence of lifetime entropy across countries is achieved not by variations across countries in terms of the growth rate of lifetime entropy, but by the mere fact that, while some countries with a higher initial entropy have exhibited a decrease in lifetime entropy, other countries, with lower initial entropy levels, have exhibited a rise in lifetime entropy over the cohorts considered.

5.2 Cohort lifetime entropy along the life cycle

Let us now examine how cohort lifetime entropy has evolved with the age of cohort members, that is, how the risk about the duration of lifetime is progressively resolved as one becomes older. Figure 7 shows cohort lifetime entropy along the lifecycle for men and women in France, England and Wales, Iceland and Sweden, for the first and the last cohorts under study in each country.

On the basis of Figure 7, three main observations can be made. First, in each country, we can see that a major change in the form of entropy age profiles concerns the very beginning of life. For cohorts born in the 18th or 19th century, infant mortality was high, and being able to survive the first few years of life was opening new possibilities in terms of life duration, and, hence, increasing (remaining) lifetime entropy H_k. This rise in (remaining) lifetime entropy at low ages is no longer observed among cohorts born in the 20th century.

A second observation concerns the overall shape of the entropy age profiles by cohorts between ages 40 and 100. For those ages, cohort entropy age profiles have tended to shift upwards over time. The overall shape of the entropy age profile between ages 40 and 100 looks pretty much the same across cohorts, but the level of cohort entropy has changed. We can see that, above age 40, cohorts born in the early 20th century faced a higher lifetime entropy at all ages in comparison to cohorts born in the 18th or 19th century. The gap varies with the age, but is equal on average to about 0.5 bits.

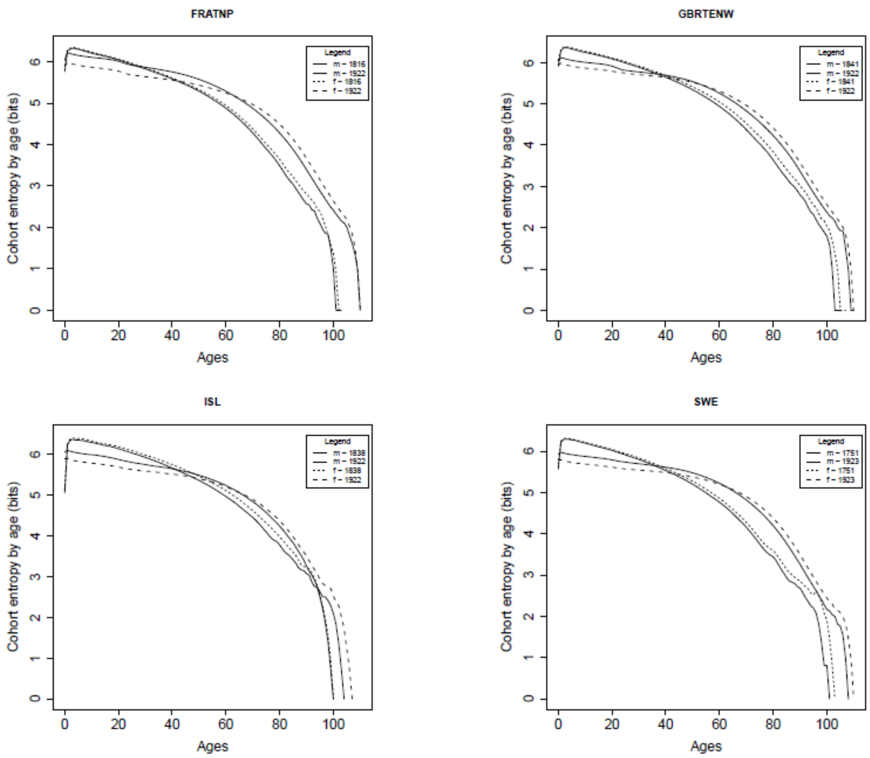

Figure 7: Cohort lifetime entropy by age in France (FRATNP), England and Wales (GBRTENW), Iceland (ISL) and Sweden (SWE), for women (f) and men (m), first and last available years.

A third observation concerns the very high ages of life. For cohorts born in the 18th or 19th century, lifetime entropy beyond age 100 was pretty low, and falling strongly to zero bits (risk fully resolved).[18] However, for cohorts born in the early 20th century, lifetime entropy at age 100 remains equal to about 2 bits. Thus, for high ages, the lifetime entropy age profile has tended to shift to the right, allowing cohorts to face a risk about the duration of their life at ages when, in the past, little risk remained. This major change can be explained by the shift of survival curves to the right, which opened new perspectives at very high ages, and, also increased risk about the duration of life at those ages.

[18]This can be explained by the fact that cohort life tables for those cohorts exhibited probabilities to reach ages beyond 95 years that were often close to zero or equal to zero when the life table was closed at a maximum age equal to 95. Note that the fact that life tables were closed at lower ages is not a pure measurement artifact: it reflects that surviving beyond those ages was, at those distant epochs, extremely rare.

6 Conclusions

Life is inherently risky: no one knows how long one will live. This paper proposed an intuitive measure of the risk about the duration of life: Shannon's lifetime entropy index to the base 2. Thanks to its measurement unit (the bit), that index makes the risk about the duration of life commensurable with the risk involved in the most common randomization device: tossing a fair coin.

On the theoretical side, we showed that Shannon's lifetime entropy index at age k can be decomposed into a weighted sum, across all future ages $i \geq k$, of Wiener's entropy terms relative to two events, which are, on the one hand, "death at age i conditionally on survival to age i", and, on the other hand, "survival to age i". This decomposition allows us to interpret Shannon's lifetime entropy index H_k as the mathematical expectation of the amount of information learnt by the event of a death at a future age $i \geq k$, information that can be decomposed in the two components $W(d_i)$ and $W(s_{i,k})$. This decomposition allowed us also to study how the risk about the duration of life is being resolved as the individual becomes older. Contrary to what one may believe at first glance, survival from age k to age $k+1$ does not necessarily resolve risk about the duration of life. Actually, if the probability of death at age k is larger than the probabilities of death at higher ages, it can be the case that survival to a higher age raises H_k.

On the empirical side, we identified 6 stylized facts using the Human Mortality Database: (1) over the last two centuries, (period) lifetime entropy at birth exhibits an inverted-U shape pattern with a maximum in the first half of the 20th century (at about 6 bits); (2) curves of period lifetime entropy at birth for men and women crossed during the 20th century; (3) over the last 150 years, Western countries have converged in terms of (period) lifetime entropy at birth towards 5.6 bits for men and 5.5 bits for women; (4) the entropy age profile has shifted from a non-monotonic profile (in the 18th and 19th centuries) to a strictly decreasing profile (in the 20th and 21st centuries); (5) men exhibit a higher lifetime entropy than women below ages 50-55, and a lower life entropy after ages 50-55; (6) the entropy age profile becomes steeper at higher ages, coinciding with a rise in the speed of learning about one's duration of life.

The main contribution of this paper was to develop a measure of the risk about the duration of life that relies on an intuitive metric. Developing such intuitive measures of life riskiness is necessary for at least two distinct reasons. First, as shown by Edwards (2013), the risk about the duration of life is costly (and is even larger when individuals do not annuitize). Hence it is important that citizens have a precise idea of that risk. Second, lack of knowledge, at the individual level, of the risk about the age of death may partly explain some puzzles in the economics of risk and insurance, such as the long term care (LTC) private insurance puzzle (Brown and Finkelstein 2007, 2011).[19] Several causes

[19] That puzzle goes as follows. Despite the large probability (about 30 to 50 percents) to enter a nursing home at the old age (Brown and Finkelstein 2009), and despite the large costs of LTC, few people purchase private LTC insurance. According to Brown et al (2007), only 9 to 10 percents of the population at risk of LTC purchase a private insurance in the U.S.

may explain this paradox, both on the supply side (high loading factors) and on the demand side (crowding out by family care or by social insurance). But lack of quantification of life riskiness - and, hence, of dependency preceding death - may also explain the LTC private insurance puzzle. From that perspective, developing more intuitive measures of life riskiness matters.

Having stressed this, the present study suffers nonetheless from some limitations, which invite further research. A first limitation concerns the intuitive character of our measure in the presence of cognitive biases. One may argue that even if individuals are familiar with the informational content revealed by tossing *one* fair coin, they may still have difficulties in understanding how large is the amount of information revealed by tossing $n > 1$ fair coins, because of cognitive biases. The construction of a more intuitive measure of life riskiness, which would be more deeply rooted in the human mind, remains thus on the research agenda.

Another important question concerns the historical conditions determining the patterns of life riskiness across centuries. The goal of this study was purely descriptive. Examining the causal mechanisms at work behind the stylized facts identified here would go far beyond the scope of this paper. However, it is important to study how the factors at work in the decline in mortality are related to the dynamics of lifetime entropy. Moreover, the perception of life riskiness may have varied depending of those factors, which, in turn, may also have been affected by mental attitudes towards death (see Ariès 1975). This motivates future studies focusing on interactions between human attitudes towards death, the perception/measurement of life riskiness, and the determinants of mortality.

Taken together, those two extensions would, on the theoretical side, require to depart from the model studied in this paper, to consider a more general framework, which would include the inputs in the perception of lifetime riskiness, the inputs in the production of mortality, and, also, the relations between those different inputs. This ambitious program is left for future research.

In sum, this paper is only a first step towards more intuitive measures of the risk about the duration of life, which would really allow humans to understand the extent of that risk, and how it is being resolved as they become older. It is crucial to develop indicators that allow individuals to have a more concrete idea of the amount of risk about the duration of life. Those indicators will also allow humans to make (more) rational decisions in the context of a risky lifetime.

7 References

Aczél, J, Daroczy, Z. (1975). *On Measures of Information and their Characterizations*. Mathematics in Science and Engineering, Elsevier, volume 115.

Ariès, P. (1975). *Essais sur l'histoire de la mort en occident*. Editions du Seuil, Paris.

Brouwer, W., van Exel, N. (2005). Expectations regarding length and health related quality of life: some empirical findings. *Social Science and Medicine*, 61(5), 1083-1094.

Brown, J., Coe, N., Finkelstein, A. (2007). Medicaid crowd-out of private long-term care insurance demand: evidence from the Health and Retirement Survey. NBER Chapters, in:

Tax Policy and the Economy, volume 21, 1-34, National Bureau of Economic Research.

Brown, J., Finkelstein, A. (2007). Why is the market for long-term care insurance so small? *Journal of Public Economics*, 91 (10), 1967-1991.

Brown, J., Finkelstein, A. (2009). The private market for long term care in the U.S.. A review of the evidence. *Journal of Risk and Insurance*, 76(1): 5-29.

Brown, J., Finkelstein, A. (2011). Insuring long term care in the United States. *Journal of Economic Perspectives*, 25 (4), 119-142.

Clark, G. (2007). *A Farewell to Alms. A Brief Economic History of the World.* Princeton University Press.

de la Croix, D., Sommacal, A. (2009). A theory of medical effectiveness, differential mortality, income inequality and growth for pre-industrial England. *Mathematical Population Studies*, 16, 1-34.

Demetrius, L. (1976). Measures of variability in age-structured populations. *Journal of Theoretical Biology*, 63, 397-404.

Dormont, B, Samsonz, A-L., Fleurbaey, M., Luchini, S., Schokkaert, E., Thebaut, C., Van de Voorde, C. (2018). Individual uncertainty on longevity. *Demography*, 55, 1829-1854.

Easterlin, R. (1999). How beneficient is the market? A look at the modern history of mortality. *European Review of Economic History*, 3, 257-294.

Edwards, R., Tuljapurkar, S. (2005). Inequality in life spans and a new perspective on mortality convergence across industrialized countries. *Population and Development Review*, 31 (4), 645-674.

Edwards, R. (2013). The cost of uncertain life span. *Journal of Population Economics*, 26 (4), 1485-1522.

Fogel, R. (1994). Economic growth, population theory and physiology: the bearing of long-term processes on the making of economic policy. *American Economic Review*, 84 (3), 369-395.

Galor, O. (2012). The demographic transition: causes and consequences. *Cliometrica*, 6 (1), 1-28.

Gigerenzer, G., Hertwig, R., Van Den Broek, E., Fasolo, B., Katsikopoulos, K. (2005). A 30 percent chance of rain tomorrow: how does the public understand probabilistic weather forecasts? *Risk Analysis*, 25 (3), 623-629.

Godelier, M. (2018). *La mort et ses au-delà.* Editions du CNRS, Paris.

Hamermesh, D. (1985). Expectations, life expectancy, and economic behavior. *Quarterly Journal of Economics*, 100 (2), 389-408.

Hamermesh, D., Hamermesh, F. (1983). Does perception of life expectancy reflect health knowledge? *American Journal of Public Health*, 73, 911-914.

Hill, G. (1993). The entropy of the survival curve: an alternative measure, *Canadian Studies in Population*, 20 (1), 43-57.

Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org

Jankélévitch, W. (1977). *La mort.* Flammarion, Paris.

Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research*, 3 (6).

Keyfitz, N. (1977). *Applied Mathematical Demography.* New-York: John Wiley and Sons.

Khinchin, A.I. (1957). *Mathematical Foundations of Information Theory.* Dover (edition 2017), London.

Lan Karen Cheung, S., Robine, J.M. (2007). Increase in common longevity and the compression of mortality: The case of Japan. *Population Studies*, 61 (1), 85-97.

Lan Karen Cheung, S., Robine, J.M., Jow-Ching Tu, E., Caselli, G. (2005). Three dimensions of the survival curve: horizontalization, verticalization and longevity extension. *Demography*, 42 (2), 243-258.

Lanciani, R. (1892). Gambling and cheating in Ancient Rome. *North American Review*, 155, No. 428, 97-105.

Lee, R. (2003). The demographic transition: three centuries of fundamental change. *Journal of Economic Perspectives*, 17 (4), 167-190.

Mirowsky, J. (1999). Subjective life expectancy in the US: correspondence to actuarial estimates by age, sex and race. *Social Science and Medicine,* 49(7), 967-979.

Nagnur, D. (1986). Rectangularization of the survival curve and entropy: the canadian experience, 1921-1981. *Canadian Studies in Population*, 13 (1), 83-102.

Noymer, A., Coleman, C. (2014). A universal pattern of the evolution of life table entropy and life expectancy. In *Extended Abstracts*, 1-7, Boston. Population Association of America

Pierce, J. (1980). *An Introduction to Information Theory. Symbols, Signals and Noise.* Dover, London.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27 (3), 379-423.

Smits, J., Monden, C. (2009). Length of life inequality around the globe. *Social Science and Medicine*, 68 (6), 1114-1123.

Tuljapurkar, S., Edwards, R. (2011). Variance in death and its implications for modeling and forecasting mortality. *Demographic Research*, 24, 497.

Tversky, A., Kahneman, D. (1975). Judgement under uncertainty: heuristics and biases. In: *Utility, Probability and Human Decision Making*, 141-162, Springer.

Van Raalte, A., Caswell, H. (2013). Perturbation analysis of indices of lifespan variability. *Demography*, 50(5), 1615-1640.

Vaupel, J., Zhang, Z., van Raalte, A. (2011). Life expectancy and disparity: an international comparison of life table data. *BMJ open*, 1(1), e000128.

Wiener, N. (1965). *Cybernetics or Control and Communication in the Animal and the Machine.* MIT Press.

Wilmoth, J., Horiuchi, S. (1999). Rectangularization revisited: variability of age at death within human populations. *Demography*, 36 (4), 475-495.

Wrigley, E.A. (1969). *History and Population.* Weidenfeld & Nicolson, London.

8 Appendix

8.1 Related entropy indicators

Among indicators of verticalization of survival curves, entropy indexes have also been widely used to measure the variance of the age at death.

The entropy indicator that is the closest to the one studied in this paper is the lifetime entropy index developed by Hill (1993), and which was applied to measure lifetime entropy in Canada in the 20th century. Hill's (1993) entropy of the age at death is defined as:

$$HI_k = -\sum_{i=k}^{115} p_{i,k} \ln(p_{i,k}) \tag{9}$$

The unique difference between Shannon's lifetime entropy index H_k and the HI_k index lies in the fact that HI_k is based on the natural logarithm, and thus consists of an entropy index defined to the base e, whereas Shannon's entropy index relies on the logarithm defined to the base 2. One can thus rewrite Shannon's lifetime entropy index in terms of Hill's entropy index as follows:

$$H_k = HI_k \frac{\sum_{i=k}^{115} p_{i,k} \log_e (p_{i,k}) \log_2(e)}{\sum_{i=k}^{115} p_{i,k} \log_e (p_{i,k})} = HI_k \log_2(e) \tag{10}$$

This formula shows that Shannon's entropy index defined to the base 2 can be rewritten in terms of Hill's entropy index, while rescaling the levels in such a way as to quantify lifetime entropy in terms of bits. This rescaling suggests that, if one focuses only on the growth patterns of lifetime entropy over time, the obtained growth rates are invariant to the measure unit chosen, so that measuring lifetime entropy by means of H_k or HI_k does not make a difference.

However, focusing on the growth patterns of lifetime entropy is not the only possible use of that measure; one may want to be able to interpret the *levels* of the measure of life riskiness at a particular point in time or at a particular age. From that perspective, adopting the base 2 rather than the base e allows to quantify risk about the duration of life in a metric that makes life riskiness commensurable with the risk involved in tossing a fair coin. On the contrary, the measurement unit of the HI_k index is less easy to connect with common situations of risk. Moreover, relying on the base 2 rather than on the base e allows us also to decompose the lifetime entropy index in terms of Wiener's entropy of the single event "death at age k conditionally on survival to age k". Wiener's entropy being defined to the base 2, it is more natural, when constructing a lifetime measure of the risk about the duration of life, to rely also on the base 2, without any re-normalization.

Having compared our indicator with Hill's entropy indicator, it is also useful to compare Shannon's lifetime entropy index with Keyfitz's entropy, which is the most widely used entropy index in demography (see Keyfitz 1977, Demetrius 1976, Nagnur 1986, Noymer and Coleman 2014). Keyfitz's entropy is:

$$K_k = -\frac{\sum_{i=k}^{115} s_{i,k} \log (s_{i,k})}{\sum_{i=k}^{115} s_{i,k}} \tag{11}$$

The denominator of Keyfitz's entropy is merely the life expectancy at age k, whereas the numerator aggregates, along the entire lifecycle, the logarithmic transform of the unconditional survival probabilities to the different ages of life.

It should be stressed that Keyfitz's life table entropy or population entropy is not, mathematically speaking, an entropy index. It measures the reactivity or elasticity of life expectancy to a proportional change of the strength of age-specific mortality. As such, it differs substantially from the lifetime entropy index that we propose in this paper. Actually, Shannon's lifetime entropy index can be rewritten in terms of Keyfitz's entropy as:

$$H_k = K_k L_k \frac{\sum_{i=k}^{115} (s_{i,k} d_i) \log_2 (s_{i,k} d_i)}{\sum_{i=k}^{115} s_{i,k} \log (s_{i,k})} \tag{12}$$

where L_k is the life expectancy at age k. Given the important difference between Shannon's lifetime entropy and Keyfitz's population entropy, it is not surprising that these two indicators exhibit quite different patterns over time.

To illustrate those differences, Figure A1 below compares, for the period life tables for France (males), the pattern of Shannon's lifetime entropy index to the base 2 (left scale) with Keyfitz's entropy index (right scale). The comparison of those two curves is made difficult by the fact that these indicators have completely different scales. Whereas the Keyfitz entropy index relies, over the period, between 0.15 and 0.6, the Shannon entropy index varies between 5.4 and 6.2. Having stressed this difficulty, one can nonetheless make some observations regarding those two patterns. First, if we focus on the second part of the 20th century, the two entropy indexes exhibit both a declining trend. However, if we focus on the period between 1890 and the Second World War, we can see that the two indicators show completely different patterns. Whereas Keyfitz's entropy exhibits a declining trend, Shannon's entropy exhibits an increasing trend. Thus the two indicators exhibit quite different patterns, which is not surprising given that these indicators measure two different things.

Figure A1. Shannon's lifetime entropy and Keyfitz's entropy, for France, males, 1816-2014.

8.2 Additional material

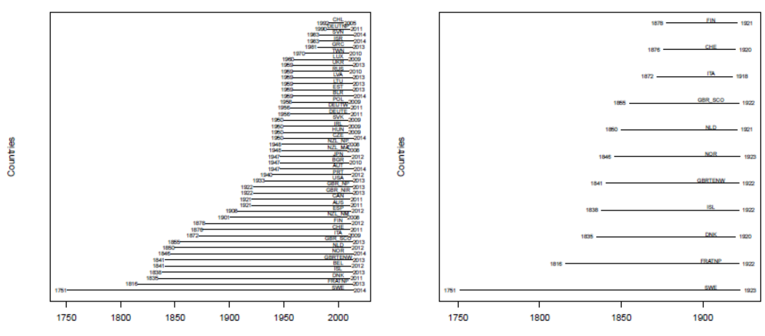

Figure A2: Ranges of period lifetables (left) and cohort lifetables (right) per country in the Human Mortality Database

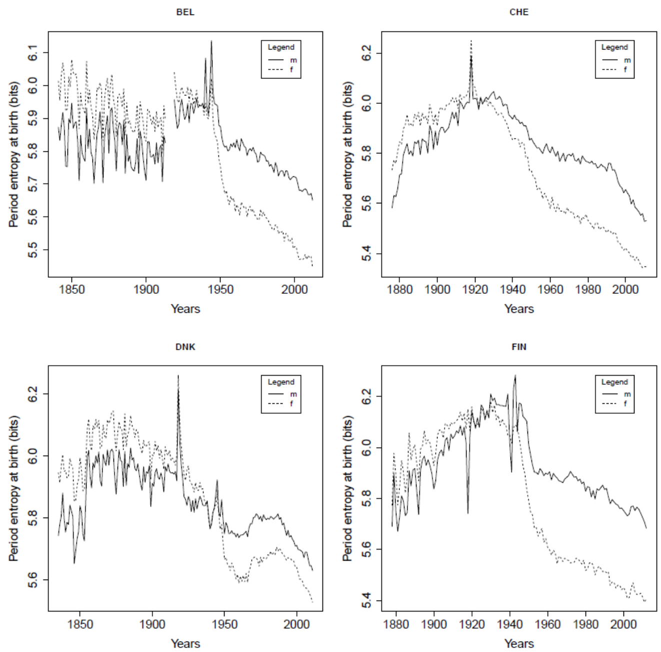

Figure A3: Lifetime entropy at birth (period) in Belgium (BEL), Switzerland (CHE), Denmark (DNK) and Finland (FIN), for women (f) and men (m).

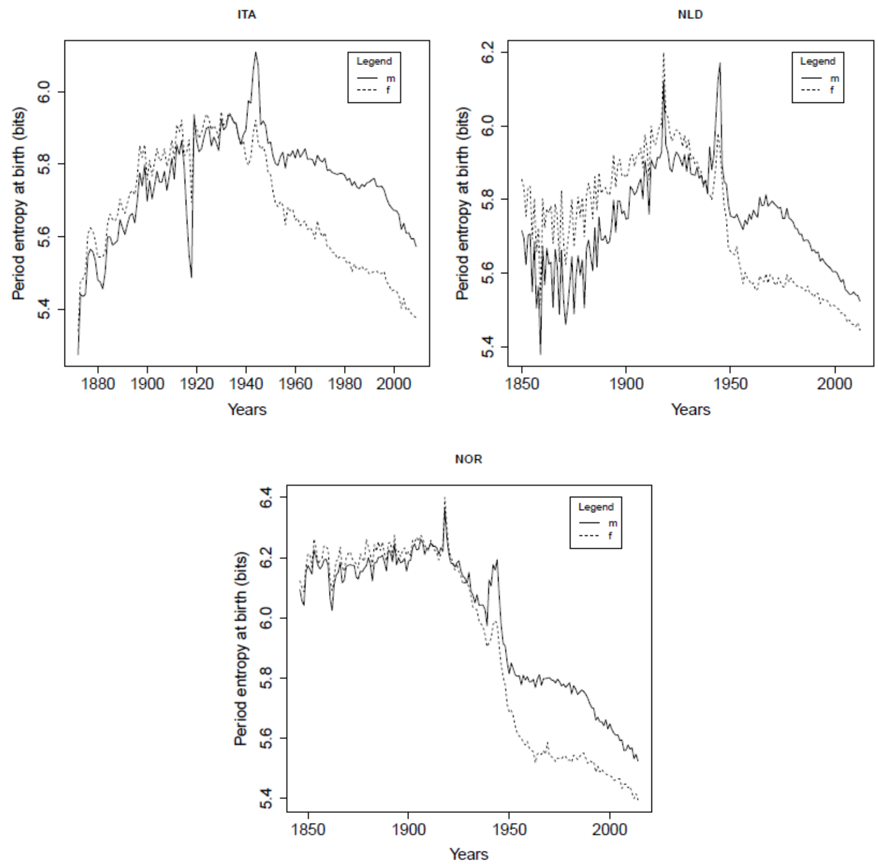

Figure A4: Lifetime entropy at birth (period) in Italy (ITA), Netherlands (NLD) and Norway (NOR), for women (f) and men (m).

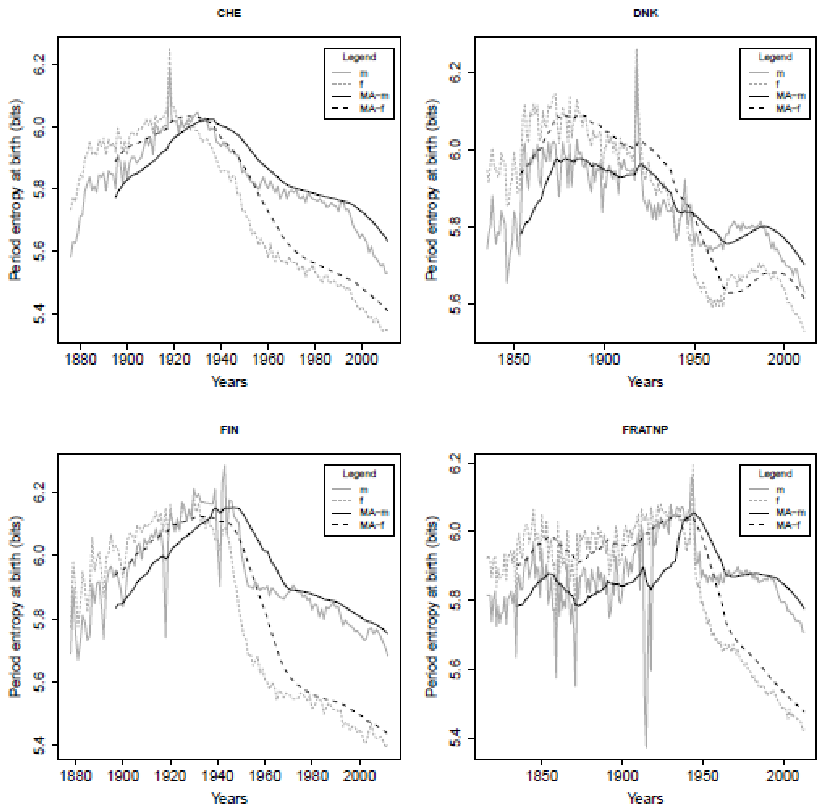

Figure A5: Life entropy at birth (period) in Switzerland (CHE), Denmark (DNK), Finland (FIN) and France (FRATNP), for women (f) and men (m) and corresponding moving averages (arithmetic mean, 20 years).

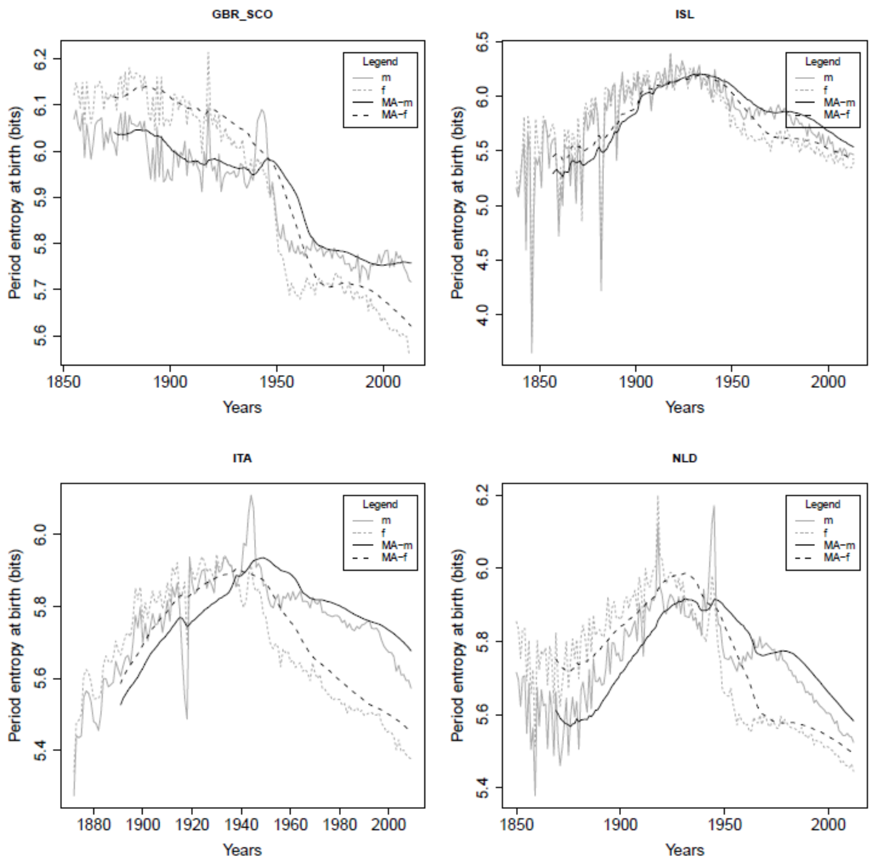

Figure A6: Life entropy at birth (period) in Great Britain and Scotland (GBR SCO), Iceland (ISL), Italy (ITA) and The Netherlands (NLD), for women (f) and men (m) and corresponding moving averages (arithmetic mean, 20 years).

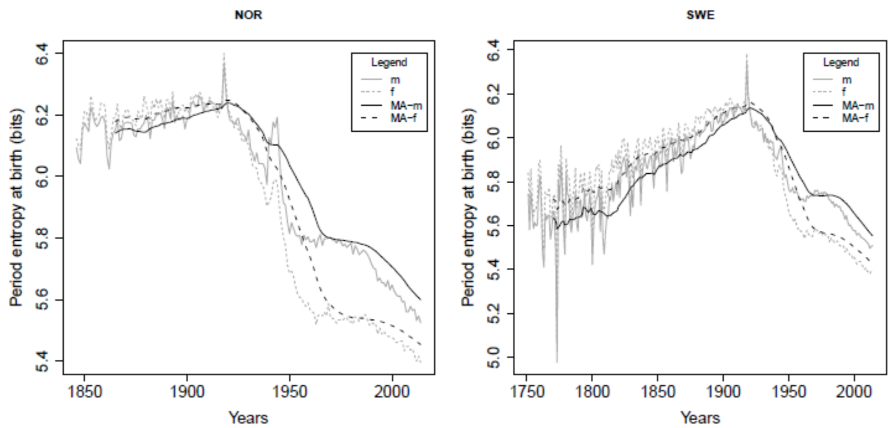

Figure A7: Life entropy at birth (period) in Norway (NOR) and Sweden (SWE), for women (f) and men (m) and corresponding moving averages (arithmetic mean, 20 years).

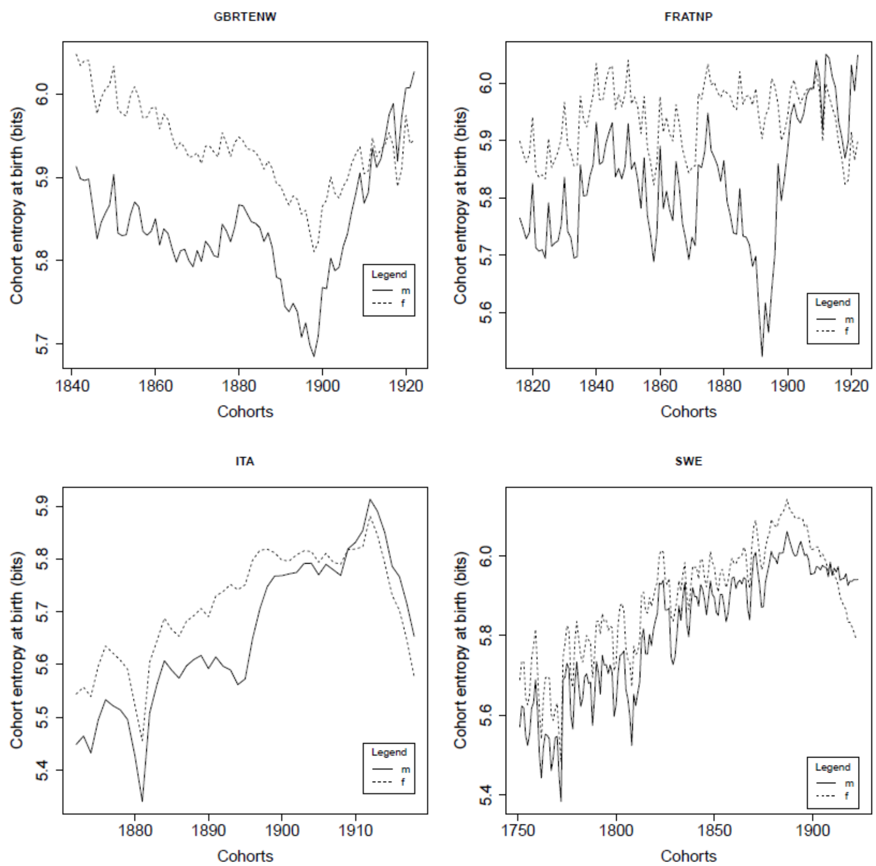

Figure A8: Lifetime entropy at birth (cohort) in England and Wales (GBRTENW), France (FRATNP), Italy (ITA) and Sweden (SWE), for women (f) and men (m).

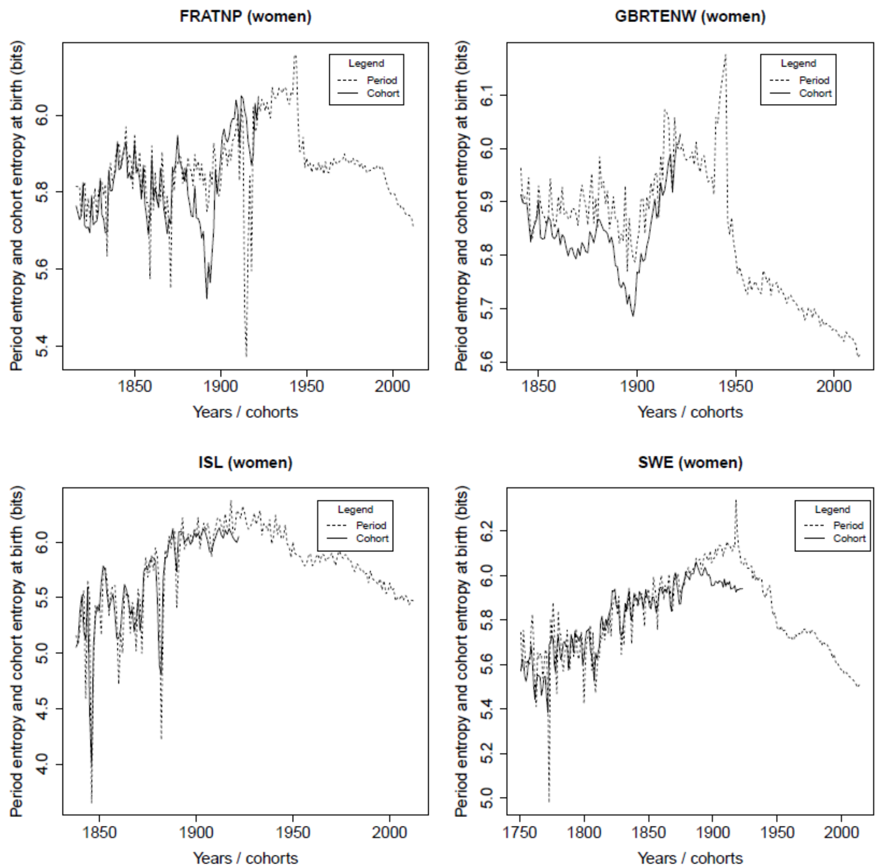

Figure A9: Period and cohort lifetime entropy at birth in France (FRATNP), England and Wales (GBRTENW), Iceland (ISL) and Sweden (SWE) for women.

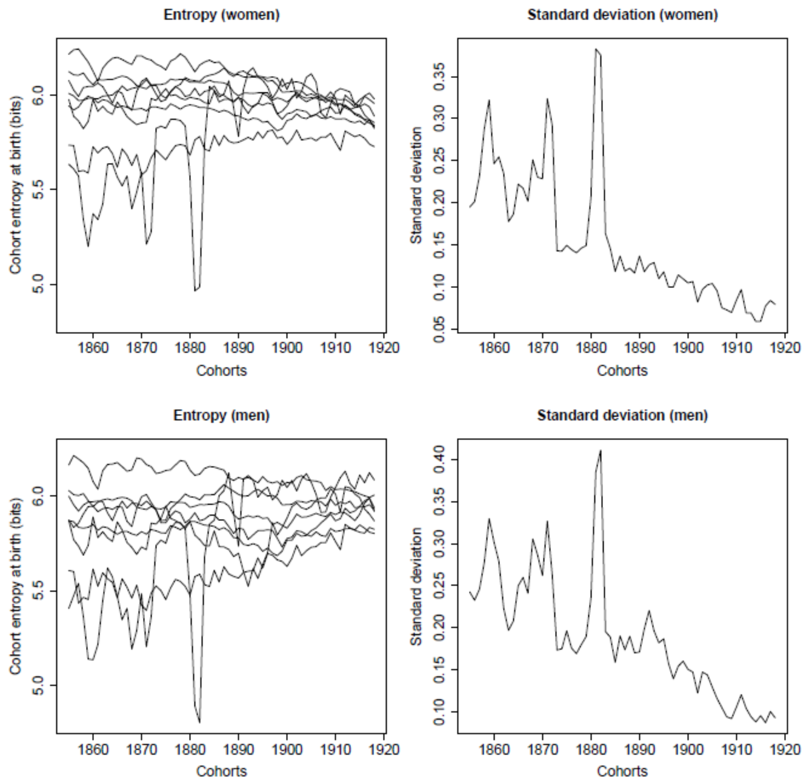

Figure A10: Lifetime entropy at birth (cohort) in Denmark, England and Wales, Finland, France, Iceland, Italy, Netherlands, Norway, Scotland, Sweden and Switzerland and associated standard deviation by year, for women (f) and men (m).

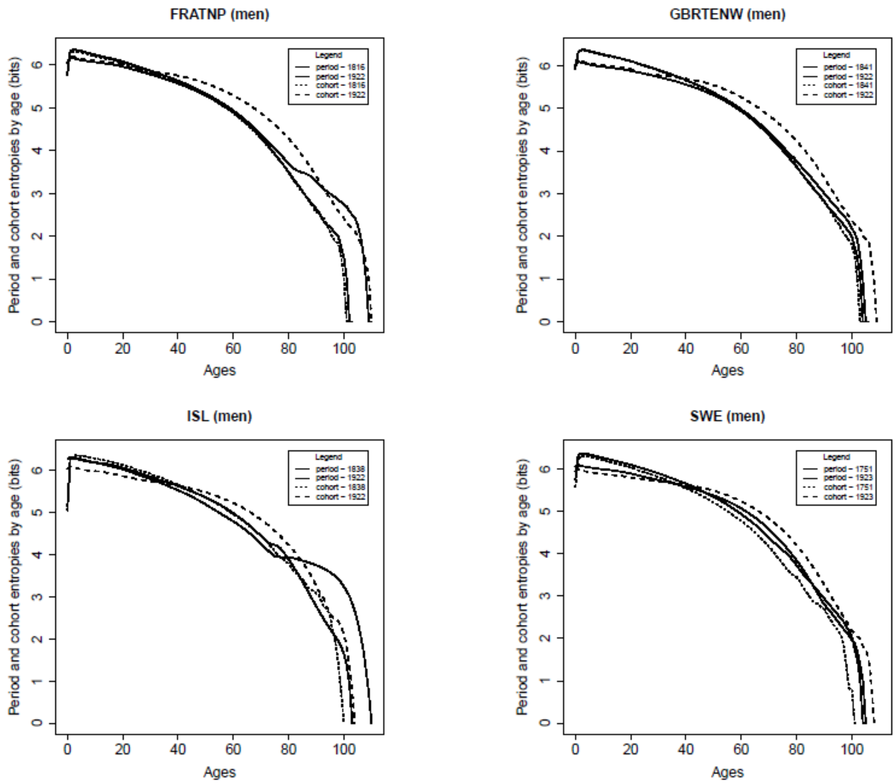

Figure A11: Cohort versus period life entropy along the life cycle in France (FRATNP), England and Wales (GBRTENW), Iceland (ISL) and Sweden (SWE) for men, first and last available years.