# Supplementary material

# Loss of Resilience in Aging Societies

## Contents

## 1. The data

Spreadsheets with the selection of data used in this study are available as separate supplementary files to the main article on the PNAS website. What follows is some background on the two data bases MOROS and Seshat, and the ways we processed the data. We plan to publish an extended and extensively documented version of the MOROS database separately.

### The Mortality of States (MOROS) Dataset

The Mortality of States Index (MOROS) provides an overview of the lifespan of different political states. It documents commonly agreed state formation and end dates for over 440 different states, covering approximately 5,000 years from 3100 BCE (Egyptian Dynasties I and II) to 2021. We define the state as a set of centralized institutions that coercively extract resources from, and impose rules on, a territorially circumscribed population. This is a necessarily broad definition. There was significant variety in how pre-modern states governed, as well as their level of administration, centralization, and coercion. It is not a surgically precise term: statehood exists on a spectrum and is not a binary (1).

The idea of the state is not without detractors. There are critiques including the sheer diversity of states, and that the idea of a state is ultimately a false projection of a modern political form onto ancient case studies (2). Nonetheless, there is general consensus across political science and archaeology that, despite significant variety, states provide a real and useful political category. While there are difficulties in drawing precise boundaries it is a commonly used category in studies of history and politics (3).

The dates for both state beginnings and ends should be seen as approximate, pivotal dates in which significant changes to state form, function, and/or sovereignty occurred. They are often indicative of processes that may have taken many years to unfold. For instance, 1177 is used as rough dating for the collapse of the Bronze Age state network, even though the process unfurled over decades (4). The potential for inaccurate dating increases further back in history due to less reliable documentary and archaeological data.

The entries have been gathered from a range of different materials. The initial, primary sources were:

- Three different surveys of historical empires and large polities by Taagepera (5-7);
- The Seshat Database (8, 9);
- The four volume 2016 Encyclopedia of Empire (10) and;
- The Correlates of War Project (11).

We used these primary sources to generate the first dataset. We compared entries to eliminate duplicates. Seshat, encompasses not only states, but broader cultural periods that are distinguished by changes in material culture. Hence, entries from Seshat were only included when they represented a distinct and established state.

Each of the primary sources focuses on overlapping, similar units, although with differing definitions. These are summarized below in Table S1. Most sources lack a distinct measurement of statehood, and hence a guide to coding for state formation, continuity, and termination. The sole exception is the Correlates of War Project which uses both political recognition from great powers and population size as proxies for state sovereignty. These are of less relevance since the Correlates of War data was not used in our analysis of pre-modern states. The arbitrary population threshold is also not appropriate for pre-modern states with often significantly lower populations and recognition by neighbors would inappropriately exclude many ancient (especially 'pristine') states.

***Table S1: Source Material Definitions***

| **Source** | **Unit of Analysis** | **Definitions** |
|---|---|---|
| Seshat | Polities and quasi-polities | Polities are defined as independent political units. These range from small scale villages through to empires.[16] |
| 2016 Encyclopedia of Empire | Empires | An expansionist polity composed of a ruling center and dominated periphery which looks to create forms of sovereignty over peoples of a people different to its own.[12] |
| Taagepera | Empires and large polities | Empires are large (25,000km2 or greater) sovereign political entities comprised of components which are not sovereign.[2] |
| Correlates of War | States | Membership in the League of Nations or UN, or a population of at least half a million, and recognition of sovereignty by two major powers (via diplomatic missions).[17] |

We then drew on a wider literature search to both validate the majority of the existing entries and to create an additional 22 profiles. Most of these were from specialty sources for Chinese dynasties (12) and Korean kingdoms (13) which were less reliably covered by the primary sources. We also consulted books focused on societal collapse for additional entries, although these were either already covered, or not suitable (14, 15).

There is often a range of different estimates for the beginning and end of a state. Where we have found competing suggestions, we input both the lowest and highest credible estimates.

During this construction phase we excluded 30 polities. Entries were excluded for one of two reasons: a) it was unclear whether the polity would qualify as a state, and/or b) the formation and termination dates were highly uncertain (spanning decades) and/or contested. These are summarized in Table S2. We expect that many of these could be clarified and included in a future version of MOROS. For now, we have taken a precautious approach.

***Table S2: Excluded Polities***

| **Polity** | **Justification for Exclusion** |
|---|---|
| Benin Edo | A pre-colonial state in modern Nigeria that appears to have evolved from a city-state into an empire over several centuries (16). While the generally accepted termination date is 1897, it has been excluded due to uncertainty over state formation dates. |
| Brunei Sultanate | The Brunei Sultanate appears to a state, even an empire. However, it's history is primarily recounted through European and Chinese sources, making dating and any understanding of continuity difficult (17, 18). |
| Chavin | Appears to be a cultural unit rather than a single continuous state. |
| Da Viet | While it evolved into a centralized state, it can be seen as a series of polities with different fragmentation periods (including a conquest period by the Ming dynasty). This poses significant problems with dating(19). |
| Elam | Elam appears to be better characterized as a cultural unit rather than a single continuous state. It experienced numerous periods of fragmentation and centralization and the dates are highly uncertain due to the sparsity of data. |
| Harrapan (Urban Indus Civilization) | Currently lacks indications of a centralized and hierarchical political structure (20). |
| Hurrian Kingdoms of Urkesh | Appears to be several competing kingdoms rather than a single state, with highly uncertain formation and termination dates. |
| Jenne Jeno | It does not appear to have the political centralization to qualify as a state and was a single urban area rather than a territorial state (21). |
| Kanem Bornu | A centralized state and unified empire that many see as having lasted over a millennium (22). However, it appears to have had two separate imperial periods, and dating was deemed highly |

| | |
|---|---|
| | uncertain due to a lack of peer-review estimates and divergent dates. |
| Maya | Not a single state, but rather a cultural zone containing multiple city-states. |
| Minoan Crete | There is significant contestation over whether Minoan Crete was characterized by a unified state or not. |
| Moche | There is significant contestation over whether the Moche was characterized by a unified state or not. |
| Mutapa (Kingdom of) | Two different state periods with a lack of dates from peer-reviewed research, especially for the state formation of the first state. |
| Ndebele | Unclear as to whether it constitutes a state, and unclear formation and termination dates. |
| Ngoni | There is a lack of clarity over whether this constitutes a state. It was a marching conquest polity with a constantly moving territory and little formalized or specialized administration. |
| Papal States | It does appear to be a state (theocracy) with a form of taxation over multiple territories, albeit with highly decentralized governance. However, it is a border case and given multiple periods of schism (including times when territories fell more under the control of Roman nobility than the papacy) we have decided to exclude it. |
| Rapa Nui | The Rapa Nui appears to lack the centralization and institutionalized authority to be considered a state. Moreover, there is significant dispute over the dating of its demographic collapse(23). |
| Republic of Pisa | Despite fluctuating territorial extent, it appears to be better characterized as a city-state during much of its duration. |
| Shona | Highly uncertain dates and appears to represent a cultural unit rather than a state. |
| Srivijaya Empire | It was a state and empire based on control of trade routes rather than land (24). However, most of the evidence comes from a few inscriptions and references from other polities making dating od state formation and termination difficult and highly uncertain. |
| Teutonic Order | A state with fluctuating territories. However, dates in both the literature and online sources are highly divergent and uncertain. |

| | |
|---|---|
| Tui-Tonga | Widespread debate as to whether this constitutes a state or empire. Some characterize this as a 'chiefdom' rather than a centralized state. This is largely due to its leaders relying on their own charisma and personality rather than the institutionalized authority of the office (25). |
| Tukolor Empire | It is frequently referred to as a 'quasi-state' and a rulership over a set of garrisons (26). While there are some compelling arguments for it to be considered a state (27), we decided to exclude it due to the expert disagreement and uncertainty. |
| Venetian Empire | While it is a state (a thalassocracy and oligarchic republic), it was often a city-state with fluctuating territorial extent (28). Hence, we decided to classify it as a city-state for now and exclude it. |
| Vishnukundina Dynasty | A state presiding over parts of South India. Dating is uncertain with few estimates, especially in peer-review literature. |
| Wahabi Empire | Appears to be more of a religious movement rather than an empire and has highly uncertain formation and termination dates. |
| Western Satrap | Unclear as to whether the Western Satraps were vassals to the Kushan Empire, and hence not an independent state (24). |
| Xianbei | Highly uncertain dates and status. |
| Yap Empire | Widespread debate as to whether this constitutes a state or empire. Some characterize this as a 'chiefdom' rather than a centralized state. This is largely due to its leaders relying on their own charisma and personality rather than the institutionalized authority of the office (25). |
| Zapotec | The existence of a Zapotec empire (although it appears to have been a state at some point) is debated and dating highly uncertain (29). |

We applied four criteria to assess statehood:

- The presence of a state apparatus that was formally (and even legally) capable of imposing rules and preventing fission.
- Institutionalized authority that could enact the functions of the state without relying on the charisma and personal attributes of the ruler.
- Continuous rule over a territory extending beyond a single city.
- The level of expert (dis)agreement.

For instance, we exclude the Tui-Tonga and Yap Empires since it appears that despite some degree of political centralization and a lack of territorial fission, the authority of leadership hinged on individual attributes, rather than the formal authority of the office. Our approach is conservative. Where there was disagreement or significant uncertainty over state status, continuity, or dating, we tended towards exclusion. This does pose a bias against areas with less well-documented states. such as Africa. Many empires in pre-colonial Africa are documented primarily from outside sources, and colonial empires had a vested interest to see these as 'quasi-polities' (27). Nonetheless, the uncertainty over dating and continuity poses a challenge, one which we hope to address in the future with expert elicitation.

In some cases, there were inconsistencies between sources over whether to merge or split certain polities. For instance, Taagepera (5) codes the Fatimid, Ayyubid, and Mamluk as a single polity, while the Encyclopedia of Empire and other sources do not (indeed, the Encyclopedia of Empire entry contests that the Ayyubid are not an empire, while the Mamluk sultanate is). In these cases, we have adopted what appears to be the most widely used periodization and, in general, opted for disaggregation over merging. The following polities were all merged in the original Taagepera dataset and split in MOROS: Muscovy-Russia-USSR (divided into Tsardom of Muscovy, Russian Empire, and USSR), Fatmid-Ayyubid-Mamluk (split into Fatimid, Ayyubid, and Mamluk) and Almoravid-Almohad (divided into Almoravid and Almohad).

The dates in MOROS do not represent any particular quantitative thresholds. Instead, they represent rough agreement by experts as to when a state can be said to have existed and ended based on interpretation of an array of sources and factors. It is a qualitative overview of common expert opinion on political periodization. This poses problems. Different experts, and different fields can implicitly deploy varying interpretations of what signifies the end of a polity or lineage. This is difficult to detect since experts frequently do not explicitly define state formation and termination. Nonetheless, this approach remains the best proxy for state formation and termination. In the future, it would be useful to craft a precise definition of state formation and termination and use expert elicitation (or even expert crowdsourcing) to screen entries.

States have been grouped according to whether they are an empire, kingdom, confederacy, or nation-state. This is accompanied by a secondary classification based on whether there are elected rulers (a republic), monarchy (kingdom), a khan (khaganate), or by a religious caliph. Table S3 provides an overview of the definitions for these different state forms. As with the formation and termination estimates, this rough and preliminary classification is based on how the state entry is typically labelled in the corresponding literature.

Note that the dates provided in MOROS say little about the exact nature of the state formation and end. An empire in the dataset could have undergone a full collapse of political, economic, and societal institutions, or just undergone a fundamental change in political form (such as the movement of Rome from Republic to Empire). It also covers a simpler change in ruling elites, such as dynastic shifts in China that were incurred by internal warlords or coups (which we have identified and marked within MOROS). We hope to use expert elicitation and systematic literature reviews in the future to provide deeper information on the exact details of each entry, including what the termination entailed, a stricter definition of state formation and termination, the purported causes for collapse/transformation, and the evidence underpinning different theories.

MOROS is a work-in-progress. Further work is needed to ensure the estimates, are robust, comparable, and provide appropriate depth in analysis. It is not entirely comprehensive of either all states throughout human history or for all types of polities. It excludes city-states, non-state polities, and cultural units such as 'civilizations'. Nonetheless, it is to the best of our knowledge the largest dataset of state lifespans in existence.

**Table S3: Definitions of the Different Political Forms Used in MOROS**

| ***State Type*** | ***Explanation*** |
|---|---|
| *Empire* | *Empires are large polities composed of previously sovereign territories and states who are formally and informally controlled through coercion by a core state to extract value (7, 30).* |
| *Confederacy* | *Large polities composed of consenting, previously sovereign territories, and states.* |
| *Khaganate* | *A political body ruled by a Khan, most frequently an empire or confederacy.* |
| *Caliphate* | *An Islamic state with a caliph as head of state.* |
| *Kingdom* | *A monarchical rule in which the state is ruled by a king or queen with successions usually determined by kinship.* |
| *Republic* | *A state in which rulers are elected and their power is derived from the wider body of citizens. This is often accompanied by an oligarchic or aristocratic form of rule and the use of elections for successions.* |

## Use of the SESHAT database

This research employed data from the Seshat Databank (seshatdatabank.info) under Creative Commons Attribution Non-Commercial (CC By-NC SA) licensing (8, 9, 31, 32). We used the Equinox-2020 Dataset to quantify the polity longevity as the difference between the 'Date.From' and 'Date.To' variables in the dataset. For the polities with multiple sub-units characterized by multiple geographic areas and/or time periods, we combined the sub-units (with same values of the 'PolId' variable) and quantified the overall longevity using the earliest start dates and latest end dates of sub-units. Importantly, in the Seshat database, long-lived societies have been split into several distinct units with typical durations of ~200 years. We also combined these deliberately split societies to quantify the overall longevity. In the database, such entries can be identified through checking the PolId with suffix of '1', '2', '3', or 'E', 'M', 'L', or 'A', 'B', 'C', which have consecutive duration periods. For our analyses we exclude the category of "quasi-polities". This construction in Seshat takes together groups of small-scale polities (e.g., independent villages, or small chiefdoms) that were not feasible to code individually but co-occur in a geographic area with some degree of cultural homogeneity.

## 2. Methods for deducing hazard functions from longevity distributions

To see how the development of risk with age should be expected to shape the probability distribution of longevity, consider an imaginary large group of independent societies that are all born at the same moment. The risk of each of the societies to collapse at a given moment (for example, in a given year) depends on their age. Thus, a fraction of the societies is terminated at each timestep. The total number of societies terminated at each timestep depends on this fraction, but also on how many societies have survived so-far. Plotting this absolute number of terminations against the time when they happened produces a curve that corresponds to the longevity distribution. The approach we take is to first derive analytically which longevity distributions correspond to different hazard functions (describing how risk of termination changes with age). This then allows us to find the parametrized hazard functions by fitting the longevity distributions.

### Analytic derivation of longevity distributions from hazard functions

Survival analysis is a branch of statistics for analyzing the duration till a particular event, such as the collapse of a society. Here we briefly summarize how we can deduct a statistical distribution from a parameterized risk (or hazard) function, for more details we refer among others  to (33).

The hazard function (h(t)) describes the instant risk of collapse as a function of the current age (t). We tested different parametric functions for this.

The following differential equation describes the survival of a population:

$$\frac{dx}{dt} = -h(t)\,x$$

If we start this differential equation with an initial condition of 1, the solution describes the survival function. The analytic solution is:

$$S(t) = e^{-\int_0^t h(y)\,dy} = e^{-H(t)}$$

Where H(t) is the cumulative hazard function, which is the integral of the hazard function from zero to t.

The survival function is directly related to the cumulative distribution function F:

$$F(t) = 1 - S(t) = 1 - e^{-H(t)}$$

The probability density function is the derivative of the cumulative distribution:

$$f(t) = h(t)\,S(t) = h(t)\,e^{-H(t)}$$

So if the hazard function can be integrated, we can easily derive both the equations that define the probability density functions. Table S4 shows the hazard functions that we tested.

***Table S4. The equations of the hazard functions and survival functions of the distributions that we tested. The names of the distributions refer to the shape of the hazard functions (in our parameter settings) and is sometimes different than the common names. As far as we know, the saturating risk distribution is not described before (though the shape resembles the gamma distribution).***

| Name | Hazard function h(t) | Survival function S(t) |
| --- | --- | --- |
| Constant risk (also known as exponential) | $\frac{1}{\mu}$ | $e^{-t/\mu}$ |
| Linear increasing risk (cf. Rayleigh distribution) | $a\,t$ | $e^{-\frac{a}{2}t^2}$ |
| Linear increasing risk with intercept | $(a\,t + r)$ | $e^{-\frac{a}{2}t^2 - r\,t}$ |
| Saturating risk | $c\frac{t}{t+h}$ | $e^{-(c\ t - c\,h\,ln(h+t) + c\,h\,ln(h))}$ |
| Saturating risk with intercept | $(c\frac{t}{t+h} + r)$ | $e^{-(c\ t - c\,h\,ln(h+t) + c\,h\,ln(h) + r\,t)}$ |
| Humped risk (also known as lognormal) | $\frac{1}{\sigma t}\,\frac{\phi\left(\frac{\log(t)-\mu}{\sigma}\right)}{1-\Phi\left(\frac{\log(t)-\mu}{\sigma}\right)}$ | $1-\Phi\left(\frac{\log(t)-\mu}{\sigma}\right)$<br>Where $\Phi$ is the cdf and $\phi$ the pdf of the Gaussian distribution |
| Exponentially increasing risk (also known as Gomperz) | $\alpha\,e^{\beta\,t}$ | $e^{-\frac{(\alpha\,e^{\beta t})}{\beta}+\frac{\alpha}{\beta}}$ |
| Exponentially increasing risk with intercept (also known as Gomperz-Makeham) | $(\alpha\,e^{\beta\,t} + \lambda)$ | $e^{-\frac{(\alpha\,e^{\beta t})}{\beta}+\frac{\alpha}{\beta}-\lambda\,t}$ |

## Why a humped longevity distribution implies rising risk before the mode

At the mode the derivative of the pdf switches sign from positive to negative.

The derivative of the pdf (S(t)*h(t)=exp(-H(t))*h(t)) is:

$$S(t)\left(\frac{\partial h(t)}{\partial t} - h(t)^2\right)$$

Note that the survival function S(t) and squared hazard $h(t)^2$ are always positive (or zero).

Thus, the derivative is positive only if $\frac{\partial h(t)}{\partial t} > h(t)^2$

This implies that risk should be increasing in the period before the mode.

After the mode $\frac{\partial h(t)}{\partial t} < h(t)^2$ implying that h(t) may still increase, but could also become constant or decrease. This analysis also illustrates why a pdf with a constant hazard $\frac{\partial h(t)}{\partial t} = 0$ is always monotonically decreasing.

## Fitting distributions to data

We fitted all these distributions to the longevity data of each of the data sets. We used an iterative maximum likelihood procedure (`mle` in MATLAB) to fit the parameters to these distributions, except for Constant Risk (exponential distribution) and Humped Risk (lognormal distribution) where faster methods are available. We used the Akaike Information Criterion (AIC) to select the best fitting distribution. This model-selection criterion is based on the log-likelihood with a penalty for the number of fitted parameters. The minimum AIC value thus selects for the most parsimonious model with a good fit.

For the Moros data set four states that were formed before 1800 still exist till today. In survival analysis such incomplete data are called right-censored (33, 34). We used the common practice to use cumulative distributions to account for their contribution to the log-likelihood function(33).

The uncertainty in our predictions were assessed using a parametric bootstrap(33). In this procedure we use the fitted distributions to generate 1000 random data sets of the same size as the original data sets. The distributions were then fitted again to each of these bootstrapped data sets. We present in figures the 0025 and 0.975 quantiles as an estimation of the uncertainty due to fitting.

To judge the fit of the parametric models, we also compare how the survival functions predicted by the fitted distributions relates to a survival curve obtained in a non-parametric way directly from the data using the Kaplan Meier approach (34). The confidence ranges around the non-parametric distributions are determined using Greenwood's formula (33) using the MATLAB function `ecdf`.

Our developed MATLAB software is available at gitlab (https://git.wur.nl/sparcs/mixeddistribution).

## 3. Selecting the best fitting hazard function, MOROS data

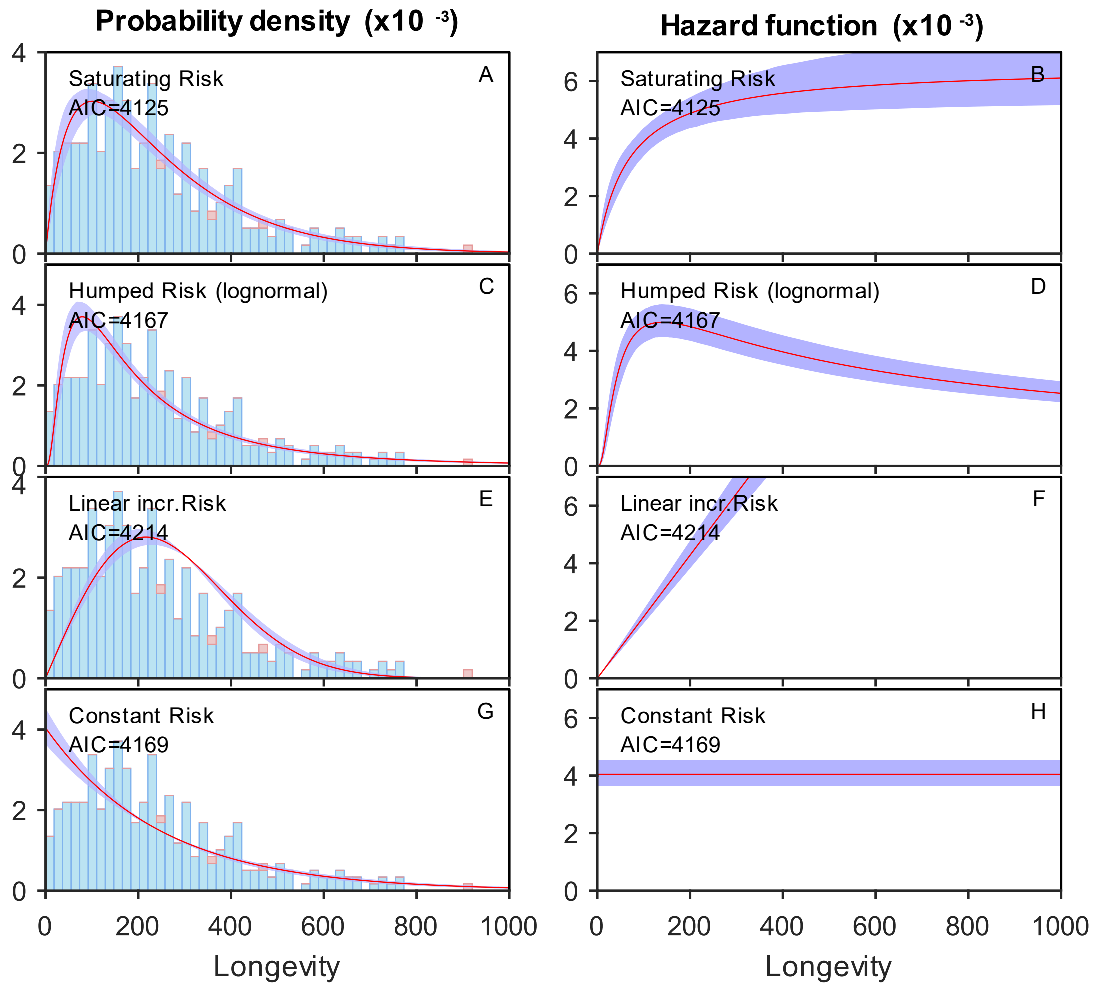


Figure S1 *Observed longevity distributions (blue bars) compared to what might be expected (red curves) assuming different distributions. The best-fitting (minimum AIC) model is Saturating Risk (A) (A,B: Saturating Risk, C,D: Humped Risk (lognormal), E,F: Linear Increasing Risk, G,H: Constant Risk. The specific optimized hazard functions relating termination risk with age are represented in the right-hand panels. Zones represent 95% uncertainty margins obtained from bootstrapping (see Methods). See Table S4 for equations. Fitted parameter values to the functions are: A,B: Distribution = Saturating; parameters: c=0.006514 h=67.29 (n=324); C,D: Distribution = Lognormal; parameters: mu=5.189 sigma=0.9051 (n=324); E, F: Distribution = Linear; parameters: a=2.14e-05 (n=324); G,H: Distribution = Exponential; parameters: mu=247.3 (n=324)*

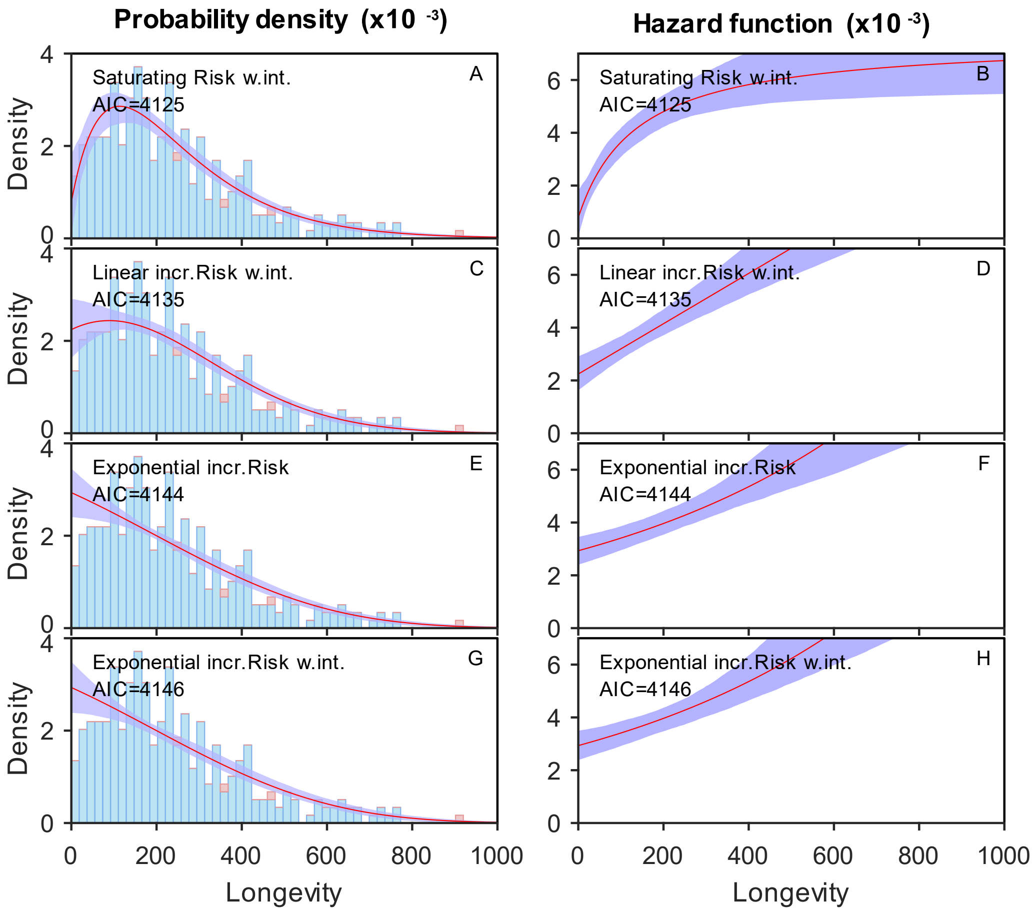


*Figure S2 Similar to fig S1, but now allowing for fitted intercepts as an expansion of the models. The fit is obviously a bit better. However, note that the AIC that includes a penalty for the number of fitted parameters is still minimal for the saturating risk without intercept (Figure S1A,B). (AB: Saturating Risk with intercept, CD: :Linear Increasing Risk with intercept, EF: Exponential Increasing Risk, GH: Exponential Increasing Risk with intercept. The specific optimized hazard functions relating termination risk with age are represented in the right-hand panels. Zones represent 95% uncertainty margins obtained from bootstrapping (see Methods). See Table S4 for equations.*

*Fitted parameter values to the functions are A,B: Distribution = SaturatingExponential; parameters: c=0.006763 h=136.9 r=0.0007851 (n=324); C, D: Distribution = LinearExponential; parameters: r=0.002243 a=9.533e-06 (n=324); E,F: Distribution = Gompertz; parameters: alpha=0.002933 beta=0.001507 (n=324); G, H: Distribution = GompertzMakeham; parameters: alpha=0.002933 beta=0.001507 lambda=9.212e-13 (n=324);*

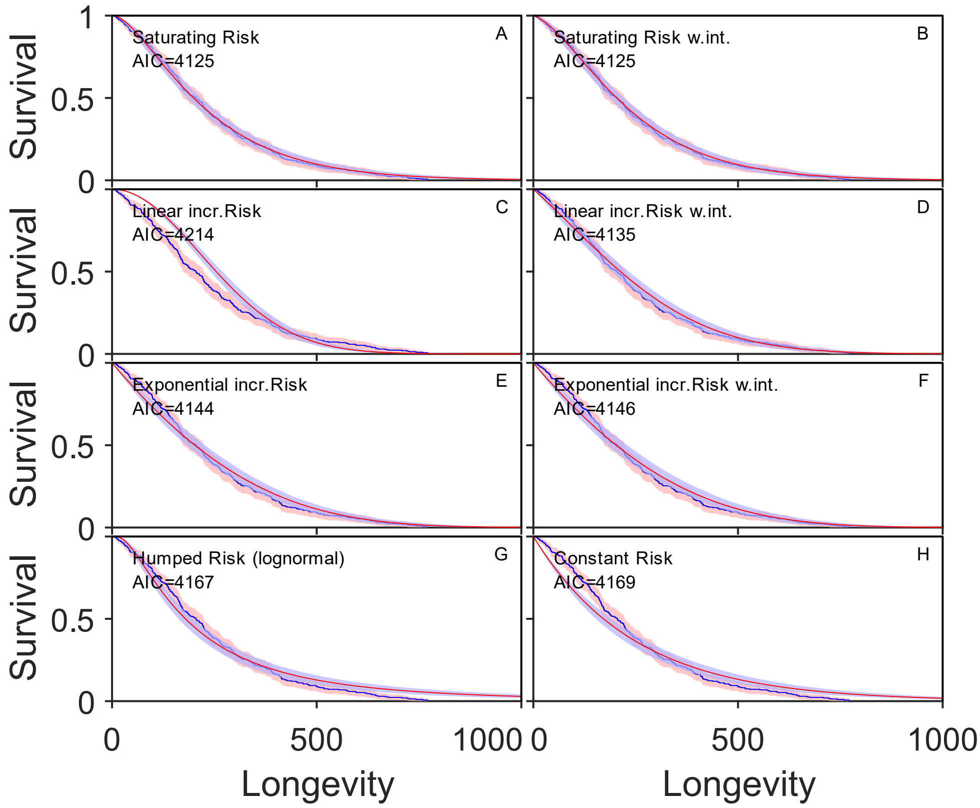


*Figure S3 Uncertainty estimates from bootstrapping suggest that the data are undistinguishable from what would be expected from a saturating risk with aging. Data are pink, models purple. Survival functions for the different fitted models (red lines with blue uncertainty zones) as compared to non-parametric survival functions and their uncertainty estimates (orange). Note that the uncertainty bands of the non-parametric and the fitted saturating risk functions overlap strongly (panels A and B). By contrast, a constant risk function deviates significantly from the non-parametric estimates (Panel H). For fitted parameter see the legends of Figure S1 and S2.*

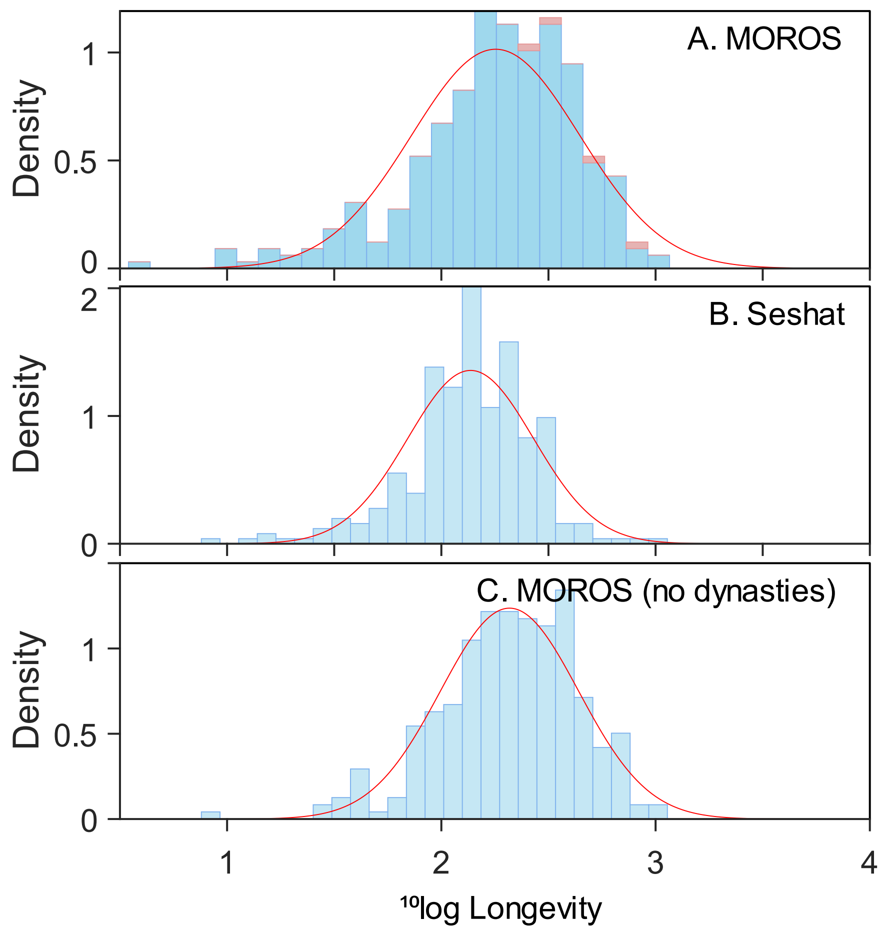


*Fig. S 4.1 Longevity follows roughly follows a lognormal distribution, both in the MOROS and SESHAT database. While the means of the distributions are similar, longevity in MOROS has a somewhat larger standard deviation, and is more left-skewed due to the presence of more short-lived societies. If we leave dynasties out (C), the two distributions are more similar.*

A. *MOROS: n=324, mean = 2.253, standard deviation = 0.393, skewness = -0.9846*
B. *SESHAT: n=291, mean = 2.137, standard deviation = 0.294, skewness= -0.677.*
C. *MOROS (without dynasties): n=274, mean = 2.3181, standard deviation = 0.3227, skewness = -0.5777*

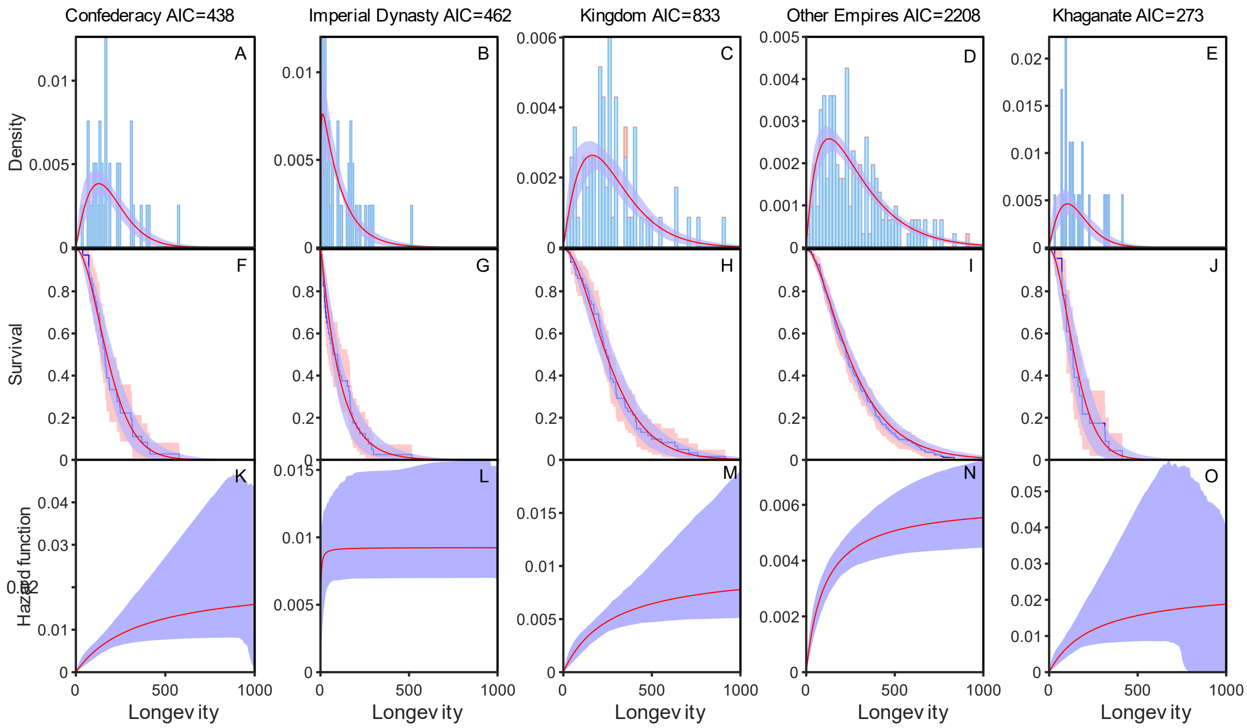


*Figure S4.2 Longevity distributions for different types of societies in MOROS. A saturating risk is the best fitting hazard function for each type, but parameters differ. Especially imperial dynasties have a remarkably low mean longevity.*

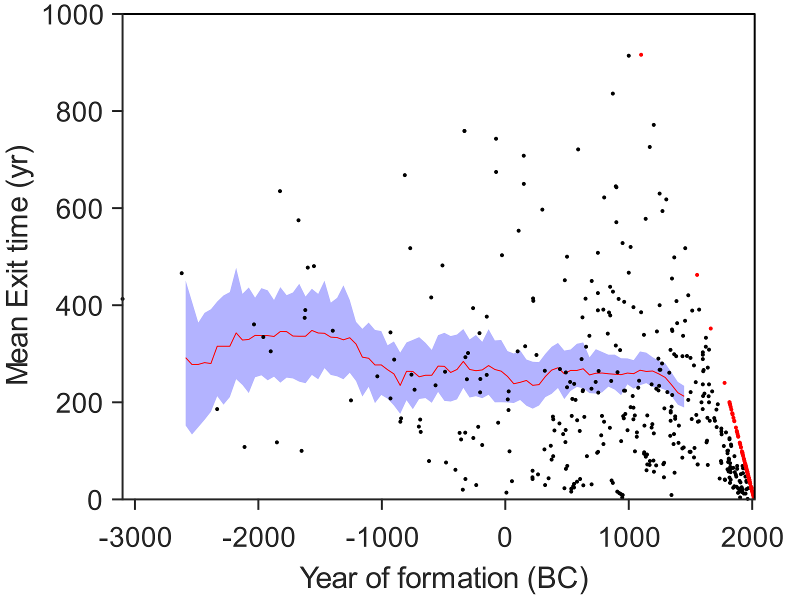


*Fig. S5 Mean longevity of the states in MOROS as a function of their origination year (calculated with a sliding window of 20 years). Red dots denote societies that are still existing ("right-censored data", see methods). Blue line represents the moving average of the number of observations in that period. Our study considers only societies that already ended and that originated before 1800.*

# 4. Selecting the best-fitting hazard function for Seshat

Explanation see comparable figures for MOROS

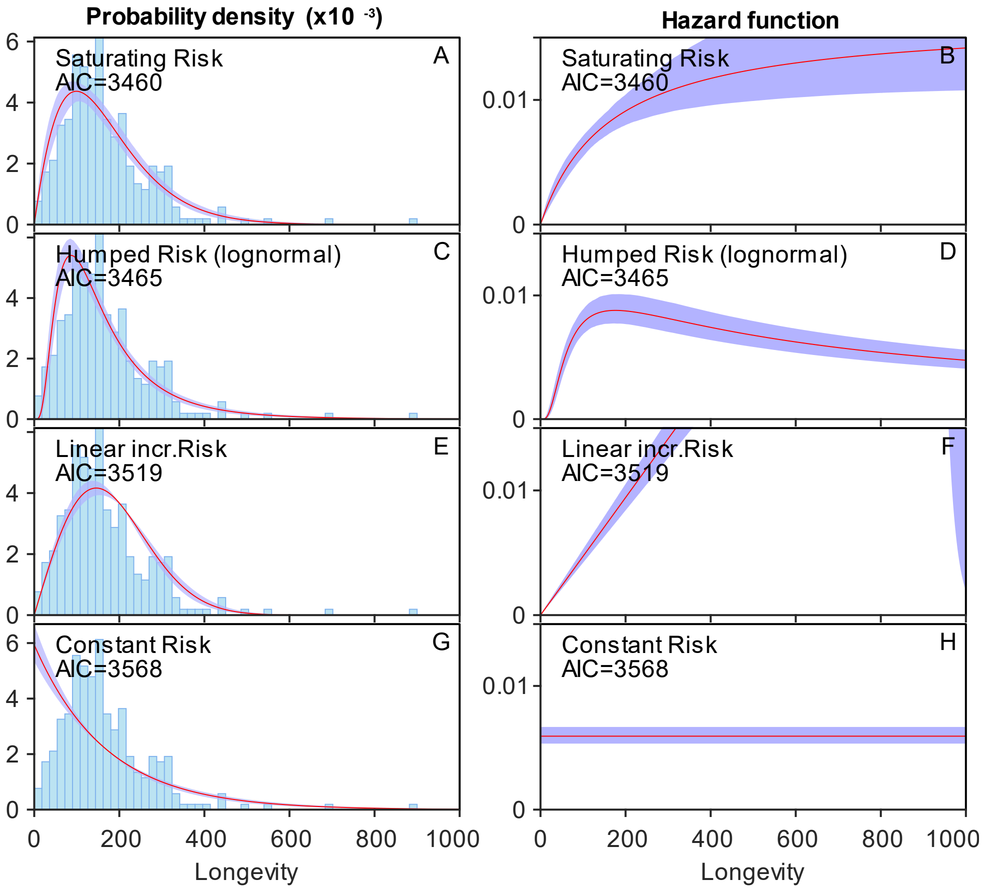


Figure S9: In the Seshat database without quasi-polities the Saturating risk function has the best fit. Fitted parameter values to the functions are A, B: Distribution = Saturating; parameters: c=0.01645 h=161.4 (n=291); C, D: Distribution = Lognormal; parameters: mu=4.92 sigma=0.6772 (n=291); E,F: Distribution = Linear; parameters: a=4.714e-05 (n=291); G,H: Distribution = Exponential; parameters: mu=168.4 (n=291);

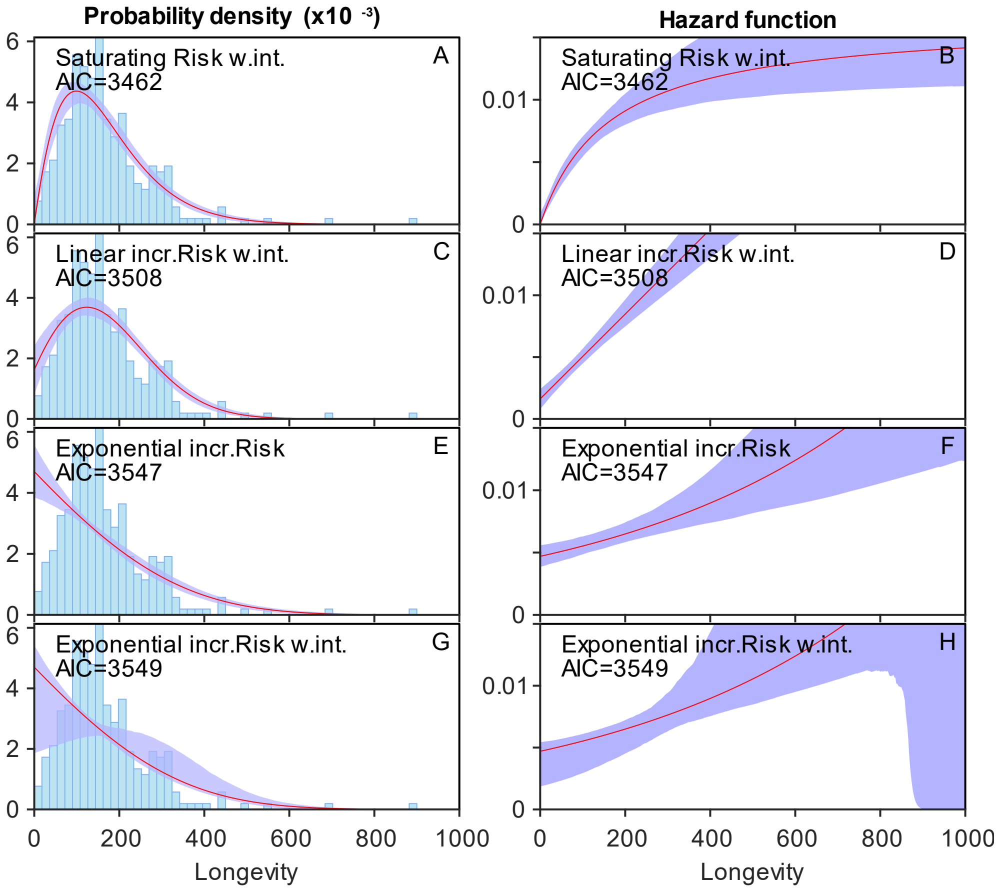


Figure S10: Adding a intercept does not improve the fit (Seshat data without quasi-polities). Fitted parameter values to the functions are:A,B: Distribution = SaturatingExponential; parameters: c=0.01645 h=161.4 r=9.343e-17 (n=291); C,D: Distribution = LinearExponential; parameters: r=0.001621 a=3.427e-05 (n=291); E,F: Distribution = Gompertz; parameters: alpha=0.004701 beta=0.001616 (n=291); G,H: Distribution = GompertzMakeham; parameters: alpha=0.004701 beta=0.001616 lambda=3.728e-14 (n=291);

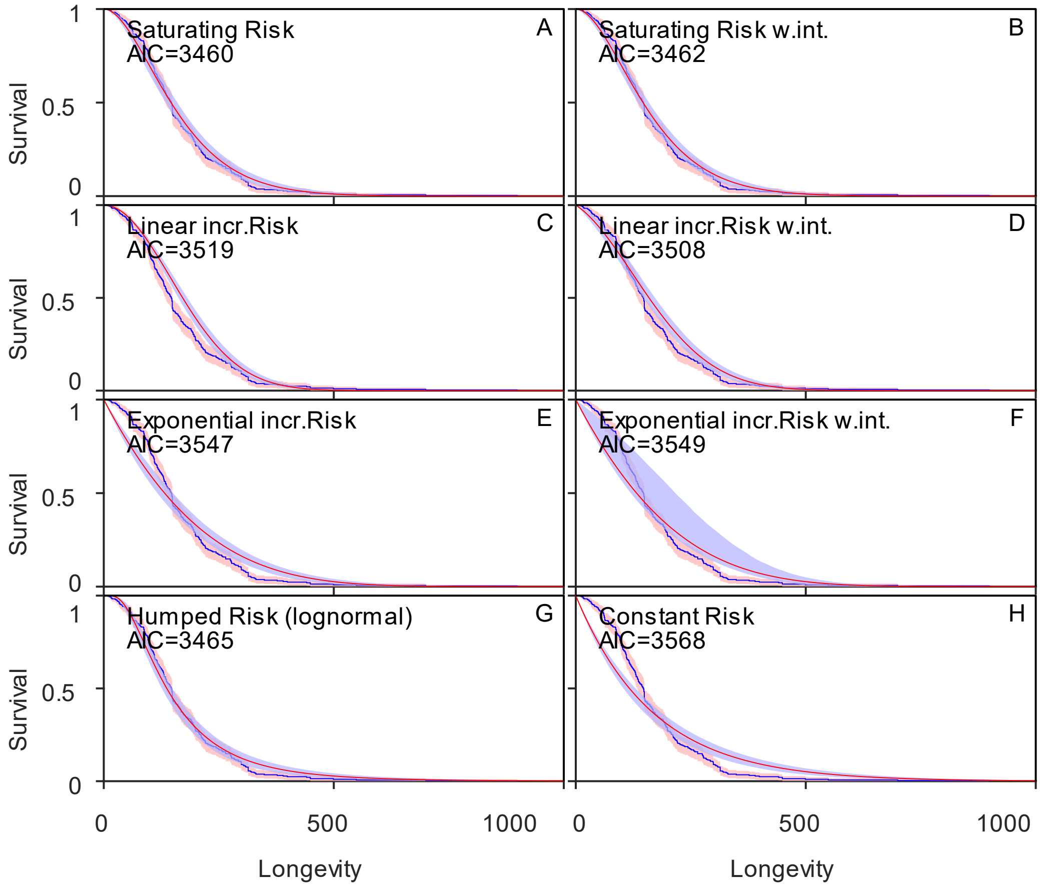


Figure S11: The Survival functions of the fitted distributions (purple) compared with the survival function directly based on data (pink) (Seshat data without quasi-polities). For parameter values see Figure S9 and Figure S10.

## 5. Selecting the best-fitting hazard function for MOROS without Dynasties

Explanation see comparable figures for MOROS.

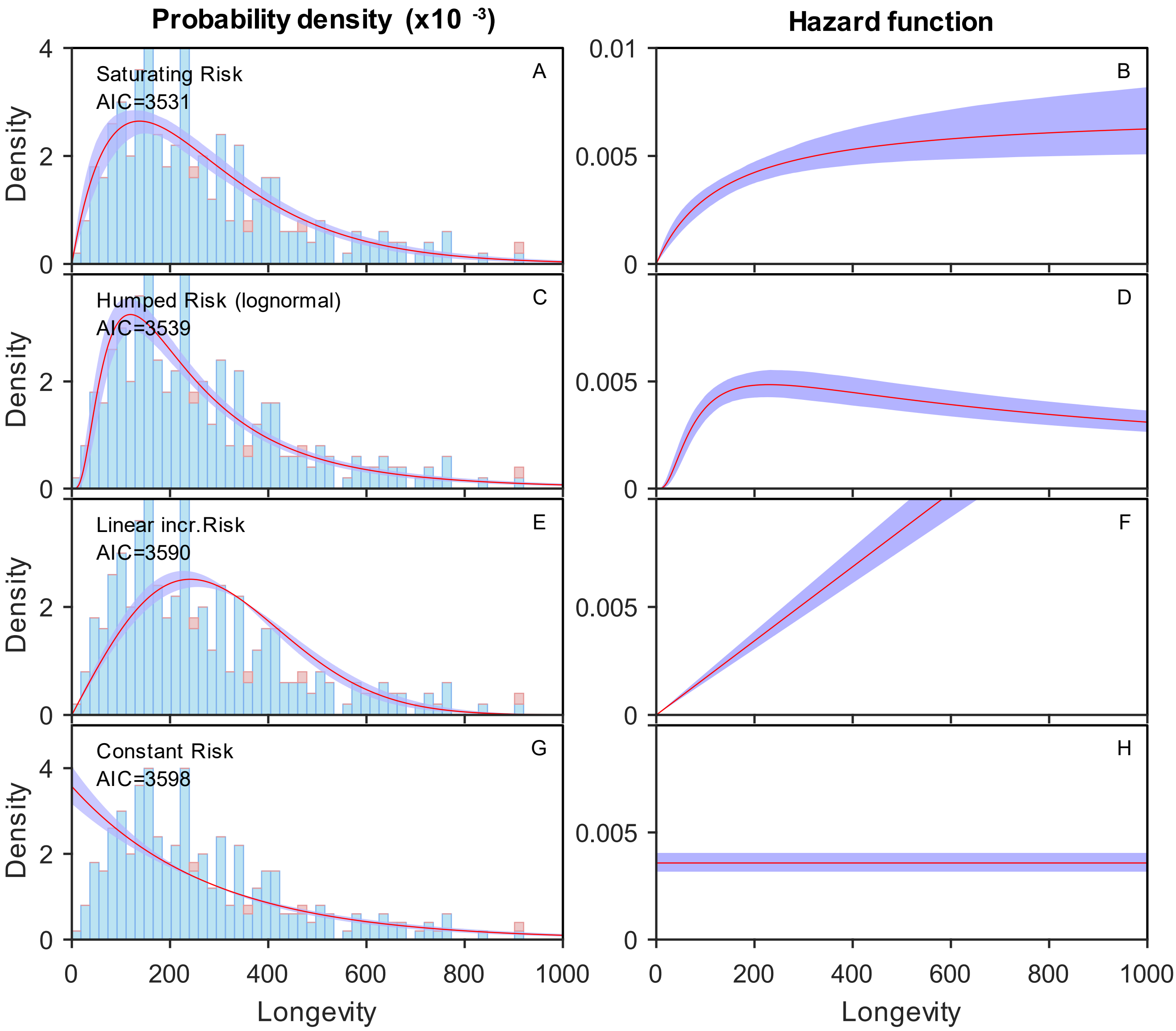


Figure S12. The fitted distributions (left) and hazard functions (right) for the Moros dataset, excluding the political classes of 'Dynasty' and 'Imperial Dynasty'. Fitted parameter values to the functions are A,B: Distribution = Saturating; parameters: c=0.007093 h=134.9 (n=275); C,D: Distribution = Lognormal; parameters: mu=5.37 sigma=0.7669 (n=275); E.F: Distribution = Linear; parameters: a=1.716e-05 (n=275); G,H: Distribution = Exponential; parameters: mu=279.8 (n=275).

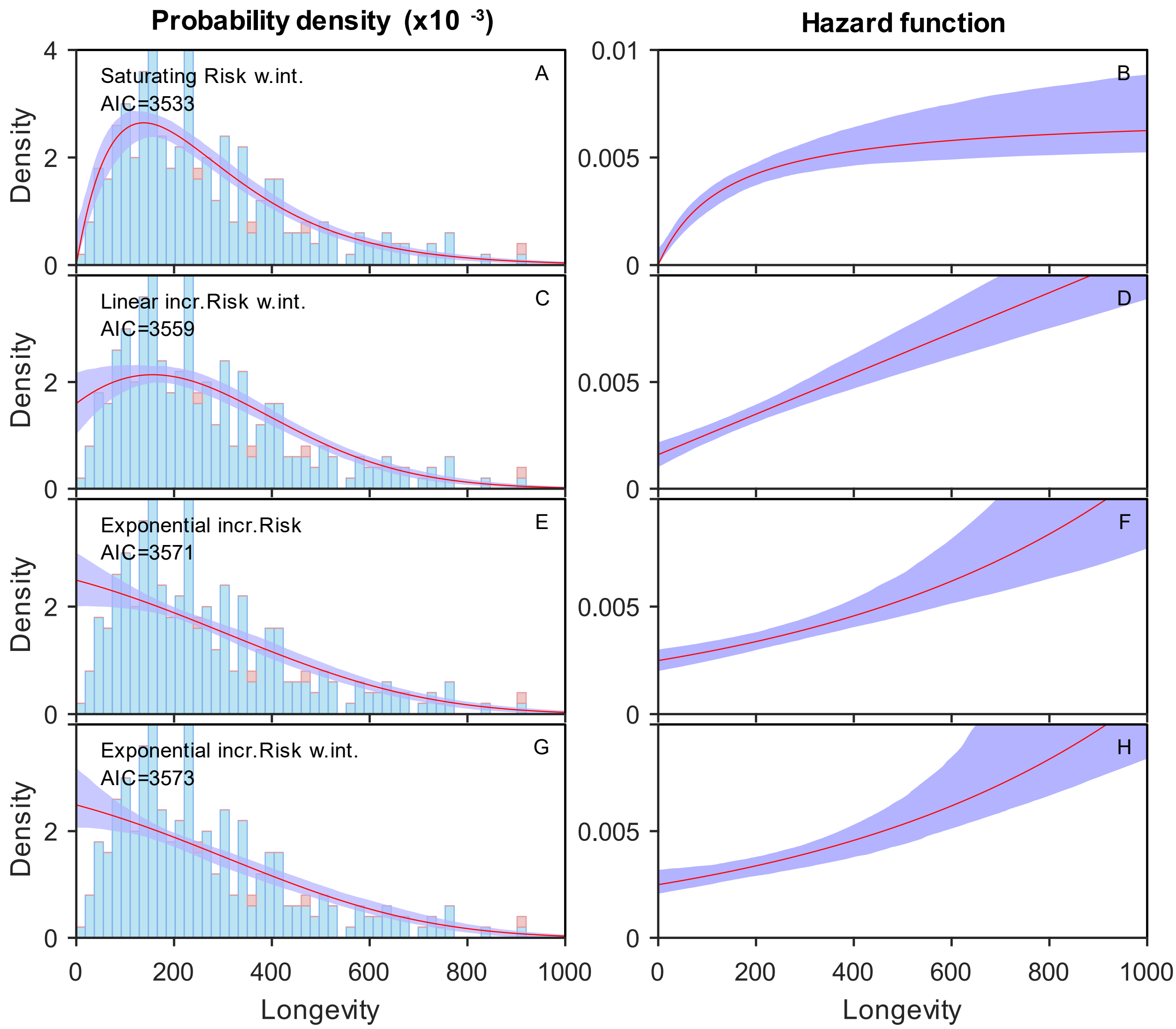


Figure S13. The fitted distributions (left) and hazard functions (right) for the Moros dataset, excluding the political classes of 'Dynasty' and 'Imperial Dynasty'. Fitted parameter values to the functions are: A,B: Distribution = SaturatingExponential; parameters: c=0.007093 h=134.9 r=5.612e-15 (n=275); C,D: Distribution = LinearExponential; parameters: r=0.001598 a=9.489e-06 (n=275); E.F: Distribution = Gompertz; parameters: alpha=0.002491 beta=0.001516 (n=275); G,H: Distribution = GompertzMakeham; parameters: alpha=0.002491 beta=0.001516 lambda=1.355e-16 (n=275).

# 6. references


1. Scheidel W (2013) Studying the state.
2. Graeber D & Wengrow D (2021) *The dawn of everything: A new history of humanity* (Penguin UK).
3. Haldon JF (1993) *The state and the tributary mode of production* (Verso).
4. Cline EH (2015) *1177 BC: The year civilization collapsed* (Princeton University Press).
5. Taagepera R (1997) Expansion and contraction patterns of large polities: Context for Russia. *International studies quarterly* 41(3):475-504.
6. Taagepera R (1979) Size and duration of empires: growth-decline curves, 600 BC to 600 AD. *Social Science History* 3(3-4):115-138.
7. Taagepera R (1978) Size and duration of empires growth-decline curves, 3000 to 600 bc. *Social Science Research* 7(2):180-196.
8. Turchin P*, et al.* (2015) Seshat: The global history databank. *Cliodynamics: The Journal of Quantitative History and Cultural Evolution*.
9. Turchin P*, et al.* (2019) An introduction to Seshat: Global history databank. *Journal of Cognitive Historiography*:115-123.
10. MacKenzie JM, Dalziel NR, Doumanis N, & Charney MW (2015) *The Encyclopedia of Empire* (Wiley Blackwell).
11. COWP (2021) Correlates of war project.
12. Wilkinson EP (2000) *Chinese history: a manual* (Harvard Univ Asia Center).
13. Lee S & Leidy DP (2013) *Silla: Korea's golden kingdom* (Metropolitan Museum of Art).
14. Tainter JA (1988) *The Collapse of Complex Societies. New Studies in Archaeology* (Cambridge University Press, Cambridge).
15. Middleton GD (2017) *Understanding collapse: Ancient history and modern myths* (Cambridge University Press).
16. Lawal B (2016) Benin (Edo city-state). *The Encyclopedia of Empire*), pp 1-6.
17. MacKenzie JM, Dalziel N, Doumanis N, & Charney MW (2016) *The Encyclopedia of Empire* (Wiley Online Library).
18. Wellen K (2016) Brunei, Sultanate of. *The Encyclopedia of Empire*), pp 1-3.
19. Lockhart BM (2016) Vietnamese (Dai Viet) Empire. *The Encyclopedia of Empire*), pp 1-6.
20. Green AS (2021) Killing the priest-king: Addressing egalitarianism in the Indus civilization. *Journal of archaeological research* 29(2):153-202.
21. Jennings J, Frenette S, Harmacy S, Keenan P, & Maciw A (2021) Cities, surplus, and the state: a re-evaluation. *Journal of Urban Archaeology* 4:15-31.
22. Hiribarren V (2016) Kanem ‐ Bornu Empire. *The Encyclopedia of Empire*:1-6.
23. Lipo CP & Hunt TL (2016) Chronology and Easter Island prehistory. *Skeletal Biology of the Ancient Rapanui (Easter Islanders),* Cambridge Studies in Biological and Evolutionary Anthropology, eds Gill GW & Stefan VH (Cambridge University Press, Cambridge), pp 39-65.
24. Bang PF, Bayly CA, & Scheidel W (2020) *The Oxford World History of Empire: Volume Two: The History of Empires* (Oxford University Press, USA).
25. Petersen G (2000) Indigenous island empires: Yap and Tonga considered. *Journal of Pacific History* 35(1):5-27.
26. Robinson DW (2016) Tukulor Empire. *The Encyclopedia of Empire*), pp 1-5.
27. Warner CM (1999) The Political Economy of'quasi-statehood'and the demise of 19th century African politics. *Review of International Studies* 25(2):233-255.
28. Fusaro M (2016) Venetian Empire. *The Encyclopedia of Empire*), pp 1-11.
29. Joyce AA (2016) Zapotec Empire. *The Encyclopedia of Empire*), pp 1-3.
30. Doyle MW (1986) *Empires* (Cornell University Press).
31. Turchin P*, et al.* (2019) Explaining the Rise of Moralizing Religions: A test of competing hypotheses using the Seshat Databank.

32. Whitehouse H, *et al.* (2021) Big Gods did not drive the rise of big societies throughout world history.
33. Lawless JF (2011) *Statistical models and methods for lifetime data* (John Wiley & Sons) p 630.
34. Kaplan EL & Meier P (1958) Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 53(282):457-481.