Economic History Working Papers No: 209/2014

Longevity and the Rise of the West: Lifespans of the European Elite, 800-1800

Neil Cummins
London School of Economics

Economic History Department, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, London, UK. T: +44 (0) 20 7955 7084. F: +44 (0) 20 7955 7730

LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE
DEPARTMENT OF ECONOMIC HISTORY
WORKING PAPERS
NO. 209- SEPTEMBER 2014

Longevity and the Rise of the West: Lifespans of the European Elite, 800-1800

Neil Cummins
London School of Economics
n.j.cummins@lse.ac.uk

Abstract

I analyse the age at death of 121,524 European nobles from 800 to 1800. Longevity began increasing long before 1800 and the Industrial Revolution, with marked increases around 1400 and again around 1650. Declines in violence contributed to some of this increase, but the majority must reflect other changes in individual behavior. The areas of North-West Europe which later witnessed the Industrial Revolution achieved greater longevity than the rest of Europe even by 1000AD. The data suggest that the `Rise of the West' originates before the Black Death.

Keywords: Key words: Mortality, Health, Nobility, Divergence

JEL Codes: N3, I15, J11

Longevity and the Rise of the West: Lifespans of the European Elite, 800-1800

Neil Cummins*

September 16, 2014

Abstract

I analyze the age at death of 121,524 European nobles from 800 to 1800. Longevity began increasing long before 1800 and the Industrial Revolution, with marked increases around 1400 and again around 1650. Declines in violence contributed to some of this increase, but the majority must reflect other changes in individual behavior. The areas of North-West Europe which later witnessed the Industrial Revolution achieved greater longevity than the rest of Europe even by 1000 AD. The data suggest that the 'Rise of the West' originates before the Black Death.

1 Introduction

The 'Rise of the West' has recently been traced to events long preceding the Industrial Revolution[1]. This paper shows how the spatial patterns of the lifespans of Europe's nobility suggest a European mortality pattern that has existed since 1000 AD. The parts of Europe that later experience the Industrial Revolution first (the North-West) have higher lifespans than those who later lag behind (the South-East). Nobles transform their behavior over the long run. Before 1550, about 30% of noble men die in battle. After 1550, the figure is less than 5%. Surprisingly, the Black Death and subsequent waves of pestilence kill nobles at a lower rate than the general population and the lethality is higher for women. There is a structural break in noble lifespan about 1400, where lifespan increases from around 50 to 55. These findings suggest that the origin of the divergence of the 'West and the rest' has its origin even earlier than recent research suggests.

*Dept. of Economic History, London School of Economics. Email: n.j.cummins@lse.ac.uk. I thank Greg Clark, Morgan Kelly, Alan Fernihough and Cormac Ó Gráda for valuable suggestions. This research would have been impossible without the incredible genealogical labors of the Church of Jesus Christ of the Latter Day Saints.

[1]See, for example, Clark (2007), Acemoglu and Robinson (2012), Voigtländer and Voth (2013) and Broadberry (2013).

The emergence of modern economic growth during the Industrial Revolution was accompanied by an explosion in Europe's population. Demographic factors are not tangential to mankind's escape from the Malthusian nightmare: they are theorized to have played a causal role (Becker et al. (1990); Clark (2007); Galor (2004)). The reasons behind the modern rise in lifespan are debated[2]. One notable absence for these debates is an international time series to characterize trends over the long run. A major issue is the fact that before 1538, individual level demographic data is sparse[3]. However, one sub-population that have left abundant evidence of their lives are the European nobility. This analysis exploits recent mass digitization of family trees to examine trends in adult lifespan over the millennium between 800 and 1800[4].

The paper is complimentary to recent work by David et al. (2010) and de la Croix and Licandro (2012)[5]. David et al. (2010) use Alison Weir's genealogy of the British Royal family to explore the evolution of life expectancy between 1500 and 1799. de la Croix and Licandro (2012) use a data-set of over 300,000 famous people from the *Index Bio-bibliographicus Notorum Hominum* examine the long time trend in lifespan. They argue that average age at death was stationary until the birth cohort of 1640. However, they decide to omit any analysis of the time-trend in lifespan before the 15th century; They only estimate trends post-1430 (see their figure 6, p.15). This analysis examines trends beginning over six centuries before either David et al. (2010) or de la Croix and Licandro (2012).

This paper has 6 sections. Section 2 discusses the data, section 3 details the methodology for the analysis while section 4 presents the results. The results section has four principle subsections: On violence (4.1), plague (4.2), time-trends (4.3), spatial patterns and time-trends by region (both 4.4). Section 5 discusses the implications of these findings and section 6 concludes. The appendix details the underlying distributions and supplementary regression results. A separate stand alone appendix (Cummins (2014)) details the data collection strategy of 1.3m records, the date coding of 402,204 string dates, the Geo-coding of 117,975 unique addresses, the categorization of nobles into 17 ranks and a sample of random sources and observations[6].

[2]See Cutler et al. (2006) for a brief summary. For more detail see Schofield et al., eds (1991).

[3]In 1538, Thomas Cromwell orders all churches in England and Wales to keep a register of births, deaths and marriages. Similar rules come into effect on the continent around this time. These parish registers dominate our understanding of the preindustrial demographic world via the seminal contributions of Wrigley and Schofield (1981); Wrigley et al. (1997); Henry (1972); Henry and Houdaille (1973); Henry (1978); Houdaille (1976).

[4]Existing demographic studies of Europe's aristocracy included Hollingsworths' analysis of the British Ducal families and peerage, Peller's analysis of Europe's ruling families and Levy and Henry's analysis of French nobility (Hollingsworth (1957, 1964, 1975, 1977); Peller (1965); Levy and Henry (1960)).

[5]Fire and Elovici (2013) is similar in terms of data collection strategy. (See stand alone appendix for details).

[6]Further, a list of the 3,133 sources will be provided in a data file on my website, neilcummins.com. Following publication, replication files and data will be provided there too.

Tree ID	Tree Title	N	Nsources
1	British Isles. Heraldic Baronage	14,799	1
2	British Isles. Peerage,		
	Baronetage, and Landed Gentry.	225,220	273
6	England. Leicester. Long Clawson.	20,739	10
7	England. London. Residents.	296,738	90
9	England. London. Visitation, 1664	4,686	1
10	England. Norfolk. 1563 Visitation.	23,098	1
12	England. Sussex. Genealogies.	20,508	2
14	Europe: Royal and Noble Houses	332,511	2,040
16	Ireland. Early Irish Families	4,295	129
17	Wales. Welsh Medieval Nobility and Gentry	267,857	713

Notes: The family tree records were constructed from published sources by the LDS church and various collaborators. They are available at https://histfam.familysearch.org.

Table 1: The Sample of Family Trees

2 Data

'Baptism for the dead' is a doctrine of the church of Jesus Christ of the Latter Day Saints (LDS). The practice is mentioned in the Bible (Corinthians chapter 15, verse 29, *The Holy Bible King James Version* (2014)). The founder of the LDS church, Joseph Smith, revived the practice in 1840 and ever since, church members have been collecting historical genealogical data and baptizing the dead by proxy. The church has been at the frontier of the application of information technology to genealogy and has digitized a multitude of historical records. Today they make the fruits of their research available online at familysearch.org. The records number in the billions.

This analysis uses records from family trees. The source of the data is histfam.familysearch.org a collaboration between the LDS church (familysearch) and individual genealogical experts. Table 1 reports the titles, sample size and number of sources for the ten family trees databases used in this paper. The individual entries are constructed from published works such as Burke and Burke's "A Genealogical and Heraldic Dictionary of the Peerage and Baronetage", (Burke and Burke (1881b)) "An Official Genealogical and Heraldic Baronage of England" Paget (1957) and Boyd's "Pedigrees with index of London citizens, abt. 1600-1800" (Boyd (1954)), numerous other published genealogical works, guild records, census records, parish registers, wills and other published family genealogies. Boyd and Burke and Burke are the leading sources (providing 295,892 and 127,269 records respectively), followed by de Sainte-Marie and de Sainte-Rosalie (1728) (73,723) and Schwennicke (2005) (70,835)[7]. The ten family trees used here are summarized in table 1.

How reliable is this data? Is it fiction? We can examine the distributions of age at death by period and see if fantastical ages are being attributed or if some average is just

[7]The complete list of the 3,117 sources will be provided as a separate pdf file (as it runs to 204 pages).

blankly applied. This does not appear to be the case - The basic shape of these distribution reveals patterns that seem to reflect a fairly consistent underlying pattern - See figures 20 and 21. It does not appear that some different process (e.g. speculative guesswork) is driving the pattern more or less as we go towards the 16th century, where we know the data is much better (and can be corroborated with parish records etc.)[8]. The connection of each individual record to its source(s) is a sign that at least the digitization and collation is of a tractable design. By making the underlying analytical data freely available I also invite other scholars to replicate my analysis here.

Family trees are an under-utilized resource for academic research. This is perhaps related to the difficulty of making the records amenable for statistical analysis. The family tree records used here contain 402,204 unique date descriptions. The entries are inconsistent, of varying quality and sometimes refer to different calendars at different points in time. I have standardized all of these dates to decimal values of years, using random attribution for missing months and days. This process is described in detail in the stand-alone appendix (Cummins (2014)). The quality of the dates recorded improved greatly over time (as indicated by the extent of an index of heaping (an excess of years ending in 0 or 5) in figure 1).

As table 1 details, the majority of the sample consists of individuals connected to European nobility. This sample is in no way representative of the general population. This is an extremely elite subsection. This paper makes no claims about the general population, just this elite sub-section. However, the nature of the selection undoubtedly changes over time. How does an individual born in 850 make it in to the data, versus someone born in 1650? Each record in the family tree data used here contains an indicator of nobility via the 'suffix' variable. I have assigned all 7,607 unique suffix values to 1 of 17 simple 'noble rank' categories from 1 (Emperor) to 17 (no suffix). The ranking is inferred from general sources such as such as Doyle (2010) and also Burke and Burke (1881a). The exclusiveness of the sample declines as the centuries pass (See the stand alone data appendix for detail on this). The family trees artificially select for 'successful' offspring and ancestors, neglecting the 'failures'. However, as the standing of each member of the family is recorded via a suffix, the section based on success is directly observed. Thus I can control for the changing exclusivity of the sample over time.

There are more selection issues: Omission is heavy. For example, there are more men recorded than women, and there is wild under representation of infant and child deaths, before 1500. I argue that this genealogical data is amenable for scientific analysis because firstly, the variable of prime interest, age at death can be restricted to those dying at age 20 and over. Secondly, the nature of the changing composition of the sample is directly observed and this can be controlled for in the analysis (e.g. sex, suffix, data type, family tree and geographic composition). Finally, I must stress that this analysis only concerns a

[8]However, the unusual shape of male age at deaths in the 8th century is perhaps suggestive of some attribution, but the sample size is relatively small here and the pattern does not seem to be any different for women.

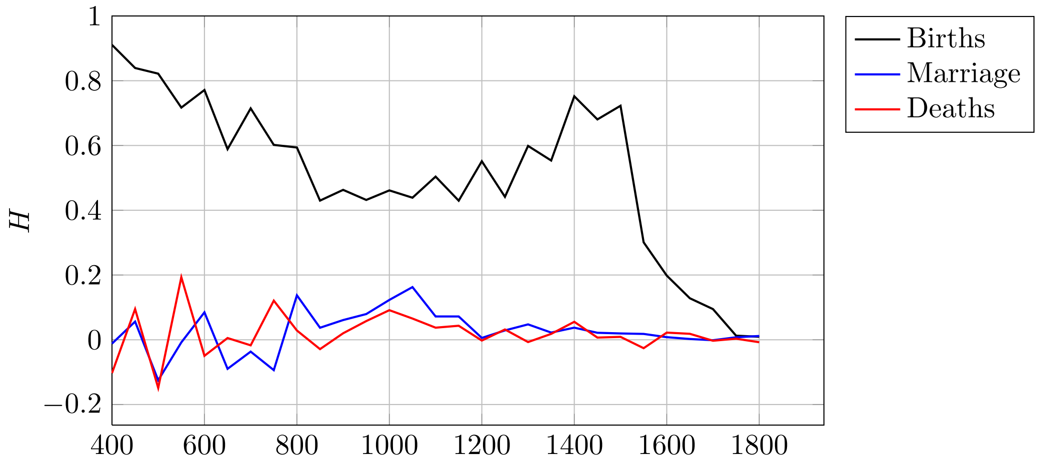


Figure 1: Year 'Heaping' over Time

Notes: $H = 5/4 * (X - 20)$ where X is the percentage of birth, death or marriage years ending in 0 or 5. Where events are reported accurately by year H will average 0. $H = 1$ implies all dates are being approximated. Births suffer severely from 'heaping' whilst marriages and deaths do not. Source: Noble sample.

very elite subsection of the human population.

Table 2 reports the counts by family tree and birth decade for those observations where both birth and death dates are recorded. Each individual has been geocoded. Figure 2 displays the geographic distribution of the full sample. In the analysis, family tree and geography are controlled for in all regressions. Of the 1,329,466 individual records, 168,167 have age at death recorded: 121,524 have an age at death over 20. 76,403 have a specific day of death. These observations form the basis of the analysis conducted here. A stand alone appendix (Cummins (2014)) describes in considerably more detail the construction of this data set.

	Tree ID									
	1	2	6	7	9	10	12	14	16	17
pre-800		11						280	162	86
800		30						311	55	69
900		39						570	81	108
1000	6	121						1,031	56	148
1100	51	215						1,910	70	249
1200	385	480				1	1	2,981	66	353
1300	350	743		1		3	9	4,768	50	492
1400	98	877		26		24	12	6,564	113	838
1500	10	2,911		6,694	41	430	569	13,193	84	3,418
1600		10,094	85	21,887	333	723	1,546	11,904	149	7,353
1700		16,399	155	8,007	10	302	1,030	2,932	92	7,392
1800		18,852	137	588		18	434	818	16	3,134
1900		933		1			6	27		27

Notes: See table 1 for tree titles. The paper examines the 800-1800 period where the data is dense enough to allow robust estimation. Source: Noble sample.

Table 2: Counts by Century and Tree ID, Age at Death Observed

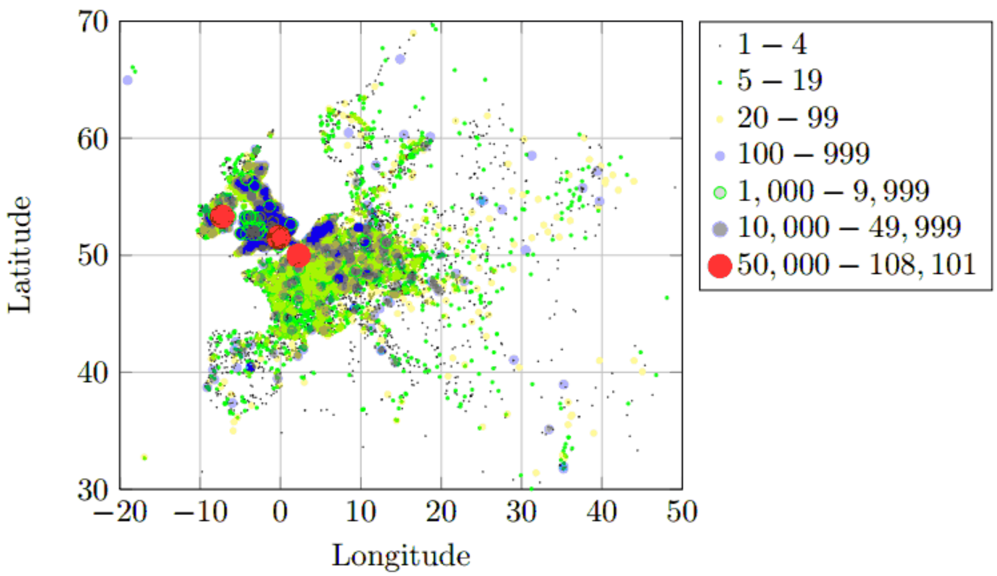


Figure 2: Density of Observations

Notes: This is a simple scatter plot of the attributed geocoded addresses. There are 117,975 variations of addresses of birth, marriage and death detailed in the 1.3m family tree records. See the stand alone appendix details on the geocoding. Source: Noble sample.

3 Methodology

This section details the empirical strategy that aims to characterize trends in lifespan amongst European nobility from 800 to 1800. Table 3 presents the summary statistics[9].

Firstly, I analyze those records with a specific day or month of death to measure the significance of violence and plague in noble mortality (sub-section 4.1 and 4.2). Both matter.

Following this, I apply sequentially, an OLS model, a quantile regression model and a Bayesian Additive Regression Tree (BART) model to an estimation equation of the form:

$$\begin{aligned} Age_D &= C + D_{Female} + Lat + Long + \sum_{i=1}^{101} D_{BirthDecade2} + \sum_{i=1}^{17} D_{NobleRank} + D_{Bastard} \\ &\quad + D_{Violent} + \sum_{i=1}^{10} D_{Tree} + \sum_{i=1}^{4} D_{DateQB} + \sum_{i=1}^{4} D_{DateQD} + \varepsilon \end{aligned} \tag{1}$$

Where Age_D = Age at Death (for those over 20), C is a constant, D_{Female} is a female dummy, Lat and $Long$ are (attributed) latitude and longitude, $D_{Birthdecade2}$ are a set of 101 categorical variables indicating the 20 year interval of birth from 240BC to 1960. (However, the data only become dense enough to estimate time-trends in the ninth century.) $D_{NobleRank}$ is a set of 17 categorical variables indicating noble rank[10], $D_{Bastard}$ equals one where an individual is the result of an illegitimate union, $D_{Violent}$ is a dummy indicating where the death is likely violent (see subsection 4.1), D_{Tree} identifies the family tree of origin and D_{DateQB} and D_{DateQD} are indicators (1-4) for the quality of the data for both birth and death respectively (Here data quality refers to the precision of the date estimate - See the stand alone appendix for detail on this).

The empirical challenge is to extract from the noisy data the major time and spatial trends in noble lifespan while controlling for the changing selectivity and composition of the sample. Equation 1 captures this by directly including controls for sex, geography, noble rank and an indicator for a violent death. Further, the family tree of origin is included as a categorical variables as are 12 separate data quality variables. Every possible co-variate that could confound our characterization of the time-trends that can be included has been included. The resulting coefficient estimates on the set of birth period dummies can therefore be interpreted as representing the controlled time trend.

[9] The average latitude and longitude correspond to a field beside Pellingford Brook in East Sussex, the South of England.

[10] Nobles are ranked in rough order of prestige: 1: Emperor, 2: King, 3: Grand Duke, ArchDuke, Ancient, 4: Duke, 5: Prince-Elector, Prince, 6: Earl, Count, 7: Marquess, Margrave, 8: Viscount, 9: Baron, 10: Baronet, 11: Knight, 12: Esquire, Gentleman and unassigned nobility, 13: Lord, 14: Geographic, 15: Military, 16: Religious, 17: Occupational and 18: No Suffix, Meaningless Suffix. See the stand alone appendix, Cummins (2014), for detail.

Statistic	N	Mean	St. Dev.	Min	Max
D_{Female}	121,478	0.35	0.48	0	1
Birth day in year	121,524	181.75	105.31	1	365
Birth year	121,524	1618.29	200.66	−229	1961
Death day in year	121,524	180.35	106.63	1	365
Death year	121,524	1674.74	203.56	−186	2013
$N_{Sources}$	121,524	2.05	2.45	1	63
Noble Rank	121,524	0.29	0.45	0	1
Latitude	121,524	50.99	7.02	−46.41	69.82
Longitude	121,524	−0.01	16.13	−162.05	179.82
Age at Death (> 20)	121,524	56.94	18.84	20.00	122.95

Source: Noble sample.

Table 3: Summary Statistics

In addition, I have employed several econometric methodologies. Distributions of age at death for the noble data sample are reported in the appendix (figures 16, 17, 20 and 21). The distributions are multi-modal. There are some striking features, particular related to pronounced female mortality during the peak child bearing years (see figure 17 for the sample as a whole, the effect is spectacularly strong in London, figure 19). This effect disappears at upper end of the age distributions. To deal with the variety of effects that could change the mortality environment at different quantiles of the age at death distribution I employ quantile regression to examine time-trends for the .1, .25, median, .75 and .9 percentiles. The estimating equation is the same as equation 1 so the results can be directly compared with those estimated for the mean by OLS.

After examining the data for a common time trend I test for spatial heterogeneity via heat-maps and the separate estimation of equation 1 across sub-periods. To deal with the potential non-linear effects of geography, I run a regression of the form:

$$\begin{aligned} Age_D &= C + D_{Female} + \sum_{i=1}^{106} D_{Lat} + \sum_{i=1}^{214} D_{Long} + \sum_{i=1}^{101} D_{BirthDecade2} + \sum_{i=1}^{17} D_{NobleRank} \\ &\quad + D_{Bastard} + D_{Violent} + \sum_{i=1}^{10} D_{Tree} + \sum_{i=1}^{4} D_{DateQB} + \sum_{i=1}^{4} D_{DateQD} + \varepsilon \qquad (2) \end{aligned}$$

Equation 2 is exactly the same as equation 1 except for the inclusion in equation 2 of dummy variables for each integer value of longitude and latitude in the sample. Finally, I divide my sample into seven separate geographic regions and use a BART model to predict noble longevity across time and space, fully accounting for the inherent non-linearity and heterogeneity.

Regression trees use algorithms to grow 'trees' that recursively divide samples of the data along values of predictor variables that best fit the observed outcome (how regressions based upon those divisions predict the rest of the data decide the fit). BART models are a sum-of-trees approach that allow interaction and additive effects using priors to keep the individual tree effects small (and approximating a different part of the unknown function). To fit the model, BART uses an iterative Markov-chain Monte Carlo algorithm (Chipman et al. (2010), Green and Kern (2012)). Here age at death can be modeled as the outcome of the predictor variables listed in equations 1 and 2 (denoted as X). The unknown function is approximated by m regression trees of structure T with terminal node parameters (leaves) L[11]:

$$Age_D = f(X) + \varepsilon \approx T_1^L(X) + T_2^L(X) + \ldots + T_m^L(X) + \varepsilon \tag{3}$$

The advantages to using this relatively new methodology is that it allows for model free variable selection (each variable can be assessed in terms of its predictive importance), no assumptions about functional form and the ability to incorporate heterogeneous, interactive and additive effects. For the description of noble longevity across one thousand years and ten million square kilometers, the choice of this approach seems appropriate[12].

4 Results

4.1 A History of Violence

European nobility specialized in the execution of violence. Their genealogies connected them to the Barbarian conquerors of Europe following the decline of the Roman Empire. We can expect that a large proportion, especially of the men, died in battle. How can we know if a individual in the data dies from violence? Where the individual has a specific date of death (an exact day), we could link that date to a list of all known battles in European history. However, many battles have been lost from history's memory. The genealogical records of noble deaths themselves may be the only remnant of minor dark age skirmishes.

To investigate how many nobles died from violence, I employ a general version of the famous birthday problem. First year statistics students are often introduced to probability via the surprisingly low number of people it takes to have a high probability of a shared birthday. If we take the number of exact-date deaths per year, n, and the observed shared death days in a given year, m, we can calculate the probability that this will occur randomly[13].

[11] See Kapelner and Bleich (2013, p.3). Birth year is included instead of period dummies.

[12] The model was estimated in **R** using the bartmachine package Kapelner and Bleich (2013). In addition quantile regression forests are used to predict the uncontrolled (reported in the appendix).

[13] For example, if there are 30 deaths in a year, the probability that two people share a death day, if death days are distributed randomly over 365 days of a typical year is about .70. The probability of three people sharing a death day is about .03.

m	Date	n	Battle
88	9 Sep 1513	148	Battle of Flodden
83	25 Oct 1415	154	Battle of Agincourt
36	9 Jul 1386	80	Battle of Sempach
23	10 Sep 1547	112	Battle of Pinkie Cleugh
22	26 Aug 1346	77	Battle of Crécy
16	10 Aug 1557	151	Battle of St. Quentin
15	1 Jul 1690	144	Battle of the Boyne
15	19 Jul 1333	58	Battle of Halidon Hill
15	11 Jul 1302	43	Battle of the Golden Spurs
14	29 Mar 1461	60	Battle of Towton

Notes: Calculated for the "Europe: Royal and Noble Houses" family tree only (to ensure no duplicates skew the calculation). All of these battles had heavy noble casualties. Source: Noble sample.

Table 4: Ten most frequent exact death dates

Table 4 reports the top 10 dates of death in the database. Each of these dates corresponds to a major European battle. Using the observed ranges of n and m (2-88 and 1-1,935 respectively), expected probabilities of coincident days of death were calculated. [14]. These values are reported in figure 3. After $m = 13$, the expected probabilities of any coincident dates of death in the sample is never above .5. All of these dates are assigned as 'likely violent'. Below 14, I assigned all dates as likely violent where the expected probability for that number of shared death days is below .5. Figure 4 reports the time-path of an index of violent deaths, averaged over family tree and death half century[15].

A useful, endogenous test of this procedure is to examine the 'violent' deaths by sex. Women were far less likely to die in battle and the assignment algorithm does a good job - See figure 22: Female violent deaths are much lower than male violent deaths and exhibit no trend over time[16].

In order to examine the determinants of a violent death and to calculated a controlled time trend, a logistic regression is run of the form:

[14] The procedure is extremely sensitive to duplicates. Therefore the m and n combinations for each exact death date were calculated by family tree.

[15] The index was calculated for each family tree, where possible. However, the family tree data set for Leicester, Long Clawson gave perverse results, e.g. nearly 32% of females dying violently (based on this procedure). This was because Long Clawson was a relatively small parish and all the death dates were actually burial dates. Therefore the tendency for burials to be clumped in smaller parishes resulted in a misattribution of these deaths to violence. Long Clawson, family tree id number 6 was therefore dropped from the ansaysis here.

[16] However, they are also non-zero. Non-violent coincident death dates will also be captured by this exercise. For example the sinking of the White Ship on the 25 November 1120.

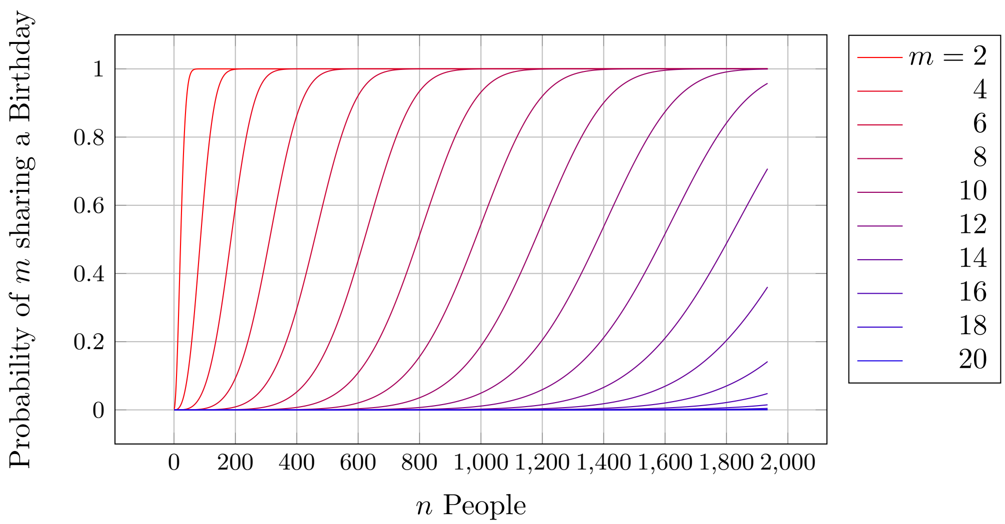


Figure 3: Expected Probabilities of Shared Death Days

Notes: $m =$ the number of people sharing a birthday, $n =$ the number of people at risk (e.g born in a given year). Calculated using the birthday command in R: http://www.inside-r.org/r-doc/stats/qbirthday

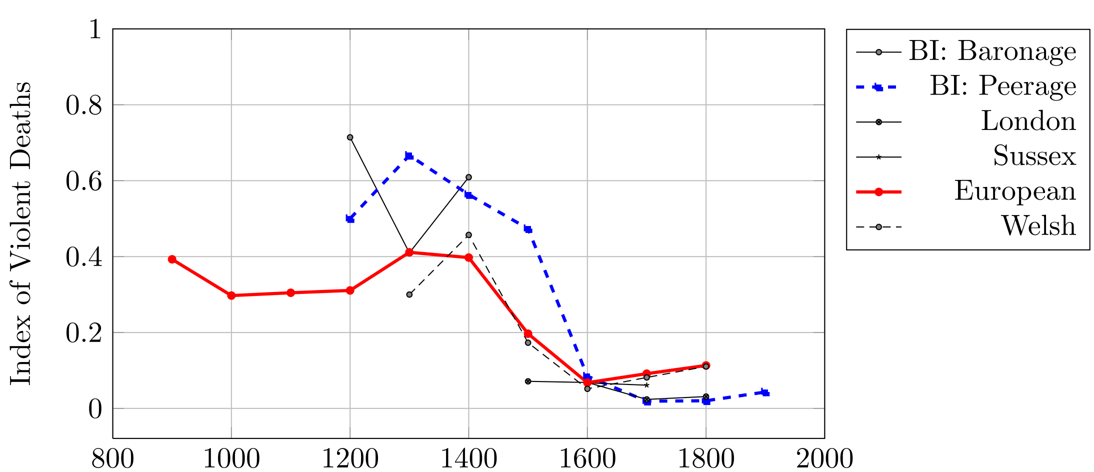


Figure 4: Proportion of Deaths from Battle

Notes: Figure reports the average proportion of male deaths from battle or violence ($N_{violent}/N_{All}$, where date of death is reported as an exact day) via the birthday probability exercise described in the text. Source: Noble sample.

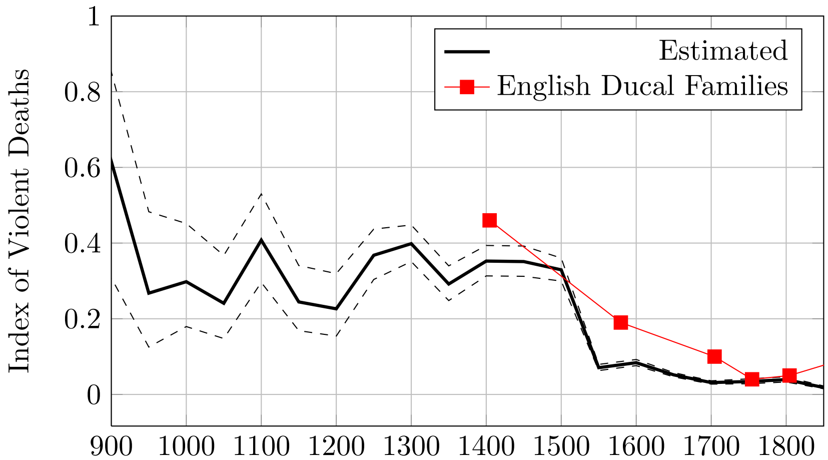


Figure 5: The Time Trend of Male Violent Deaths, from the Logistic Regression
Notes: The figure reports the predicted probabilities of a violent death (based on the birthday probability exercise) from a logistic regression controlling for sex and geography. Source: Noble sample and (Hollingsworth, 1957, p.8).

$$D_{Violent} = D_{Female} + D_{Bastard} + Latitude + Longitude + DYR \tag{4}$$

Where the notation is as equation 1 and DYR is the year of death. Table 5 reports the determinants of a violent death for those who have an exact death date in the sample[17]. Figure 5 reports the time path of predicted violent deaths. There is a sharp decline from 1500 to 1600. The trend corresponds closely to that reported by Hollingsworth (1957, p.8) for the English Ducal families.

A confounding factor in calculating violent deaths this way is the sudden return of plague to Europe in the 14th century. However, the principal result that violent deaths decline for the nobility is robust as this occurs during exactly the period during which we would expect the wave of plagues to bias this violent death index upwards.

4.2 ***Dance of Death***

Recent estimates of the lethality of the Black Death suggest a toll of 50 million, or about 60% of Europe's population (Benedictow (2004, p.383)). Emerging from the East in 1346,

[17] Table 13 in the appendix reports the noble rank effects. It is notable that military suffixes are associated with a higher risk of dying from violence (although the standard error is large), and religious occupations are significantly negative associated with a violent death.

	Dependent variable:	
	$D_{Violent}$	
	(1)	(2)
D_{Female}	$-.298^{***}$	$-.215^{***}$
	(.031)	(.033)
$D_{Bastard}$	.204	.256
	(.261)	(.257)
Latitude	$.016^{***}$	$.010^{**}$
	(.005)	(.004)
Longitude	$-.004^{**}$	$-.001$
	(.002)	(.002)
Death Year	$-.006^{***}$	
	(.0001)	
Controls		
Family Tree	*N*	*Y*
Noble Rank	*N*	*Y*
Non-Linear Time Trend	*N*	*Y*
Constant	6.350^{***}	-1.232^{***}
	(.294)	(.230)
N	76,403	76,403
Log Likelihood	-17,982.360	-17,342.070
Akaike Inf. Crit.	35,976.720	34,804.140
Note:	*p<.1; **p<.05; ***p<.01	

Notes: $D_{Violent}$ is estimated via the birthday probability exercise. The model is estimated for only those observations with an exact day of death. Source: Noble sample.

Table 5: Determinants of a Violent Death, Logistic regression

plague remained in Europe until at least 1815[18]. Despite the vast difference in death rates between the Black Death and modern outbreaks of bubonic plague, genetic and molecular testing has revealed that the cause of both is the bacillus *Yersinia Pestis* (Raoult et al. (2000); Haensch et al. (2010); Schuenemann et al. (2011)).

An simple index of mortality is constructed for the family tree sample via the formula:

$$M_t = \frac{Nd_t}{(\sum_{i=1}^{5} Nd_{t-i}+Nd_{t+i})/10} \tag{5}$$

Where M_t is an index of mortality in year t. Nd is the number of deaths in a given year. A value of one means that mortality is exactly equal to a moving window of the annual average for 5 years before and after. Figure 6 plots this simple index from 1200-1800. The plague era is immediately obvious. There are no years in the 13th century where M_t exceeds 2. After the arrival of the Black Death there are 8 years when the number of deaths is over double what we would expect: 1349, 1361, 1369, 1415, 1513, 1563, 1603 and 1625 and 5 years when M_t is over 1.5: 1375, 1471, 1540, 1593 and 1665[19]. All of these years, apart from 1415, correspond to well known plagues (Biraben (1975)). The battle of Agincourt took place on Friday, October 1415.

There is nothing in the family tree data to suggest that the Black Death or any subsequent pestilence killed anything like the proportions claimed for the rest of the population. For comparison, testaments were 21 times normal in urban Tuscany and Umbria in 1438, burials were almost 30 times normal in Siena during the plague year of 1363 Cohn (2002, p.201-2). The highest levels of crisis mortality are significantly below those for 17th century London, where plague mortality has recently been measured at 5-6 times normal mortality and never killed more than 20-25% of the city's population (Cummins et al. (2013)). Further, the simple mortality index suggests that plague mortality was consistent from the Black Death until its disappearance; consistently shocking mortality to between 1.5 to 2 times normal for this elite section of Europe's population.

Quantitative examination of Black Death mortality reveals a strong uptick in deaths during early summer with a peak in late summer (See chapter 7 in Cohn (2002))[20]. This summer peak in deaths is a distinctive marker of bubonic plague. Cummins, Kelly and Ó Gráda use this to track the disappearance of Plague mortality in London. The summer peak marker persist in poorer parishes of the city until 1720, over 60 years after the 1665 plague. (See figure 23 in the appendix for the seasonality of deaths in London and rural England from parish burial registers (Cummins et al. (2013)).

[18]According to Cohn (2008), the last Western European plague was at Noja (near Bari in Italy) in 1815. Plague may have persisted in Eastern Europe until 1897.

[19]There are three years in the 13th century where $M_t > 1.5$; 1230, 1242 and 1265. It is not clear (by eyeballing the individual observations) why mortality spikes in these years (it may be attributable to heaping and small numbers).

[20]See also Biraben (1975) and Gras (1939). For the seasonal pattern of later plagues see Slack (1977) and Schofield (1977).

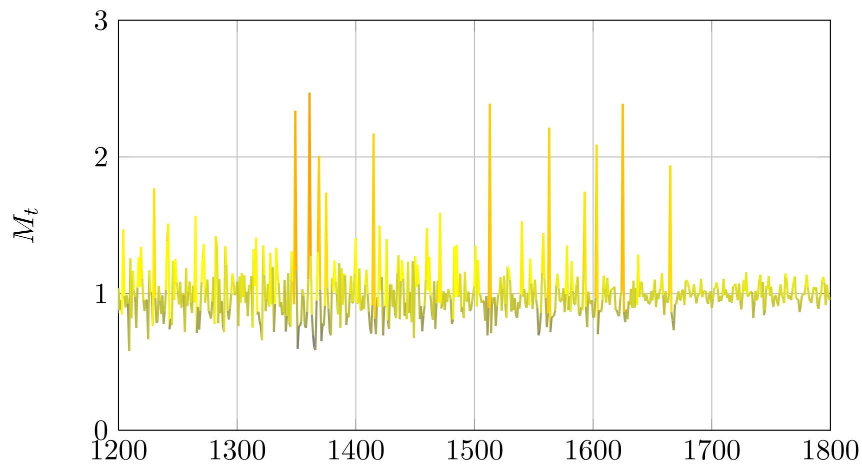


Figure 6: Index of Mortality, 1200-1800

Notes: Figure reports an index of mortality where 1 equals normal mortality. The plague era, 1346-1700, is associated with more mortality spikes than 1200-1346, or post-1700. Source: Noble sample.

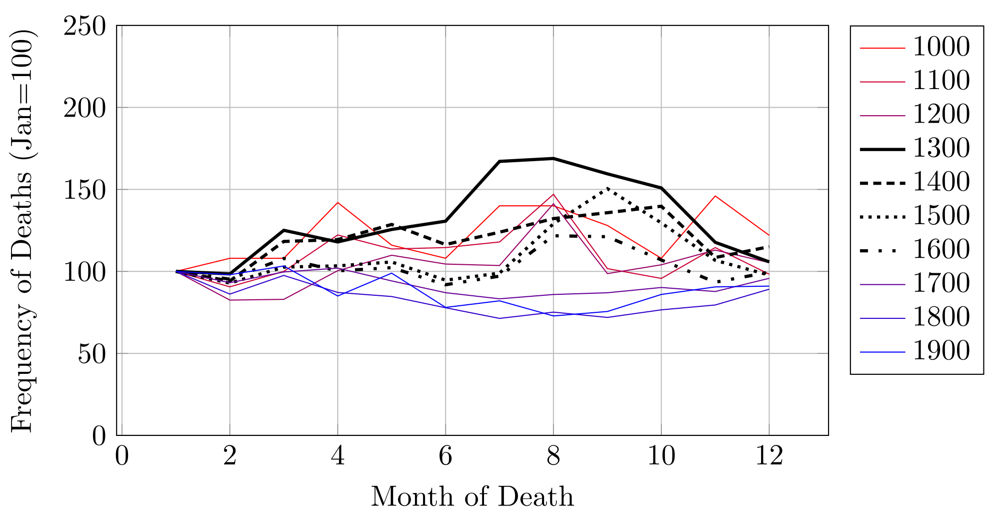


Figure 7: Seasonality of Deaths, by Century

Notes: Only records with the season of death recorded are used here. Plague typically hits in late summer. Source: Noble sample.

Figure 7 reports the frequency of month of death for those records where the month of death is recorded (data quality codes 1 and 2). The impact of the second global plague pandemic, starting with the Black Death in 1346 is evident from the changing seasonality of deaths. In particular the 1300s have a distinctive summer peak. Over time, the trend is for a disappearance of this summer peak. From the 18th-20th centuries, the summer months record significantly less deaths than the Winter months.

A well know stylized fact of the Black Death was its indiscriminate nature. King Death cut equally from the rich and the poor, men and women, old and young. Cohn (2002, p.213) quotes the chronicler of Cologne, the monk Albert:

> there was no disparity in sex or age, taking men, women, the old, the young, plebs and nobles, paupers, the rich and powerful, priests and the laity

I have shown above that the family tree data suggest otherwise: Black Death mortality was mild relative to that estimated for the rest of the population. However, how did plague discriminate within the noble family tree data? Further, was plague more virulent in the east relative to the West? Using the distinctive summer peak of plague deaths and the fact that the plague years are well known, a dummy variable (D_{plague}) was coded for death during a plague period (June-September) in one of the plague years listed above. A logistic regression was run of the form;

$$D_{Plague} = C + D_{Female} + Lat + Long + D_{Bastard} + \sum_{i=1}^{17} D_{NobleRank} \tag{6}$$

Where the notation is as equation 1. The data is restricted to the boundaries of the significant plague years recorded in the data; 1346-1666. Further, the model is run twice; once including a control for a violent death and once for those deaths that were not violent. (The likely misattribution was the concern here.)

The results are detailed in table 6. There are no consistent geographic or noble rank effects (the coefficients and standard errors for the 17 noble ranks are reported in the appendix; table 13). The Black Death and subsequent plagues were discriminately indiscriminate. Surprisingly, it appears that women faced an increased probability of a plague death. The result is significant and large in all model formulations. The lowest estimated effect in model two suggest that women faced an odds ratio of 1.16 relative to men for the risk of dying in one of the major plague seasons[21]. (The misattribuition of some violent deaths at Agincourt in 1415 to plague could only be expected to bias the analysis against this result.) The regression also indicates that older people were more likely to die during the summer plague season.

[21] Odds ratios were calculated using Fernihough (2011). For D_{Female}, they are 1.18 (.05), 1.16 (.06) and 1.26 (.09) for columns 1-3 respectively in the results table (standard errors in parentheses).

		Dependent variable:	
		D_{Plague}	
	(1)	(2)	(3)
Death Year	$-.004^{***}$	$-.010^{***}$	$-.011^{***}$
	(.001)	(.001)	(.001)
D_{Female}	$.167^{***}$	$.148^{***}$	$.228^{***}$
	(.045)	(.053)	(.075)
Age_D			$.009^{***}$
			(.002)
$D_{Bastard}$	.239	$-.324$	$-.148$
	(.686)	(1.027)	(1.051)
Latitude	$.042^{**}$	$-.013$	.002
	(.018)	(.015)	(.026)
Longitude	$-.011^{*}$	$-.003$	.002
	(.006)	(.008)	(.013)
$D_{Violent}$	3.482^{***}		
	(.051)		
Constant	$-.860$	12.238^{***}	12.810^{***}
	(1.219)	(1.176)	(1.785)
Controls			
Family Tree	*Y*	*Y*	*Y*
Noble Rank	*Y*	*Y*	*Y*
Non-Linear Time Trend	*Y*	*Y*	*Y*
N	35,318	31,498	17,350
Log Likelihood	-7,948.902	-5,94.242	-3,03.829
Akaike Inf. Crit.	15,969.800	11,95.480	6,125.657
Note:			*p<.1; **p<.05; ***p<.01

Table 6: Logistic Regression on Plague Deaths, 1346-1666

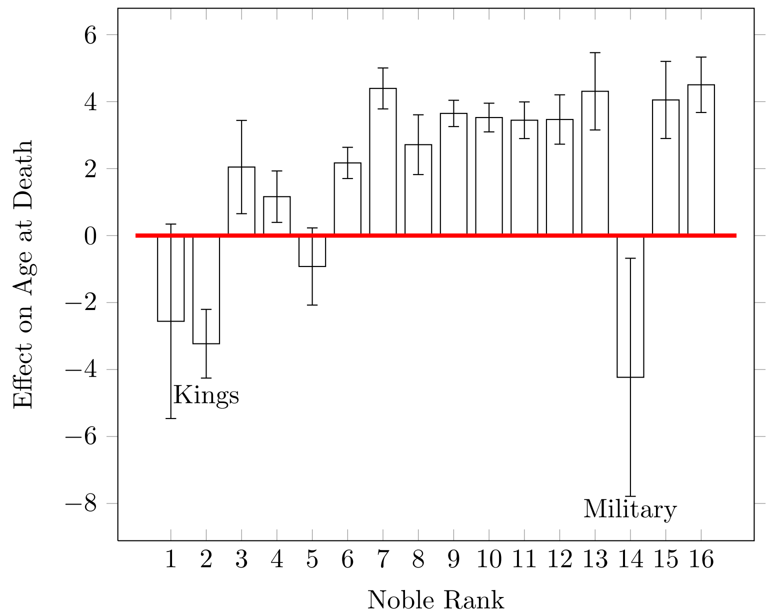


Figure 8: Occupational Hazards

Notes: Values are from the OLS estimation of equation 1. Error bars indicate 95% confidence intervals. Source: Noble sample.

4.3 Trends in Lifespan Over Time

4.3.1 Trends in the conditional, controlled Mean: OLS

Table 7 reports the results of an OLS regression of equation 1 on the noble sample (column 2)[22]. The assigned dummy for a likely violent death ($D_{Violent}$) has a strong negative effect. Geography matters too, with a strong effect of latitude and longitude on age at death. The noble rank effects are reported in figure 8. In the main individuals with a noble suffix are more likely to die older (almost 4 years older for ranks 7-13,15 and 16). Unsurprisingly, those with a military suffix die younger (although the standard errors here are large). Kings die around 3 years younger than non-nobles.

Figure 9 reports the expected age at death for each 20 year birth period from 800 to 1800. The 95% confidence intervals are too wide to allow over-interpretation of any trends in noble longevity before 1400 but after 1400 there does appear to be a sudden and sharp uptick in noble longevity - From a mean of around 50 to 54. After 1500, lifespan seems to decline until around 1650 where an uninterrupted rise begins. Noble lifespan exhibits significant oscillations across the millennium of the Dark ages to the Early modern period.

[22]Column 1 reports a version of equation 1 without any noble, data quality or family tree control variables and a a linear time trend.

	Dependent variable:	
	Age_D	
	(1)	(2)
Death Year	.018***	
	(.0003)	
D_{Female}	−.063	1.197***
	(.112)	(.119)
Latitude	.098***	.061***
	(.008)	(.008)
Longitude	−.029***	−.026***
	(.003)	(.003)
$D_{Violent}$	−7.194***	−7.585***
	(.456)	(.456)
$D_{Bastard}$	.064	−1.822**
	(.891)	(.889)
Constant	21.937***	47.558***
	(.606)	(.872)
Controls:		
Noble Rank	*N*	*Y*
Data Quality	*N*	*Y*
Family Tree	*N*	*Y*
Non-Linear Time Trend	*N*	*Y*
N	121,478	121,478
R^2	.044	.060
Adjusted R^2	.044	.059
Residual Std. Error	18.423	18.281
F Statistic	929.736***	56.548***
Note:	*p<.1; **p<.05; ***p<.01	

Table 7: OLS Regression of Equation 1

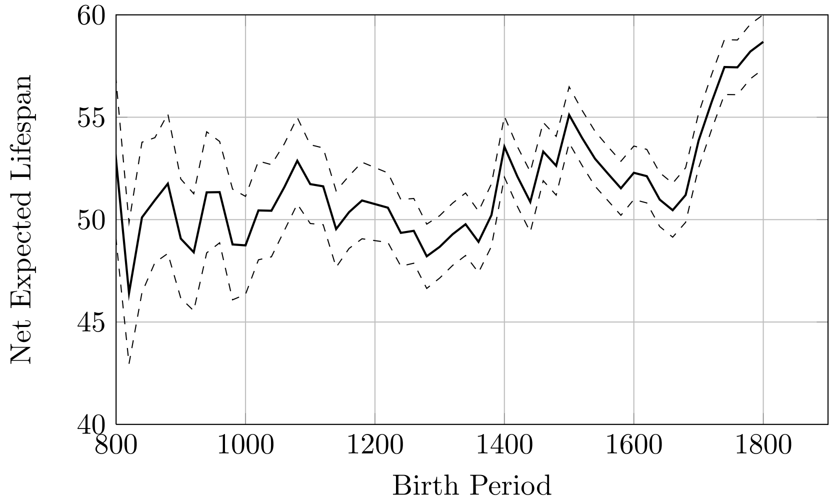


Figure 9: Expected Lifespan 800-1800

Notes: Expected values are from the birth period coefficents estimated by OLS estimation (equation 1). Error bands indicate 95% confidence intervals. Average longitude and latiude is applied (50.09, -.015), all other controls are set to 0. Source: Noble sample.

4.3.2 Trends across the distribution: Quantile Regression

Table 8 reports the quantile regression results for equation 1 for the .1, .25, .5, .75 and .9 percentiles of the age at death distribution. The non linear effects of sex are picked up by this approach. The coefficient on the female dummy variable (D_{Female}) is negative at younger ages and turns significantly positive as the percentile increases. Geography matters at every quantile of the distribution.

Figure 10 reports the expected age at death for each of the percentiles reported in table 8. The pattern agrees broadly with that indicated by OLS: there is a sharp rise in noble longevity after 1400. The time patterns are broadly consistent across the age at death distribution: increases from 1300-1400, decline from 1500-1600 and a strong rise after 1650[23].

4.4 Trends in Lifespan across space and time

The results from table 7 and 8 suggest a European Mortality Pattern. Figure *11* plots heat-maps of median lifespan (over 20) by geographic coordinates. The pattern is easiest to detect in panel (b) where the median is calculated over integer longitude and altitude coordinates. Length of noble life follows a strong South-North, East-West gradient.

Is this European Mortality Pattern constant over time? Table 9 estimates equation 1

[23]The coefficient estimates and their respective standard errors are reported in the appendix.

	.1	.25	.5	.75	.9
D_{Female}	-.451**	-.477**	1.436***	2.345***	2.668***
	(.179)	(.200)	(.171)	(.148)	(.153)
Latitude	.043***	.065***	.078***	.041***	.024**
	(.012)	(.013)	(.011)	(.010)	(.010)
Longitude	-.008	-.020***	-.033***	-.035***	-.027***
	(.005)	(.006)	(.005)	(.004)	(.004)
$D_{Violent}$	-4.753***	-8.122***	-10.298***	-7.310***	-4.583***
	(.685)	(.763)	(.652)	(.566)	(.582)
$D_{Bastard}$	-1.765	-3.230**	-1.736	-.816	-2.470**
	(1.336)	(1.487)	(1.271)	(1.104)	(1.135)
Constant	27.989***	38.032***	46.348***	59.957***	71.679***
	(1.909)	(2.125)	(1.816)	(1.578)	(1.622)
Controls:					
Noble Rank	Y	Y	Y	Y	Y
Data Quality	Y	Y	Y	Y	Y
Family Tree	Y	Y	Y	Y	Y
Non-Linear Time Trend	Y	Y	Y	Y	Y
N	121,478	121,478	121,478	121,478	121,478
Pseudo R^2	.025	.037	.045	.041	.032

Standard errors in parentheses
*** p<.01, ** p<.05, * p<.1

Table 8: Quantile Regression of Equation 1

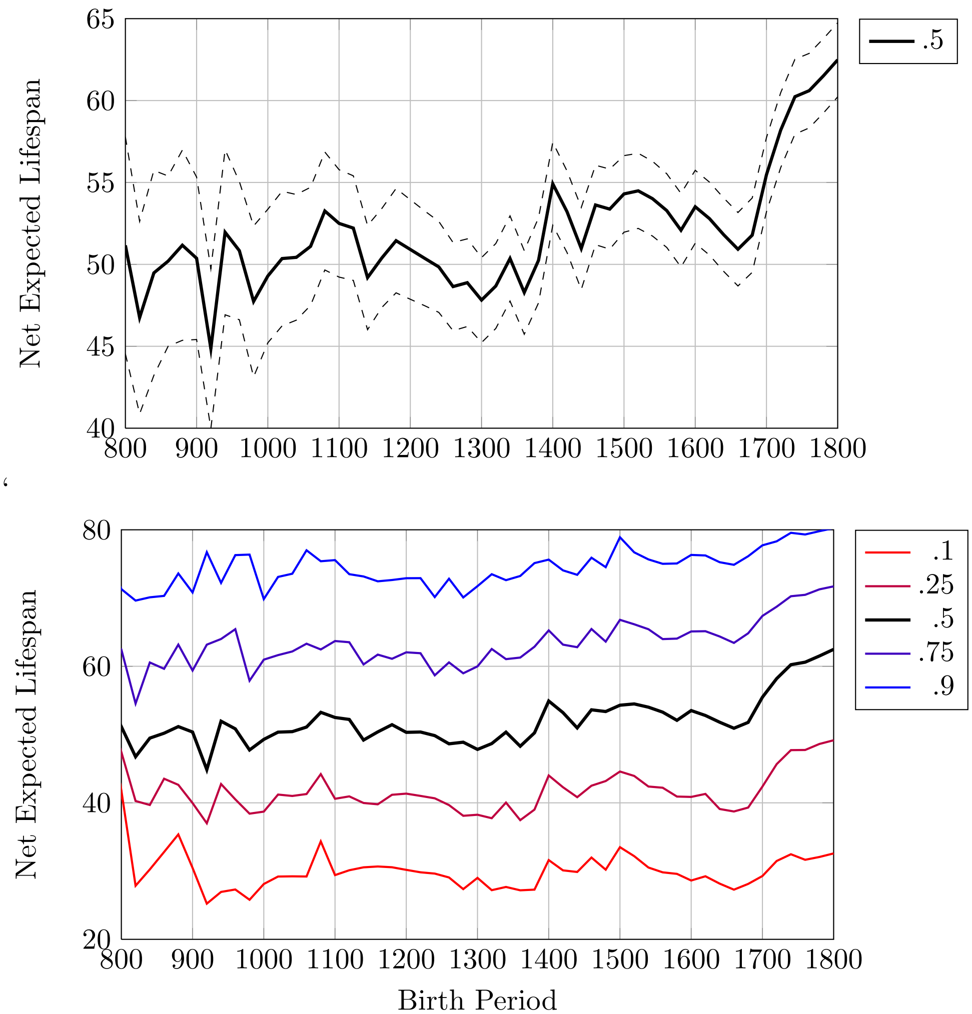
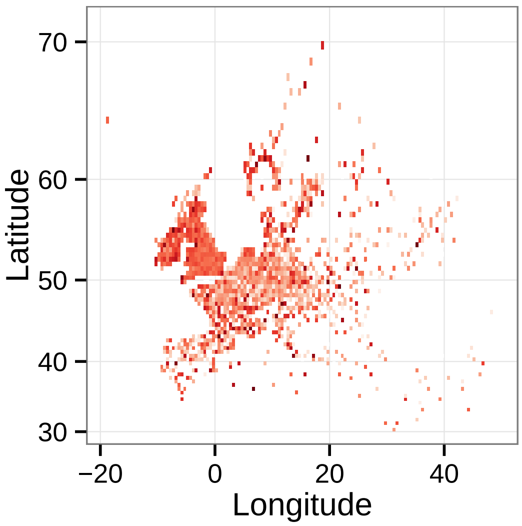
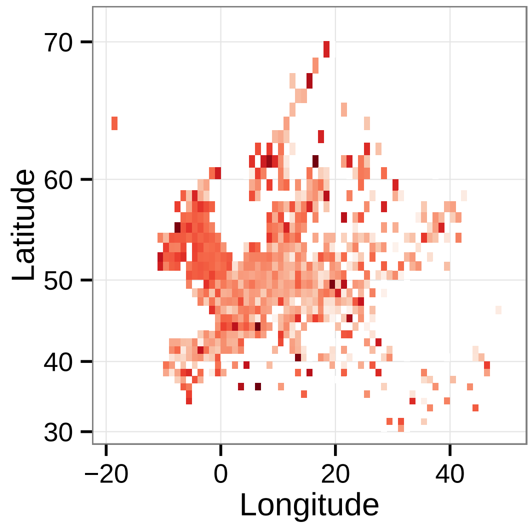


Figure 10: Expected Lifespan 800-1800, by Quantile

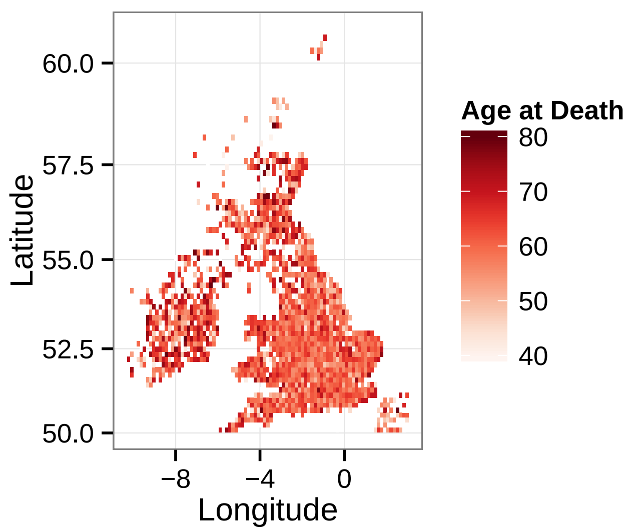
Notes: Expected values are from the birth period coefficents estimated by quantile regression (equation 1). Error bands indicate 95% confidence intervals. Average longitude and latiude is applied (50.09, -.015), all other controls are set to 0. Source: Noble sample.



(a) Europe, .5 Degree Resolution

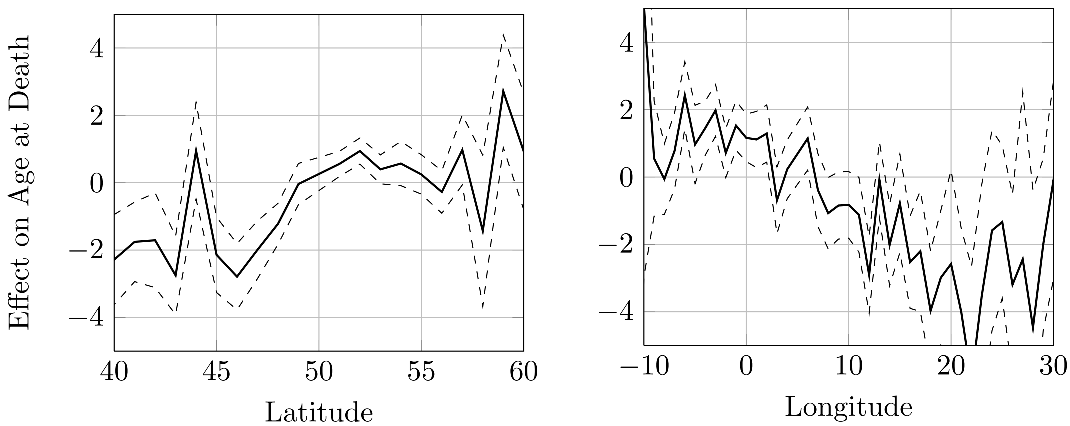


(b) Europe, 1 Degree Resolution



(c) Ireland and the UK, .15 Degree Resolution

Figure 11: The European Mortality Pattern: Heat-maps of Median Age at Death



(a) South–North

(b) West–East

Figure 12: Controlled Geographic Effects

by sub-period. The standard errors on latitude and longitude are too large for us to be sure of any real effect before the first millennium but the coefficient estimates are suggestive. Geography matters in all periods after 1000. It is notable that in the pre-Black Death era this European Mortality Pattern is present.

Is the effect of geography completely linear? To answer this, equation 2 was estimated (allowing dummy values for each integer value of longitude and latitude in the data). This fully controlled effect of geography, along with 95% confidence intervals is reported in figure 12. The effect is broadly linear.

4.4.1 Time Trends in Noble Lifespan by Region

Are there different time-trends in noble lifespan in different regions of Europe? A simple geographic bounding box is calculated for each region of Europe where the data is dense enough to allow analysis by 20 year birth period. Seven 'regions' numbered in order from North to South are drawn in figure 13 and details are reported in table 10.

A BART model is estimated based on equation 3. Figure 14 and 15 report predictions of adult noble lifespan for each of these seven regions, based upon their average latitude and longitude[24]. Predicted lifespan is stationary everywhere before 1400, where just as with the OLS and quantile regression estimates (see figures 9 and 10), lifespan suddenly rises. There is heterogeneity within Europe after 1400 however; in Scotland, Ireland and England and Wales, lifespan rises from 1400 to 1500. Everywhere in Europe, the modern

[24]The top three variables in the BART model (by average variable inclusion proportions across 100 models) are birth year, longitude and latitude. See figure 25 in the appendix.

	Dependent variable:					
	Age at Death (>= 20)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Pre 1000	1000-1340	1340-1500	1500-1600	1600-1700	post 1700
Latitude	.053	.118***	.156***	.170***	.115***	.094***
	(.078)	(.032)	(.023)	(.022)	(.023)	(.010)
Longitude	−.034	−.114***	−.096***	−.087***	−.071***	−.005
	(.045)	(.017)	(.013)	(.011)	(.009)	(.005)
D_{Female}	1.846	.899**	.333	−.286	−.282	3.644***
	(1.274)	(.427)	(.305)	(.240)	(.210)	(.195)
$D_{Violent}$	−1.560	−8.885***	−8.282***	−9.405***	−9.643***	−4.172***
	(4.344)	(1.181)	(.769)	(.712)	(.717)	(.966)
$D_{Bastard}$	−2.852	−4.594***	−2.596**	−2.993**	−3.048*	.774
	(3.601)	(1.661)	(1.227)	(1.239)	(1.561)	(3.819)
Constant	48.303***	41.096***	42.143***	42.949***	47.032***	50.331***
	(4.877)	(1.770)	(1.255)	(1.160)	(1.190)	(.598)
Controls						
Noble Rank	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
Family Tree	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
Data Quality	*Y*	*Y*	*Y*	*Y*	*Y*	*Y*
N	1,690	10,050	19,570	30,592	41,211	46,648
R^2	.033	.044	.042	.047	.036	.034
Adjusted R^2	.017	.041	.041	.046	.035	.033
Note:						*p<.1; **p<.05; ***p<.01

Table 9: Estimating by Sub periods

		Longitude		Latitude			Avg.	
	Region	Min	Max	Min	Max	N	Longitude	Latitude
1	North and NorthEastern Europe	2	50	52	75	6,987	54.59	12.3
2	Scotland	-8	-1	55	59	8,279	56.16	−3.59
3	Ireland	-11	-6	51	56	4,875	53.15	−7.49
4	England and Wales	-6	2	50	55	73,489	52.09	−1.41
5	France	-6	8	43	52	13,680	49.19	3.44
6	Central and Eastern Europe	8	50	44	52	8,111	49.24	13.19
7	Southern Europe	-11	50	36	44	2,262	40.87	4.75

Table 10: Regions

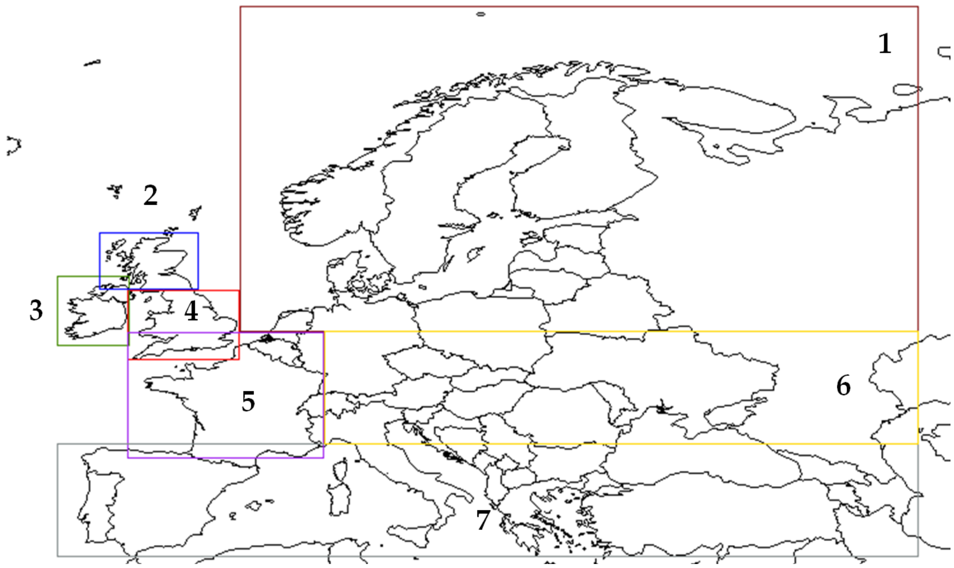
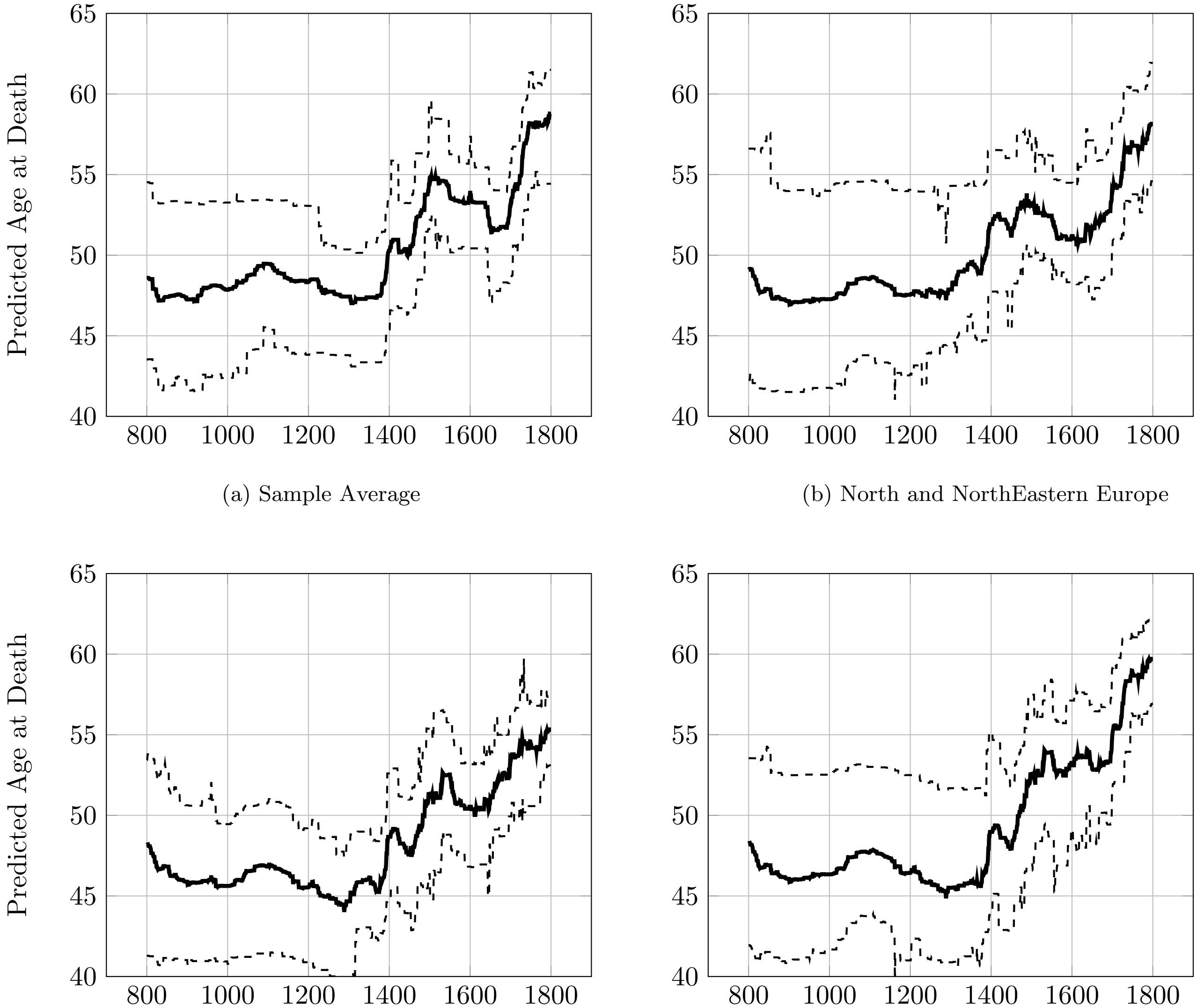


Figure 13: Regions used for Separate Quantile Regressions

Notes: Overlapping regions were assigned to the Northern most region.

increase of longevity originates around 1650 for European nobility. The impact of the thirty years war is evident in figure 15. The three models agree on the path of noble longevity between 800 and 1800.



(a) Sample Average

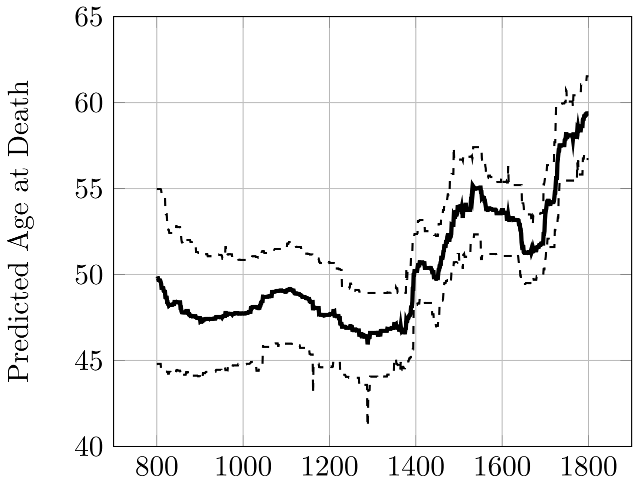
(b) North and NorthEastern Europe

(c) Scotland

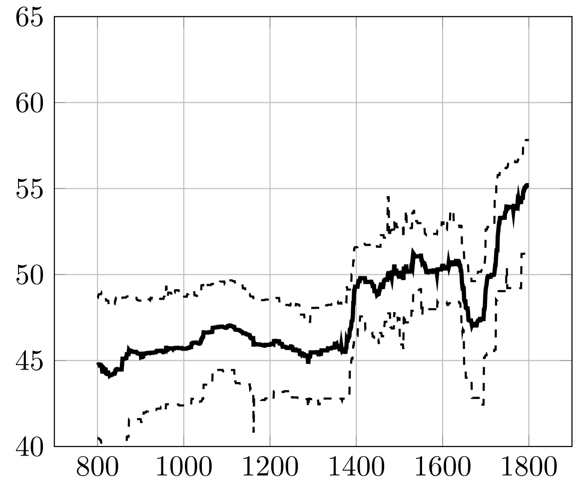
(d) Ireland

Figure 14: BART Machine predictions for Adult Longevity, by Region (Sample avg. and 1-3)

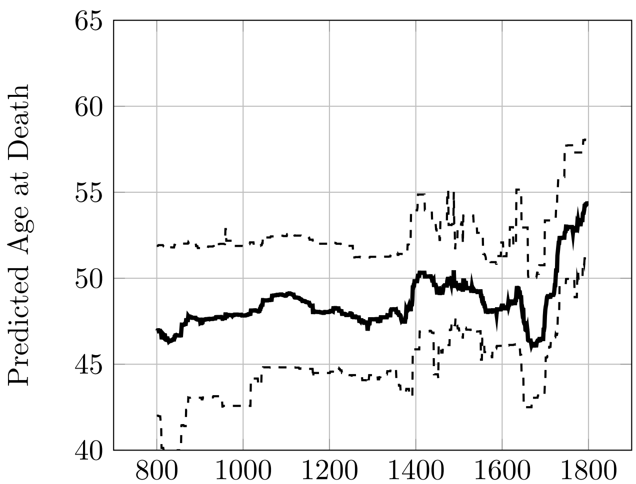
Notes: Values are predicted values from a BART model, N=151,324 with 40 predictors, 200 trees, 250 burn-in and 1000 post. samples, Pseudo $R^2 = .103$. The best model was chosen via a gridsearch over a set of hyperparameter combinations, including the number of trees, m. $m = 200$ gave the best results. The three most important variables in the BART model are birth year, longitude and latitude. To understand the raw trends in the data, uncontrolled quantile regresion forest predictions are reported in the appendix.



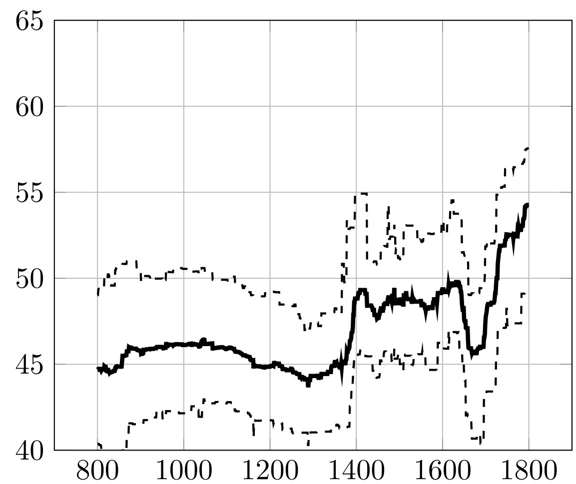
(a) England and Wales



(b) France



(c) Central and Eastern Europe



(d) Southern Europe

Figure 15: BART Machine predictions for Adult Longevity, by Region (4-7)

Notes: See notes to the previous figure for details on the construction of the predictions.

5 Discussion

This study has characterized noble lifespans from 800 to 1800. The results have many implications. Firstly, the sharp decline in the proportion of male nobles dying from violence, from at least 600 years of a steady 30% to less than 5% in the 16th century, predates the arrival of the Industrial Revolution by two centuries. The long run decline in violence[25] is cited as one of the principal correlates of the emergence of the modern World. Why did violence decline among European nobility? Was it a 'bottom-up' behavioral change (perhaps as a result of natural selection, as Clark (2007) suggests for the general population) or was it a response to changing 'top-down' institutional incentives (as argued by Acemoglu and Robinson (2012))? This speaks to the core of modern debates about the wealth of nations.

Long before the decline of violence there are significant changes in noble longevity. I have examined the impact of plague on Europe's elites. The consistent and large association uncovered between sex and plague mortality runs counter to the indiscriminate reputation of the Black Death and counter to recent paleodemographic analysis on skeletons from 14th century London (DeWitte (2009)[26]). Perhaps sex differentials in historical plague mortality strengthened women's position in the marriage market? Could a simple supply-side effect explain the origin of the European Marriage Pattern? (Hajnal (1965), Voigtländer and Voth (2013)). These questions will be pursued in future research[27].

Thirdly, this paper estimates the time-trend of noble lifespan over the millennium between 800 and 1800. The findings on the timing of the modern rise in age at death agree almost exactly with de la Croix and Licandro (2012) (the birth cohort of 1640-9). The nobility are forerunners of Europe's mortality transition (as David et al. (2010) argue too). This provides an important clue for those who seek to explain exactly why mortality declined. There could be an important role for individual behavior and a demonstration effect (e.g. hygiene and other behavioral traits) as this rise predates modern medicine or any public health measures. It also predates the Industrial Revolution[28]. Whilst modern evidence suggests that life expectancy does not matter for economic growth (Acemoglu and Johnson (2007)), the case has not been proved for the preindustrial era. Is this rise in age at death also evident for other groups within the population as well?

However, unlike de la Croix and Licandro (2012), this study argues that lifespan was not a stationary trend before 1650. There are significant oscillations, most importantly the sharp Europe-wide rise in lifespan after 1400. The rise is stronger over the 1400-1600 interval in Ireland, Scotland and in particular, England and Wales (figures 14 and 15).

[25] As evidenced by Gurr (1981), Eisner (2003), Clark (2007) and popularized recently by Pinker (2011).

[26] However, the fact that older people faced a higher probability of death from plague is consistent with DeWitte (2010).

[27] Cummins et al. (2013) details the construction of dataset of 3.5m deaths from plague era London that could be used to further explore this.

[28] It is also striking how the post 1600 rise is weaker in Central and Eastern Europe, and Southern Europe (figures 14 and 15).

This pattern has remained hidden as only long and deep time series of at least a millennia in length could uncover this.

Why did noble lifespan increase so much after 1400? It was probably not because of a Black Death 'survivor' effect as plague mortality was relatively low amongst the nobility. Absent a previously unnoticed medical revolution of the 15th century, this rise, as with the later rise of 1650, must reflect some change in individual behavior[29].

Finally, this paper documents a previously unknown European mortality pattern, Similar to that for marriage first documented by Hajnal (1965), the mortality gradient runs South-North and East-West, and has existed since before the Black Death[30]. The long existence of such a geographic effect has implications for recent work which stresses the 'little divergence' between the North-West of Europe and the South-East (Voigtländer and Voth (2013), Broadberry (2013) and de Pleijt and van Zanden (2013)). The Black Death is not the first turning point. There was something about the North-West of Europe long before 1346 that led to nobles living longer lives.

These results suggests that the 'Rise of the West' does not solely originate in institutional innovations of the 17th century (Acemoglu and Robinson (2012)) nor in social reactions to the Black Death (Voigtländer and Voth (2013)). Western exceptionalism exists in individual behavior differences that are present since at least the first millennium AD.

6 Conclusions

This paper makes four principle contributions. Violence declines for nobles in the 16th century, plague kills noble women at a higher rate than men. There is a structural break in noble lifespan about 1400 and there is a European mortality pattern that has existed since the year 1000. The 'Rise of the West' can be traced to the centuries before the Black Death. These new stylized facts may or may not only apply to this elite subgroup. Future research can test whether the patterns are more general.

References

Acemoglu, Daron and James Robinson, *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, Crown Publishing Group, 2012. 1, 29, 30

_ **and Simon Johnson**, "Disease and Development: The Effect of Life Expectancy on Economic Growth," *Journal of Political Economy*, December 2007, *115* (6), 925–985. 29

[29]This change, a generation after the Black Death, coincides with Voigtländer and Voth's dating of the origin of the European Marriage Pattern (2013).

[30]This pattern appears to have been completely unknown to previous research; a Google scholar search of' "European marriage pattern" yields 1,410 results while a search of "European mortality pattern" yields 16 (scholar.google.co.uk [23 June 2014]).

Becker, Gary S., Kevin M. Murphy, and Robert Tamura, "Human Capital, Fertility, and Economic Growth," *The Journal of Political Economy*, 1990, *98* (5), S12–S37. 2

Benedictow, Ole Jørgen, *The Black Death, 1346-1353: the complete history*, Boydell & Brewer, 2004. 12

Biraben, Jean-Noël, *Les hommes et la peste en France et dans les pays européens et méditerranéens*, Vol. 2, Mouton Paris-La Haye, 1975. 14

Boyd, Percival, *Pedigrees with index of London citizens, abt. 1600-1800*, Manuscript, filmed by the Genealogical Society of Utah, 1954. 3

Breiman, Leo, "Random forests," *Machine learning*, 2001, *45* (1), 5–32. 48

Broadberry, Stephen, "Accounting for the great divergence," Economic History Working Papers 54573, London School of Economics and Political Science, Department of Economic History November 2013. 1, 30

Burke, Bernard and John Bernard Burke, *The Book of Precedence: The Peers, Baronets, and Knights, and the Companions of the Several Orders of Knighthood, Placed According to Their Relative Rank, Together with a Scale of General Or Social Precedence*, Harrison, 1881. 4

_ **and John Burke**, *A Genealogical and Heraldic Dictionary of the Peerage and Baronetage: Together with Memoirs of the Privy Councillors and Knights*, Harrison, 1881. 3

Chipman, Hugh A, Edward I George, Robert E McCulloch et al., "BART: Bayesian additive regression trees," *The Annals of Applied Statistics*, 2010, *4* (1), 266–298. 9

Clark, Gregory, *A Farewell to Alms: A Brief Economic History of the World*, Princeton, NJ:: Princeton University Press, 2007. 1, 2, 29

Cohn, Samuel K, *The Black Death transformed: disease and culture in early Renaissance Europe*, Arnold, 2002. 14, 16

_ , "4 Epidemiology of the Black Death and Successive Waves of Plague," *Medical history. Supplement*, 2008, (27), 74. 14

Cummins, Neil, "Stand Alone Appendix to "The Economic History of European Nobility Project"," 2014. Available at neilcummins.com. 2, 4, 5, 7

_ **, Morgan Kelly, and Cormac Ó Gráda**, "Living Standards and Plague in London, 1560-1665," Working Papers 201308, School Of Economics, University College Dublin Jul 2013. 14, 29, 43

Cutler, David, Angus Deaton, and Adriana Lleras-Muney, "The Determinants of Mortality," *The Journal of Economic Perspectives*, 2006, *20* (3), pp. 97–120. 2

David, Paul, S. Ryan Johansson, and Andrea Pozzi, "The Demography of an Early Mortality Transition: Life Expectancy, Survival and Mortality Rates for Britain's Royals, 1500-1799," Economics Series Working Papers Number 83, University of Oxford, Department of Economics 2010. 2, 29

de la Croix, David and Omar Licandro, "The Longevity of Famous People from Hammurabi to Einstein," Working Papers 666, Barcelona Graduate School of Economics October 2012. 2, 29

de Pleijt, Alexandra M. and Jan Luiten van Zanden, "Accounting for the 'Little Divergence' What drove economic growth in preindustrial Europe, 1300-1800?," Working Papers 0046, Utrecht University, Centre for Global Economic History 2013. 30

de Sainte-Marie, Anselme and Ange de Sainte-Rosalie, *Histoire généalogique et chronologique de la Maison Royale de France, des pairs, des grands officiers de la Couronne & de la Maison du Roy: & des anciens barons du Royaume... Par le P. Anselme,... continuée par M. Du Fourny. Troisième édition, revûë, corrigée & augmentée par les soins du P. Ange & du P. Simplicien...*, par la compagnie des libraires associez, 1728. 3

DeWitte, Sharon N, "The effect of sex on risk of mortality during the Black Death in London, AD 1349–1350," *American journal of physical anthropology*, 2009, *139* (2), 222–234. 29

_ , "Age patterns of mortality during the Black Death in London, AD 1349–1350," *Journal of archaeological science*, 2010, *37* (12), 3394–3400. 29

Doyle, William, *Aristocracy: a very short introduction*, Oxford University Press, 2010. 4

Eisner, Manuel, "Long-term historical trends in violent crime," *Crime and Justice*, 2003, pp. 83–142. 29

Fernihough, Alan, "Simple logit and probit marginal effects in R," 2011. 16

Fire, Michael and Yuval Elovici, "Data Mining of Online Genealogy Datasets for Revealing Lifespan Patterns in Human Population," *arXiv preprint arXiv:1311.4276*, 2013. 2

Galor, Oded, "From Stagnation to Growth: Unified Growth Theory," 2004. CEPR Discussion Papers 4581. 2

Gras, Pierre, "Le registre paroissial de Givry (1334-1357) et la peste noire en Bourgogne," *Bibliothèque de l'École des Chartes*, 1939, *100* (1), 295–308. 14

Green, Donald P and Holger L Kern, "Modeling heterogeneous treatment effects in survey experiments with Bayesian additive regression trees," *Public Opinion Quarterly*, 2012, p. nfs036. 9

Gurr, Ted Robert, "Historical trends in violent crime: A critical review of the evidence," *Crime and Justice*, 1981, pp. 295–353. 29

Haensch, Stephanie, Raffaella Bianucci, Michel Signoli, Minoarisoa Rajerison, Michael Schultz, Sacha Kacki, Marco Vermunt, Darlene A Weston, Derek Hurst, Mark Achtman et al., "Distinct clones of Yersinia pestis caused the black death," *PLoS pathogens*, 2010, *6* (10), e1001134. 14

Hajnal, John, "European marriage patterns in perspective," in Eversley DE Glass DV, ed., *Population in history: essays in historical demography*, 1965. 29, 30

Hastie, Trevor, Robert Tibshirani, Jerome Friedman, T Hastie, J Friedman, and R Tibshirani, *The elements of statistical learning*, Vol. 2, Springer, 2009. 48

Henry, Louis, "Fécondité des mariages dans le quart sud-ouest de la France, de 1720 à 1829," *Annales. Histoire, Sciences Sociales*, 1972, *27* (3), pp. 612–640. 2

_ , "Fécondité des mariages dans le quart Sud-Est de la France de 1670 a 1829," *Population (French Edition)*, 1978, *33* (4/5), 855–883. 2

_ **and Jacques Houdaille**, "Fécondité des mariages dans le quart nord-ouest de la France de 1670 a 1829," *Population (French Edition)*, 1973, *28* (4/5), 873–924. 2

Hollingsworth, Thomas H, "A demographic study of the British ducal families," *Population Studies*, 1957, *11* (1), 4–26. 2, 12

_ , *The demography of the British peerage*, Population investigation committee, 1964. 2

_ , "A note on the mediaeval longevity of the Secular peerage 1350–1500," *Population Studies*, 1975, *29* (1), 155–159. 2

_ , "Mortality in the British peerage families since 1600," *Population (French Edition)*, 1977, pp. 323–352. 2

Houdaille, Jacques, "Fécondité des mariages dans le quart nord-est de la France de 1670 a 1829.," in "Annales de Demographie Historique" 1976, pp. 341–392. 2

Kapelner, Adam and Justin Bleich, "Bartmachine: A powerful tool for machine learning," *arXiv preprint arXiv:1312.2171*, 2013. 9

Levy, Claude and Louis Henry, "Ducs et pairs sous l'Ancien Régime. Caractéristiques démographiques d'une caste," *Population (French Edition)*, 1960, pp. 807–830. 2

Meinshausen, Nicolai, "Quantile regression forests," *The Journal of Machine Learning Research*, 2006, *7*, 983–999. 48

_ , "quantregForest: quantile regression forests," *R package version 0.2-3*, 2007. 48

Paget, Gerald, *An Official, Genealogical, and Heraldic Baronage of England. An account of the ancient nobility of this realm from the time of the Norman Conquest till the close of the fifteenth century* 1957. 3

Peller, Sigismund, "Births and deaths among Europe's ruling families since 1500," *Population in History*, 1965, pp. 87–100. 2

Pinker, Steven, *The better angels of our nature: The decline of violence in history and its causes*, Penguin UK, 2011. 29

Raoult, Didier, Gérard Aboudharam, Eric Crubézy, Georges Larrouy, Bertrand Ludes, and Michel Drancourt, "Molecular identification by "suicide PCR" of Yersinia pestis as the agent of medieval black death," *Proceedings of the National Academy of Sciences*, 2000, *97* (23), 12800–12803. 14

Schofield, Roger, "An anatomy of an epidemic: Colyton November 1645 to November 1646," *Local Population Studies*, 1977, (Supplement 4), 95–126. 14

_ **, David Reher, and Alain Bideau, eds**, *The decline of mortality in Europe.*, Oxford England Clarendon Press 1991., 1991. 2

Schuenemann, Verena J, Kirsten Bos, Sharon DeWitte, Sarah Schmedes, Joslyn Jamieson, Alissa Mittnik, Stephen Forrest, Brian K Coombes, James W Wood, David JD Earn et al., "Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of Yersinia pestis from victims of the Black Death," *Proceedings of the National Academy of Sciences*, 2011, *108* (38), E746–E752. 14

Schwennicke, Detlev, *Europäische Stammtafeln: neue Folge. Die fränkischen Könige und die Könige und Kaiser, Stammesherzoge, Kurfürsten, Markgrafen und Herzoge des Heiligen Römischen Reiches Deutscher Nation*, Vol. 1, Vittorio Klostermann, 2005. 3

Slack, Paul, "The local incidence of epidemic disease: the case of Bristol 1540-1650," *Local Population Studies*, 1977, (Supplement 4), 49–62. 14
The Holy Bible King James Version

***The Holy Bible King James Version**, Online Version 2014. 3*

Voigtländer, Nico and Hans-Joachim Voth*, "How the West "Invented" Fertility Restriction," American Economic Review, 2013, 103 (6), 2227–64. 1, 29, 30*

Wrigley, Edward. A. and Roger S. Schofield, *The Population History of England 1541-1871, Cambridge Universtiy Press, 1981. 2, 43*

Wrigley, Edward A., Ros S. Davies, Oeppen James E., and Roger S. Schofield, *English population history from family reconstitution, 1580-1837, Cambridge: Cambridge University Press, 1997. 2*

A For Online Publication: Distributions, Supplementary Regression and Analysis Tables

List of Tables

List of Figures

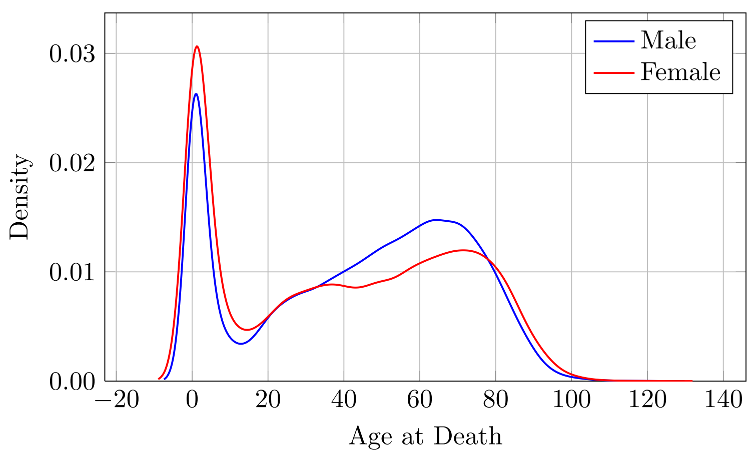


Figure 16: Distribution of Age at Death, All

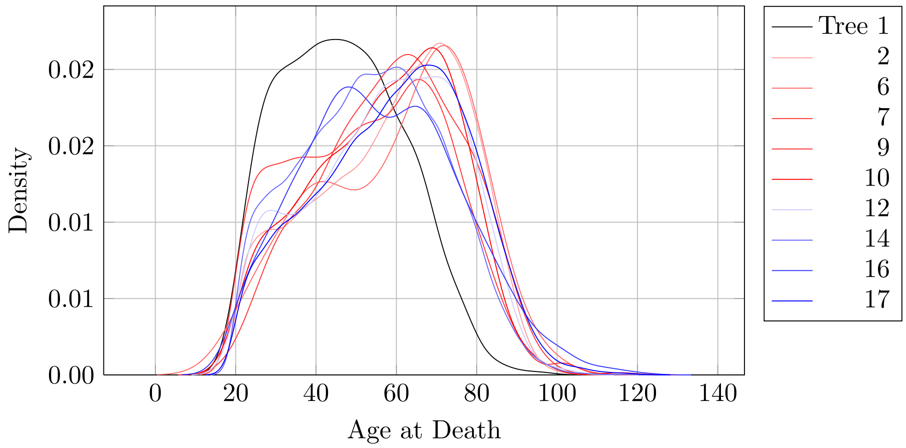


Figure 18: Distribution of Male Age at Death, by Tree

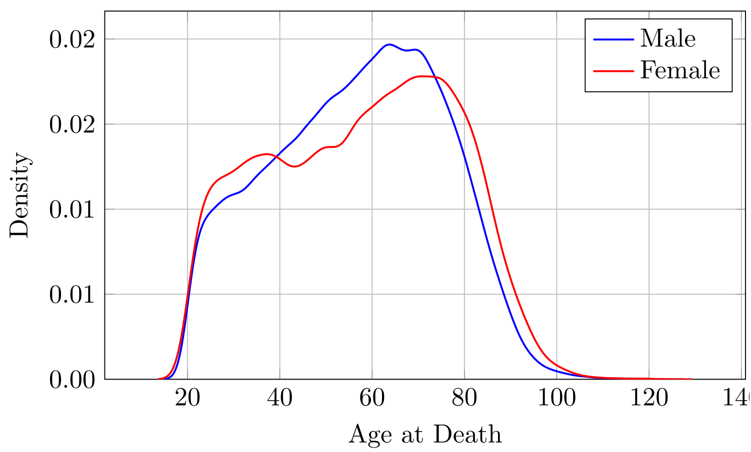


Figure 17: Distribution of Age at Death, over 20

.

A.1 Supplementary Regression/Analysis Tables

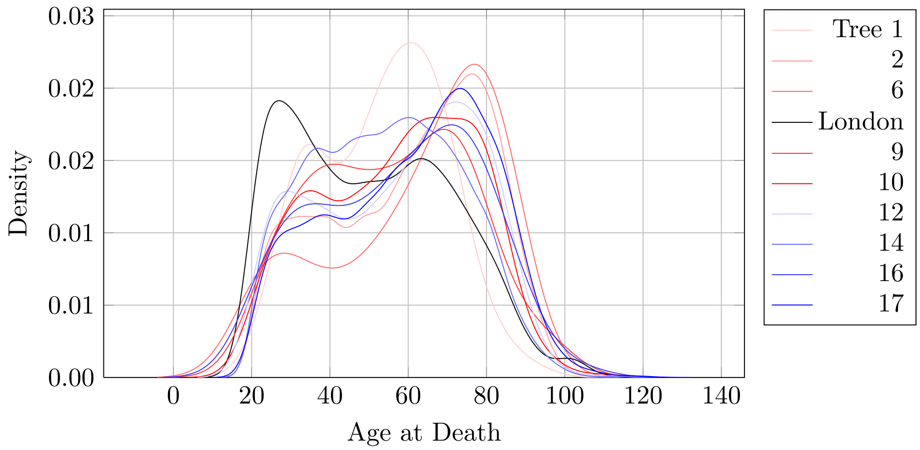


Figure 19: Distribution of Female Age at Death, by Tree

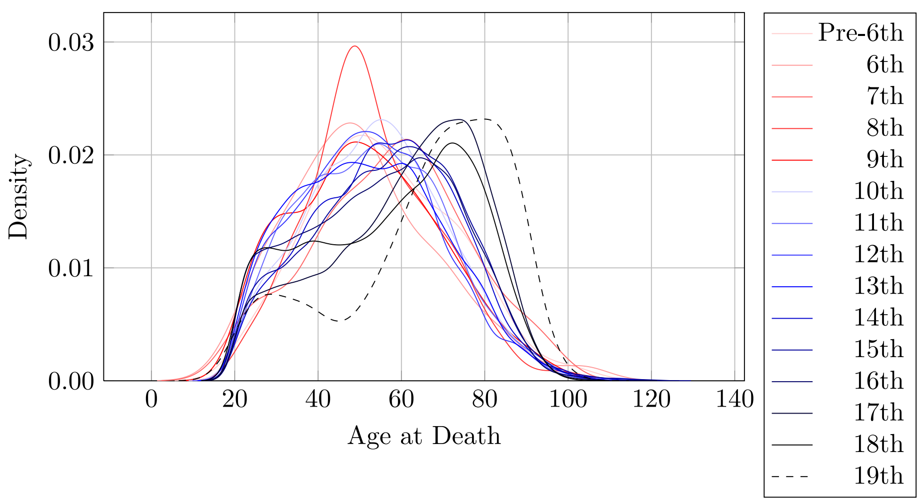


Figure 20: Distribution of Male Age at Death, by Birth Century

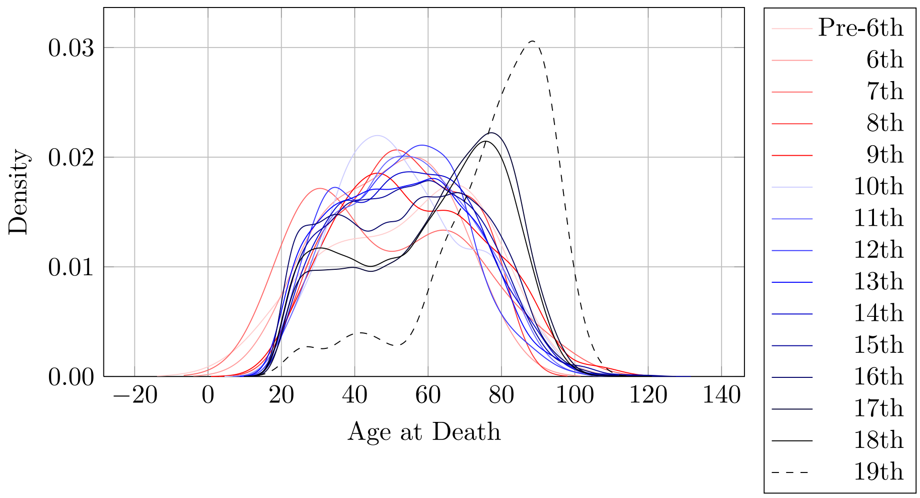


Figure 21: Distribution of Female Age at Death, by Birth Century

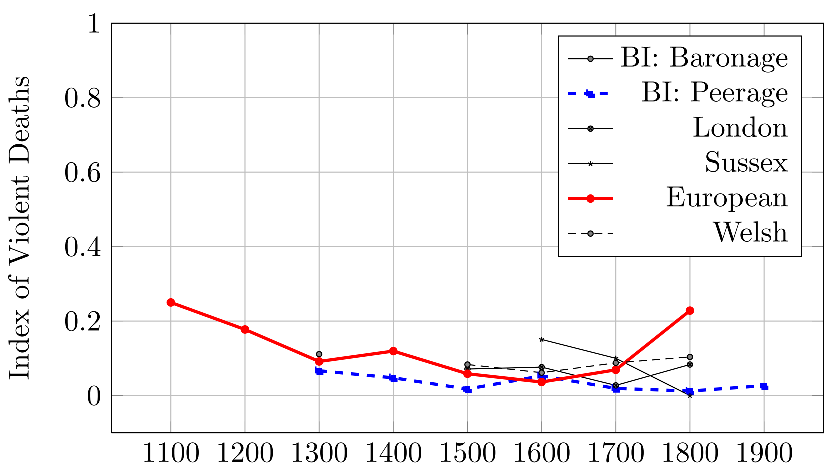


Figure 22: Index of Violent Deaths, Female

	Dependent variable:	
	violent	
	(1)	(2)
Emperor		−1.227
		(1.066)
King		0.055
		(0.214)
Grand Duke, ArchDuke, Ancient		−0.264
		(0.412)
Duke		−0.102
		(0.167)
Prince-Elector, Prince		0.025
		(0.255)
Earl, Count		0.116
		(0.092)
Marquess, Margrave		0.698***
		(0.123)
Viscount		0.100
		(0.224)
Baron, Lord		0.465***
		(0.074)
Baronet		−0.422***
		(0.151)
Knight		0.212**
		(0.107)
Esquire, Gentleman and unassigned nobility		0.050
		(0.161)
Geographic		1.028***
		(0.163)
Military		0.607
		(0.597)
Religious		−0.629**
		(0.248)
Occupational		−0.332**
		(0.164)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 11: Noble Rank Correlations with a Violent Death, Logistic Regression

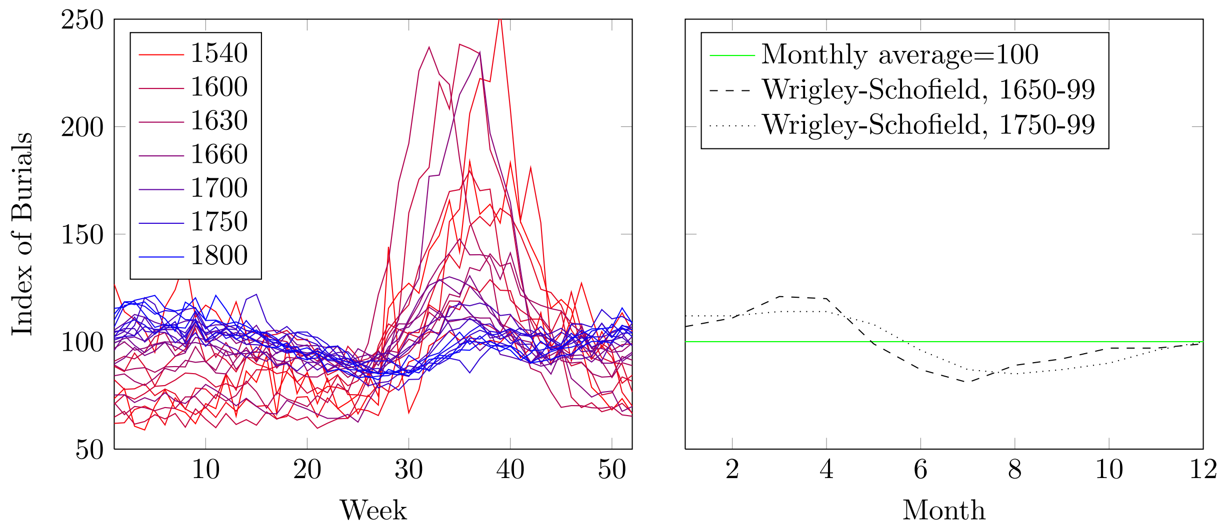
	Dependent variable:
	agedeath
Birth Date Quality 2	-1.610^{***}
	(0.469)
Birth Date Quality 3	2.464^{***}
	(0.138)
Birth Date Quality 4	1.858^{***}
	(0.166)
Death Date Quality 2	-2.400^{***}
	(0.279)
Death Date Quality 3	-1.262^{***}
	(0.145)
Death Date Quality 4	-0.055
	(0.341)
Tree ID 1	-4.941^{***}
	(0.668)
Tree ID 2	2.990^{***}
	(0.168)
Tree ID 6	2.198
	(1.443)
Tree ID 7	-1.437^{***}
	(0.238)
Tree ID 9	4.598^{***}
	(1.029)
Tree ID 10	2.001^{***}
	(0.585)
Tree ID 12	2.478^{***}
	(0.384)
Tree ID 16	3.221^{***}
	(0.649)
Tree ID 17	4.027^{***}
	(0.179)
Note:	*p<0.1; **p<0.05; ***p<0.01

Notes: The ommited categories are 1 (birth and death quality types) and 14 (Tree ID). See the stand alone appendix for detail.

Table 12: OLS Control Results

	Dependent variable:		
	plague2		
	(1)	(2)	(3)
Emperor	−12.416	−12.148	−12.867
	(1,016.254)	(1,348.173)	(2,229.220)
King	−.196	−.700	−.548
	(.668)	(1.044)	(1.049)
Grand Duke, Arch Duke, Ancient	−.545	−12.159	−13.011
	(1.153)	(423.822)	(716.361)
Duke	−1.048*	−.269	−.120
	(.628)	(.601)	(.613)
Prince-Elector, Prince	−12.666	−12.427	−13.345
	(22.285)	(268.398)	(456.692)
Earl, Count	−.221	−.250	−.166
	(.248)	(.342)	(.415)
Marquess, Margrave	−.807*	−.223	.372
	(.472)	(.744)	(.769)
Viscount	−1.214	−.226	−13.673
	(1.057)	(1.025)	(708.390)
Baron, Lord	.145	.377*	.410
	(.170)	(.225)	(.316)
Baronet	−13.155	−12.285	−13.130
	(20.716)	(237.966)	(453.584)
Knight	−.745***	−.374	−1.211*
	(.254)	(.355)	(.727)
Esquire, Gentleman and unassigned nobility	−.263	.280	.053
	(.417)	(.474)	(.613)
Geographic	2.041***	.388	−12.851
	(.268)	(.728)	(552.573)
Military	1.688	1.512	2.787**
	(1.190)	(1.088)	(1.223)
Religious	0.597	0.250	0.299
	(0.565)	(0.531)	(0.752)
Occupational	0.686**	0.167	0.600
	(0.325)	(0.437)	(0.612)
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 13: Noble Rank Plague Death Correlations



(a) London

(b) Rural England

Figure 23: Seasonal Distribution of Deaths in England, 1540-1800

Notes: Sources: Cummins et al. (2013) and Wrigley and Schofield (1981).

Table 14: Quantile Time Dummies

	Quantile				
	.1	.25	.5	.75	.9
800	11.647***	6.097	0.836	0.424	-1.594
(67)	(3.551)	(3.952)	(3.378)	(2.935)	(3.018)
820	-2.33	-1.08	-3.574	-7.508***	-3.276
(88)	(3.150)	(3.506)	(2.997)	(2.604)	(2.677)
840	0.064	-1.661	-0.846	-1.489	-2.797
(76)	(3.354)	(3.733)	(3.191)	(2.773)	(2.851)
860	2.614	2.185	-0.128	-2.414	-2.589
(117)	(2.787)	(3.101)	(2.651)	(2.304)	(2.368)
880	5.191*	1.289	0.844	1.131	0.695
(90)	(3.113)	(3.465)	(2.962)	(2.574)	(2.646)
900	0.287	-1.381	0.046	-2.669	-2.088
(131)	(2.658)	(2.959)	(2.529)	(2.198)	(2.259)
920	-4.951*	-4.341	-5.464**	1.123	3.816*
(137)	(2.611)	(2.906)	(2.484)	(2.158)	(2.219)
940	-3.234	1.407	1.634	1.952	-0.684
(126)	(2.701)	(3.006)	(2.569)	(2.233)	(2.295)
960	-2.877	-0.84	0.512	3.391*	3.386*
(198)	(2.261)	(2.516)	(2.151)	(1.869)	(1.922)
980	-4.383*	-2.933	-2.587	-4.155**	3.460*
(157)	(2.470)	(2.749)	(2.350)	(2.042)	(2.099)
1000	-2.074	-2.635	-1.035	-1.073	-3.060*
(217)	(2.186)	(2.433)	(2.079)	(1.807)	(1.858)
1020	-0.98	-0.135	0.03	-0.41	0.191
(212)	(2.203)	(2.451)	(2.095)	(1.821)	(1.872)
1040	-0.95	-0.341	0.104	0.125	0.658
(257)	(2.056)	(2.289)	(1.956)	(1.700)	(1.748)
1060	-0.974	-0.059	0.775	1.263	4.094**
(308)	(1.934)	(2.152)	(1.840)	(1.598)	(1.643)
1080	4.167**	2.862	2.933	0.418	2.487
(310)	(1.930)	(2.148)	(1.837)	(1.596)	(1.641)

Standard errors in parentheses

*** p<.01, ** p<.05, * p<.1

Continued on next page

Table 14 – continued from previous page					
	Quantile				
	.1	.25	.5	.75	.9
1100	-0.769	-0.753	2.176	1.654	2.653*
(419)	(1.757)	(1.956)	(1.672)	(1.453)	(1.494)
1120	-0.058	-0.402	1.89	1.478	0.591
(459)	(1.712)	(1.905)	(1.629)	(1.415)	(1.455)
1140	0.369	-1.369	-1.131	-1.773	0.258
(466)	(1.705)	(1.898)	(1.622)	(1.409)	(1.449)
1160	0.493	-1.564	0.064	-0.335	-0.454
(564)	(1.620)	(1.803)	(1.541)	(1.339)	(1.377)
1180	0.379	-0.167	1.122	-0.947	-0.261
(464)	(1.707)	(1.900)	(1.624)	(1.411)	(1.451)
1220	-0.359	-0.333	0.049	-0.132	0.011
(675)	(1.551)	(1.727)	(1.476)	(1.282)	(1.319)
1240	-0.547	-0.685	-0.48	-3.370***	-2.763**
(820)	(1.489)	(1.657)	(1.416)	(1.230)	(1.265)
1260	-1.117	-1.675	-1.681	-1.475	-0.077
(955)	(1.444)	(1.607)	(1.373)	(1.193)	(1.227)
1280	-2.823**	-3.246**	-1.442	-3.072***	-2.817**
(986)	(1.435)	(1.597)	(1.365)	(1.186)	(1.219)
1300	-1.177	-3.092**	-2.504*	-2.058*	-1.161
(1175)	(1.391)	(1.549)	(1.324)	(1.150)	(1.183)
1320	-2.976**	-3.600**	-1.647	0.489	0.588
(1177)	(1.390)	(1.547)	(1.322)	(1.149)	(1.181)
1340	-2.506*	-1.311	0.034	-0.999	-0.285
(1154)	(1.394)	(1.552)	(1.327)	(1.153)	(1.185)
1360	-2.997**	-3.886**	-2.038	-0.768	0.328
(1301)	(1.368)	(1.522)	(1.301)	(1.131)	(1.162)
1380	-2.899**	-2.34	-0.079	0.82	2.216*
(1163)	(1.393)	(1.550)	(1.325)	(1.151)	(1.183)
1400	1.431	2.663*	4.592***	3.213***	2.726**
(1363)	(1.358)	(1.511)	(1.292)	(1.122)	(1.154)
1420	-0.075	0.94	2.891**	1.125	1.150
(1533)	(1.335)	(1.485)	(1.270)	(1.103)	(1.134)
1440	-0.31	-0.52	0.644	0.751	0.499

Standard errors in parentheses
*** p<.01, ** p<.05, * p<.1

Continued on next page

Table 14 – continued from previous page					
	Quantile				
	.1	.25	.5	.75	.9
(1468)	(1.343)	(1.495)	(1.278)	(1.110)	(1.141)
1460	1.803	1.161	3.300***	3.415***	3.006***
(1889)	(1.300)	(1.447)	(1.237)	(1.075)	(1.105)
1480	0.024	1.845	3.048**	1.567	1.619
(1776)	(1.311)	(1.459)	(1.247)	(1.084)	(1.114)
1500	3.324***	3.239**	3.975***	4.759***	6.002***
(2720)	(1.253)	(1.395)	(1.192)	(1.036)	(1.065)
1520	2.007	2.592*	4.164***	4.108***	3.792***
(3485)	(1.230)	(1.369)	(1.170)	(1.017)	(1.045)
1540	0.34	1.041	3.686***	3.384***	2.755***
(3825)	(1.224)	(1.362)	(1.164)	(1.012)	(1.040)
1560	-0.364	0.866	2.958***	1.938*	2.110**
(5009)	(1.206)	(1.342)	(1.147)	(0.997)	(1.025)
1580	-0.589	-0.428	1.759	2.008**	2.161**
(5287)	(1.205)	(1.341)	(1.147)	(0.996)	(1.024)
1600	-1.561	-0.481	3.192***	3.058***	3.412***
(6674)	(1.193)	(1.328)	(1.135)	(0.986)	(1.014)
1620	-0.94	-0.041	2.468**	3.092***	3.334***
(6814)	(1.194)	(1.329)	(1.136)	(0.987)	(1.015)
1640	-2.018*	-2.252*	1.484	2.320**	2.345**
(6310)	(1.201)	(1.337)	(1.143)	(0.993)	(1.021)
1660	-2.903**	-2.614*	0.599	1.379	1.965*
(6224)	(1.202)	(1.338)	(1.144)	(0.994)	(1.022)
1680	-2.060*	-2.039	1.458	2.788***	3.221***
(5126)	(1.217)	(1.354)	(1.158)	(1.006)	(1.034)
1700	-0.903	1.034	5.149***	5.330***	4.818***
(5295)	(1.215)	(1.352)	(1.156)	(1.004)	(1.033)
1720	1.299	4.320***	7.871***	6.681***	5.405***
(4687)	(1.226)	(1.365)	(1.167)	(1.014)	(1.042)
1740	2.285*	6.373***	9.910***	8.215***	6.652***
(4816)	(1.226)	(1.364)	(1.166)	(1.013)	(1.042)
1760	1.474	6.395***	10.278***	8.423***	6.408***
(5515)	(1.220)	(1.358)	(1.160)	(1.008)	(1.037)

Standard errors in parentheses

*** p<.01, ** p<.05, * p<.1

Continued on next page

Table 14 – continued from previous page

	Quantile				
	.1	.25	.5	.75	.9
1780	1.886	7.276***	11.185***	9.224***	6.891***
(5900)	(1.218)	(1.355)	(1.159)	(1.007)	(1.035)
1800	2.407**	7.819***	12.160***	9.663***	7.333***
(6954)	(1.212)	(1.349)	(1.153)	(1.002)	(1.030)

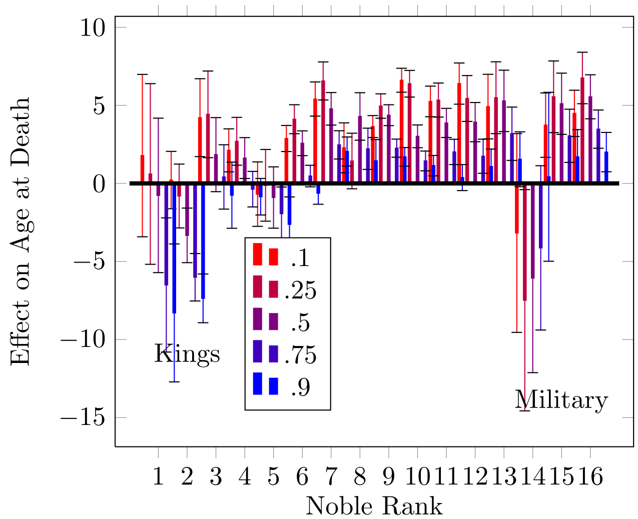


Figure 24: Coefficient on Noble Rank Dummy by Quantile, with 95% CI

A.1.1 Missing the Forest for the Trees

To understand the raw trends in the data this section uses Quantile Regression Forests (Meinshausen (2006)), which build upon Random Forests; Random Forests were introduced by Breiman (2001) and have the advantage of making no distributional assumptions, handling complex non-linear interactions and having error rates that have been shown to compete very well with alternative methods[31] (Hastie et al. (2009, p.590)).

Quantile Regression Forests are a robust non-parametric method to estimate the conditional distribution of a response variable[32]. Figures 26 and 27 plot annual estimates for the .1, .25, median, .75 and .9 quantile for the sample average and the 7 regions of Europe defined in table 10[33]. Looking at the raw trends, it is apparent that the Black Death hits harder in England than anywhere else: All quantiles decrease apart from the .1 quantile who are stable. This result, for England and Ireland, that that age mattered during the Black Death is consistent with the results from table 6, reported earlier.

[31] Such as *bagging* (bootstrap aggregation) and gradient boosting.
[32] See Meinshausen (2006) for an numerical evaluation for the algorithm.
[33] Estimation is via the R package quantregforest Meinshausen (2007).

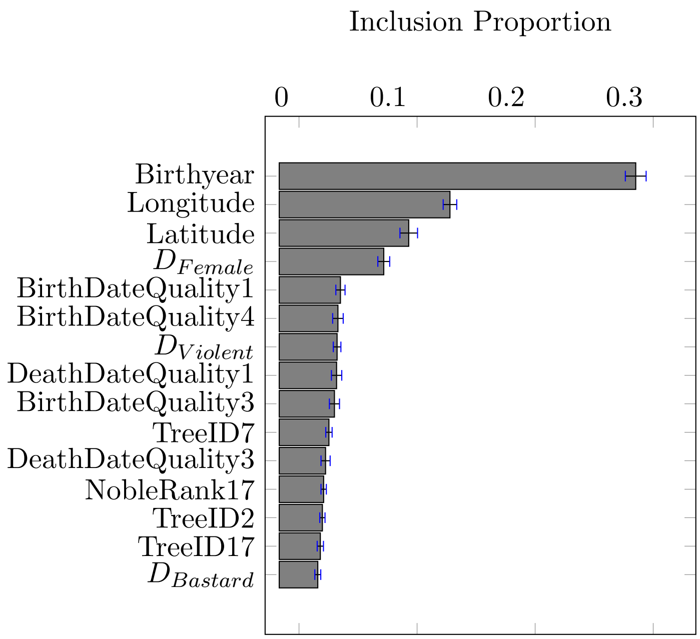
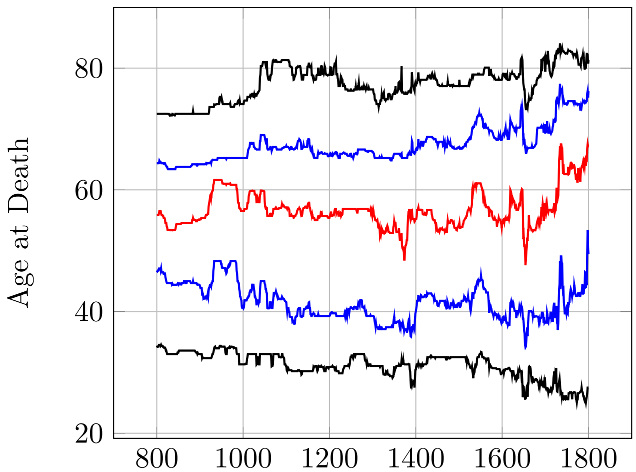
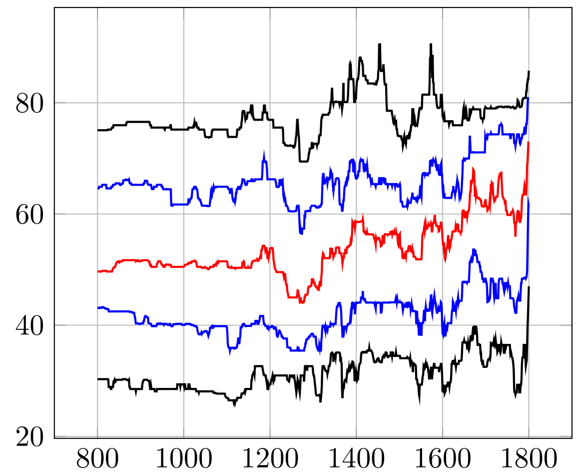


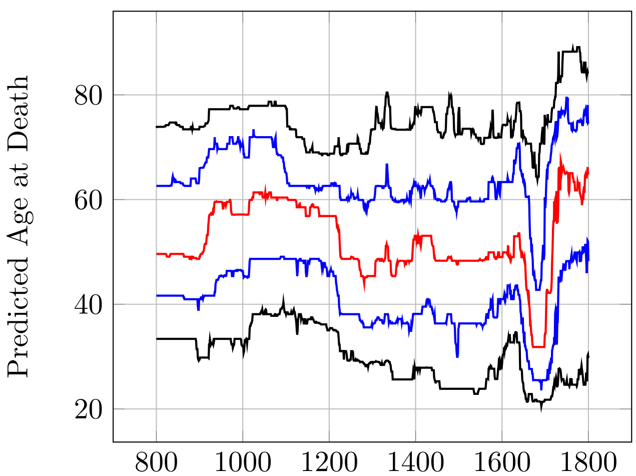
Figure 25: Top 15 Variables in BART model



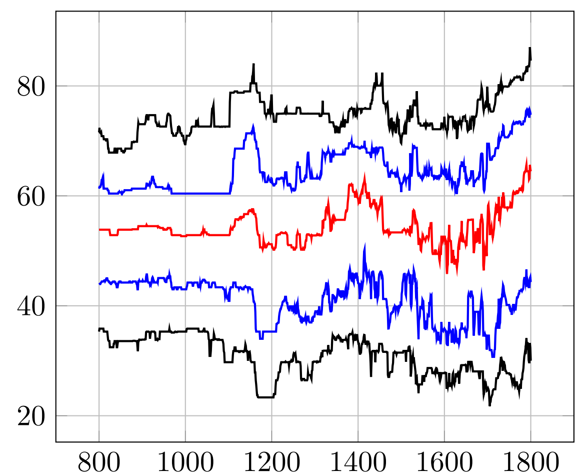
(a) Sample Average



(b) North and North-Eastern Europe

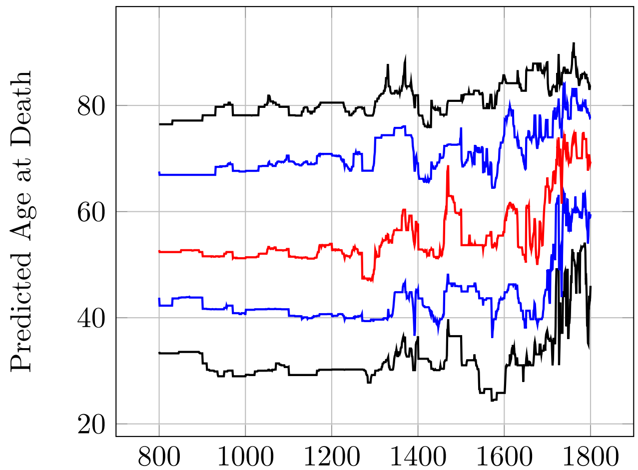


(c) Scotland

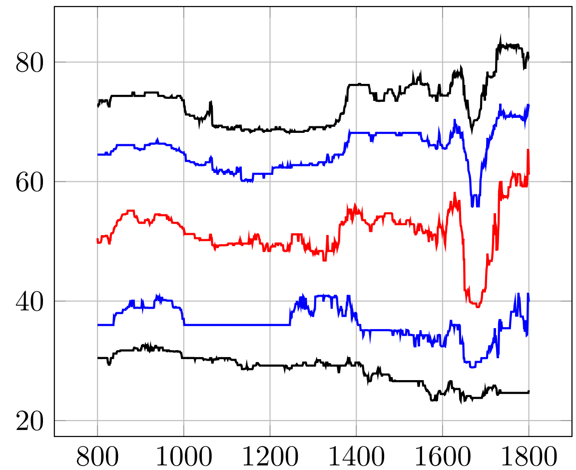


(d) Ireland

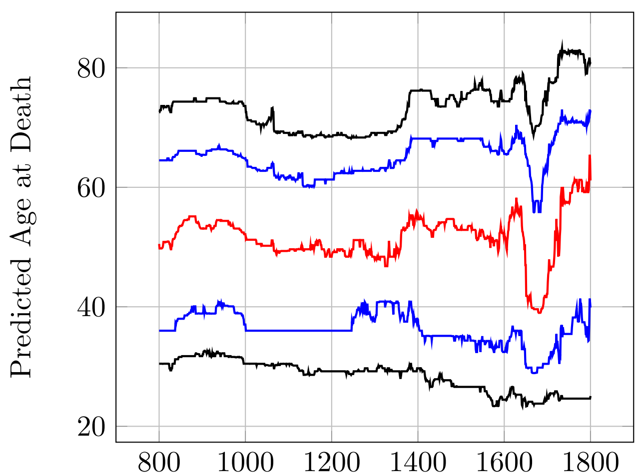
Figure 26: Quantile Regression Forests predictions for Adult Longevity, by Region (Sample average and 1-3)



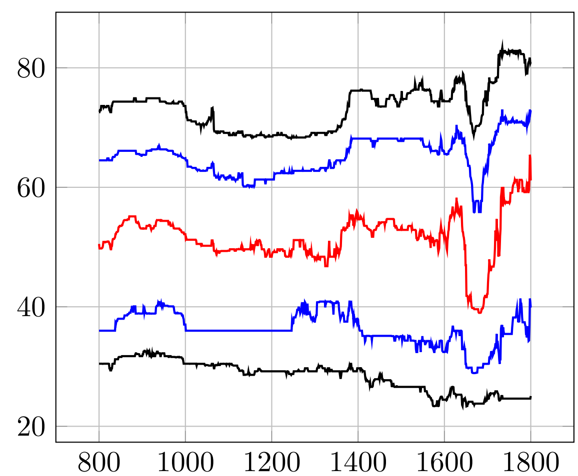
(a) England and Wales



(b) France



(c) Central and Eastern Europe



(d) Southern Europe

Figure 27: Quantile Regression Forests predictions for Adult Longevity, by Region (4-7)