New Evidence on the Anatomy of the Health-Status Gradient from Imperial China[1]

Karen Eggleston — Stanford University and NBER

Wolfgang Keller — University of Colorado, NBER, and CEPR

Carol H. Shiue — University of Colorado, NBER, and CEPR

July 5, 2025

Abstract

Using matched lifespan-status data on multiple generations, we study the evolution and reproduction of the health-status gradient in a 1675-1775 sample from Central China. Top status men live about five years longer than men with low status. Exploiting our intergenerational setting, we present new evidence on the role of health for status by showing that long-living fathers tend to have rich sons, whereas sons of rich fathers do not live longer than sons of poorer fathers. Intergenerational status persistence is much higher than intergenerational lifespan persistence, but money cannot buy much health in this society. In contrast, a long lifespan enables men to rise in social status by participating in China's civil service examination and gaining official positions. We also show that child and maternal mortality are subject to a status gradient during the sample period.

Keywords: Lifespan mobility, well-being, status persistence, Preston curve

[1] Contact: karene@stanford.edu, kellerw@colorado.edu, shiue@colorado.edu.

Ted Telford initiated the data collection underlying this work. Support from NIH is gratefully acknowledged (Shiue).

I. Introduction

The negative correlation of income with premature mortality is virtually ubiquitous -- it has been found wherever and whenever it is sought (Deaton 2016). Precisely because its emergence is obscure, fully understanding the negative income-mortality correlation, aka Preston curve, has been a challenge.[2] Using Chinese family histories spanning more than a century, this paper sheds new light on the negative income-mortality correlation by exploiting variation in intergenerational health and status persistence from earlier generations. We show that a strong Preston curve in the child generation stems primarily from parent-child health persistence, not from status persistence. This finding provides new evidence on the role of good health for high incomes.

We analyze genealogical data from central China for birth cohorts from 1675 to 1775, the early Qing dynasty. Our primary health measure is completed lifespan, while social status can be distinguished based on more than 20 standardized descriptors in four classes for a more parsimonious analysis. Status is measured primarily with the educational attainment and official position of the male head of household, based on performance in the civil service examination system, with investments, occupation, household structure, and other indicators of wealth determining status as well. Consistent with population data for similar periods, the large majority of households never attained any status. Unlike historical studies using occupation to measure status in Europe, the civil service exam is a distinctive setting where male educational attainment directly conveys social status, and this status is not hereditary, but in principle and in practice open to a wide swath of the population.

Our data provides several advantages. Compared to the cross-sectional data of most influential studies of the health gradient – from the Marmot et al. (1978) seminal study of British civil servants to Chetty et al. (2016) on the contemporary US – we not only study the evolution of the gradient over a century but also its relationship to intergenerational transmission of health and status. Relative to most widely-cited historical cohort data (e.g. the Human Mortality Database, HMD) and compared to prominent studies of intergenerational health transmission (such as Black et al. 2024), our data includes status and thus we can analyze the status gradient in intergenerational health transmission. Finally, relative to earlier studies of China, which for data availability reasons tend to focus on the elite and special settings (e.g. Lee et al. 1994, Chen and Preston 2022), our sample includes more commoners and thus captures a fuller span of the status distribution in society.

Our analysis reveals a status gradient in adult lifespan of modest magnitude, as well as gradients in child and maternal mortality. Probing the roles of intergenerational health and status transmission, we find a salient role for intergenerational health transmission. Specifically, our results present new evidence for the significant role that health plays for social status by showing that long-living fathers tend to have rich sons, whereas sons of

[2] Preston (1975).

rich fathers do not live longer than sons of poorer fathers. Intergenerational persistence in status is relatively high; however, even the rich son of a rich father cannot 'buy' himself a significantly longer lifespan in our historical setting, and evidence that status raises lifespan is limited.

In contrast, we find evidence for a role of health raising status because in the case of strong intergenerational health transmission, the long-living son of a healthy father can rise in social status by participating in and passing successive levels of China's civil service examination. A longer life also allows earning more years of income, building wealth, and investing in children's human capital. While we do not claim causality, both reduced-form and instrumental-variables results provide evidence for an effect of health on status. More broadly, our unique panel data linking intergenerational health and status mobility to the health-status gradient provides novel evidence consistent with a much stronger relationship from health to status than vice versa in Qing imperial China.

The positive health-status gradient during the sample period 1675-1775 is about five years of lifespan difference between the highest and the lowest status men. Equivalently, the chance of survival to 65 conditional on surviving to 15 (i.e., S(65|15)) was 41% for top status graduates of the civil service exam (*jinshi*) and 29% for commoners. That range includes the S(65|15) survival estimate of 35% for Swedish men born around 1750 from the Human Mortality Database, though their status is unknown. The gradient is steeper at top status levels, and it becomes significant during the 18th century.

Consistent with existing estimates for later periods, the intergenerational transmission of lifespan is relatively weak. In the well-known framework of regressing son on parent outcomes, the intergenerational coefficient for lifespan is one fourth or less of the size of the intergenerational status coefficient (about 0.15 versus 0.6, respectively). Mother lifespan plays a stronger role for son lifespan than does father lifespan, and intergenerational lifespan persistence for top status families is higher than for commoners. Also, over the sample period, both lifespan and status mobility tended to increase.

Our estimates on the extent of the health-status gradient are conservative, because if anything the vital statistics recorded in the genealogies underestimate the premature mortality of commoners. We present suggestive evidence of status being protective against the highest source of mortality throughout this historical period: infant and child mortality. Both child and maternal mortality are subject to a status gradient as well. For example, during the 1720s the son of a *jinshi* had about twice the chance to live past age 10 compared to the son of a commoner. This pattern suggests that the modest adult mortality gradient that we document reflects in part the selective survival of commoners with higher health endowment, relative to their counterparts growing up in higher-status households under less selective pressure.

The paper, first, contributes to the understanding of the strength of the health-status gradient. A large literature studies the relationship between socio-economic status (SES)

and health (overviews include Cutler, Lleras-Muney, and Vogl 2011, Deaton 2016). A frequent starting point to study the SES-health gradient is the cross-country relationship described originally by Preston (1975), though within-country analysis, as in the present paper, has produced additional insights (Chetty et al. 2016). We present novel evidence about how adult lifespans differed by social status for a time and region with little previous evidence, but with great importance for understanding the well-being component of the 'great divergence' (Pomeranz 2001) between China and the west.[3] Unlike most previous studies, our cohorts span a century, allowing us to examine the evolution and reproduction of the Preston Curve across multiple generations. We also contribute to the historical and economic literature by analyzing data with a relatively time-invariant, multiple-category metric of SES – so that we are documenting a gradient, not merely a gap between a binary stratification of rich versus poor, or elite versus commoner. As emphasized by Cutler, Vogl, and Lleras-Muney (2011) and others, the SES-health gradient reflects a bundle of factors in SES that potentially have distinct independent influences on health (e.g., short-term income fluctuations versus long-term income or wealth, or education which is not subject to short-term fluctuations). In our setting, these dimensions of status are deeply inter-connected through the social prestige of ranking on the national civil service examinations (e.g., higher educational attainment leading to stable employment, assortative matching with wealthier marriages, and bureaucratic or social power over consequential community resources). The robust if modest gradients in lifespan for married men are supplemented with evidence on differential child and maternal mortality by household status. Moreover, Qing dynasty China represents an interesting period to study because of its dramatic population growth and stagnant or declining living standards; by contrast, many previous studies focus on later periods with better vital statistics, especially cohorts living through improving living standards and overall rising life expectancy in the 20th century.

Second, we contribute to the small but growing literature on intergenerational health transmission, and the role of intergenerational persistence in shaping the SES-health gradient. In particular, research has documented a positive causal effect of parental income on child health (Case, Lubotsky, and Paxson 2002), but a relatively low intergenerational health association across a variety of adult health metrics (Halliday 2024), including lifespan (e.g. Black et al. 2024). Evidence from pre-demographic-transition Asia is sparse. Maystadt and Migali (2021) study intergenerational health transmission in China over a later period (the long 19th century), finding coefficients around 0.25 to 0.5.[4] However, they do not document status gradients, intergenerational status mobility, nor the relationship between health and status persistence, as we do. Our estimate of intergenerational health transmission is lower, and remarkably similar to that found for the late 19th and early 20th century US (Black et al. 2024). Our estimates confirm findings that mothers are especially salient for intergenerational health transmission and that lifespan

[3] See also Telford (1992) for results on the status gradient in mortality from a related sample.

[4] These unusually high estimates may be due to the sample, which is related to the Eight Bannerman, which were an elite military force of the Qing. Lavely (1998) describes the setting as a tightly controlled military colony where one studies mobility under a special set of circumstances.

information from both parents as opposed to only father is important for obtaining unbiased intergenerational mobility estimates.

Most distinctively, we contribute to understanding the origins of the health-status gradient by probing the relationship with 'wellbeing' persistence: intergenerational mobility of status as well as lifespan, and their relationships to each generation's Preston Curve. We are not aware of any previous studies analyzing intergenerational status and health persistence over multiple generations and their relationship to the health-status gradient for a given generation. Our empirical results highlight the salience of intergenerational health transmission for the gradient, even though status is more persistent across generations. Overall, our results reinforce the finding that a strong determinant of the gradient is poor health limiting earning potential (e.g., health shocks reinforcing cycles of illness-induced poverty). Especially in the historical era when few health technologies existed to protect even rich families from ubiquitous morbidity and premature mortality from infectious disease and injury, the 'survival dividend' from intergenerational persistence of good health strongly shaped the fortunes of Chinese across the status distribution.

We also contribute to a small but growing literature using genealogical data to study long-run questions of economic status or health. An early contribution is Fogel (1986), who compiled a sizable sample of genealogical records to better understand the evolution of mortality in North America. Recent work on Malthusian population responses in China using genealogical data includes Hu (2023), and data on lineages from this part of China have been used to address questions of the fertility quantity-quality trade-off, the evolution of social mobility as the return to human capital changes, as well as the long-run impact of big shocks on the power of elite (Shiue 2017, 2024, and Keller and Shiue 2024, respectively). Surveys of Chinese genealogies include Liu (1978) and Shiue (2016).

The results presented here are based on a small number of genealogies from a particular county in Qing China. Because broad administrative data for comparing our sample with population statistics is not yet available, it is impossible to know to what extent our results generalize to other parts of China. At the same time, census-like data that guarantees representativeness typically does not exist until the late 20th century; before, even studies with large samples are subject to unknown selection. For example, recent work on lifespan mobility with millions of observations for 1880-1920 US includes no more than 1/3 of the target population, and 99% of the sample is white (Black et al. 2024).[5] Learning from historical mobility patterns and health gradients requires understanding how data availability might affect the results. In the present sample, the individuals are on average wealthy compared to the typical Chinese during this period, in part because mortality of poor people is underreported (Telford 1992). Section II as well as the appendix examines this and other questions of sample composition. Relative to most studies for similar times,

[5] Reweighting the sample to obtain census weights will only work if one has good reasons to believe that the 1% remaining non-whites in the sample are representative of all non-whites in the population. See also Bailey, Cole, Henderson, and Massey (2020) on sample composition in ex-post linked US census data.

here commoners are in the majority in the sample, a key advantage for studying broad patterns of mobility, as opposed to mobility of only the elites.

The paper is organized as follows. The next section describes the data and setting. Then we present our empirical findings, starting with evidence on the health-status gradient, and then focusing on intergenerational transmission of health and of status, and their relationship to the evolution of the health-status gradient. The final sections discuss our results and briefly conclude.

II. Historical Setting and Data

II.1 Historical Setting

This study covers the relatively early part of the Qing Dynasty (1644-1911). Total population grew from about 200 million in the mid-17th century to 400 million by 1900 (Cao 2000). The effective power of government was organized around an autocratic central authority and lower administrative regions. The state taxed lightly in international comparisons with other states. Moreover, the scope of markets for allocating scarce resources was more limited by technology (e.g., transport technology, financial instruments) than by government regulation, in part because effective enforcement would have required more state resources than were in fact allocated.

In this structure of governance by bureaucratic-scholar officials, education was a cornerstone of entry into higher income classes and a key aspect of the social contract between the state and local management. Consistent with low central taxation, the capacity of the state was limited. The provision of public goods was instead delegated to local governments and the leaders of local clans.

Thus, local elites exercised paternal authority over their extended families, while the participation of local elites in the political decision-making body of the state was legitimized through the state-run civil service examination system. Throughout the Ming-Qing period, participation in the tournament-style civil service exams was the most direct path towards status and income. Passing the exam and obtaining a government office held substantial rewards. Social status derived from passing the civil service examinations was the “ultimate source of power” (Ho 1962). These returns were so attractive that non-scholar officials who obtained wealth also invested in their sons’ education.

On the one hand, local leadership by elites who passed the civil service exam strengthened and ensured the legitimacy of the central state. On the other hand, local elites who supported the state through participating in the civil service earned high financial rewards. There thus arose, in effect, a partnership between the central state and local elites, to mutual advantage.

II.2 Data

II.2.A Chinese Genealogies as a Source of Information

In this paper, we use one of the main sources of information produced by clans, the genealogy. Genealogies allow for a more expanded definition of status compared to many other sources of SES proxies, because they include not only status information mirroring the ranking of exam degrees and official titles used by the state, but in addition, genealogies record non-official types of wealth that clearly mattered.

Chinese genealogies are documents that record a family's history, providing information in the form of an annotated family tree where male individuals are listed in terms of an intergenerationally-linked structure along with the names and vital statistics of their spouses and children. Genealogies originated from beliefs surrounding ancestral ceremonies and Confucian teachings dating from at least the Tang Dynasty (618-907 A.D.). By the Ming and Qing, the practice had spread widely across the empire. As such, the Chinese genealogy is a well-known type of a historical document that falls under the category of household documents (Faure and He 2020). Like other household documents—which also include contracts, prayers for religious rituals, religious manuals, account books, tax payment records, almanacs, and textbooks—genealogies preserved information that was important to families within the social order of late imperial China. Wang (2008) catalogs approximately 50,000 publicly available genealogies, the bulk of which were produced in the Ming and Qing Dynasties.

Because the documents are self-reported, issues of representativeness, selection bias, and accuracy need to be considered. Not every family had a genealogy; some historical genealogies have survived while others have not. Also, Chinese genealogies are organized patrilineally—the family tree follows the marriages of the males of the family—implying less information on daughters than for sons. Compilation also required that at least some family members were literate, in which case the set of families for which genealogies exist might be better off than the set of families for which they do not exist (wealth bias). The principle of Chinese genealogical compilation was such that all males in the patrilineal line were included for reasons of ritual.

Clan membership was self-professed, but it was important because it defined and clarified allegiances and responsibilities during times when conflict might create uncertainty regarding obligations. Family laws were not immutable, and they could adapt to new situations, but the recording of details about members in genealogical charts—biographies, obituaries, official degrees, appointments to office—reflected the lineage as a force for communal organization. Unlike official household registers that linked population counts to tax levies (and which resulted in a built-in disincentive for accurate reporting), members of clans were more invested in the accuracy in their genealogical documents.

Other economic functions of Chinese genealogies are related to the way China was governed during the sample period. Whether implicitly or explicitly, the state delegated

much of the administration of local public goods to local clans. Tasks such as the management and upkeep of irrigation systems, poor relief, and grain storage for famines would oftentimes be passed down from the district magistrate to local clans. The amount that clans could contribute depended in turn on the resources at their disposal and the number of members in each clan. For example, lineage school fees were assessed on lineage members, with fees rising in an individual's income, providing reliable information on income differences between clan members (Chang 1955).

In summary, there is a considerable amount of information in family records that were privately maintained. Genealogies were common not only because of their centrality for ancestral rituals, but also because of their role documenting a written record that describes private forms of economic and political organization.

II.2.B Sample Data: Coverage and Representativeness

The data are from genealogies of clans that resided in Tongcheng County of Anhui Province. In the late 18th century, Tongcheng county had about 1.3 million inhabitants (Beattie 1979). The county was relatively developed and densely settled, one of the many thriving regions located near the lower Yangzi River Basin. Tongcheng was about 150 miles away from Nanjing, the early Ming Dynasty capital; Beijing, the later Ming and Qing capital, was about 650 miles away. The region was mainly a rice-producing area where the wealthiest families were typically landowners. Over the Ming and Qing Dynasties, the region gained some fame for having produced a number of the highest officials of the empire.

Genealogies exist for numerous clans in Tongcheng, with one account relying on more than 60 genealogies (Beattie 1979), and another study utilizing 39 genealogies (Telford 1992). The sample we use is created from data collection initiated by Ted Telford in the 1980s. Because employing all known genealogies from Tongcheng would lead to the oversampling of elites, the present sample is based on a targeted approach, where our goal is to obtain a broadly representative sample in the raw data. Given the focus on lifespan and status, an important part of the approach is to target the fact that most of the population in the genealogy had no social status ('commoner') while top-status holders may have been around 2%.

To achieve that, we select the genealogies of seven clans, the Chen, Ma, Wang, Ye, Yin, Zhao, and Zhou. For the entire period covered in their genealogies, this amounts to a total of about 40,000 men, women, and children. Taken together, these genealogies yield a sample that is broadly representative for China as a whole. For an analysis of various aspects of selection and bias, see Shiue (2025).

Of the more than 11,000 couples recorded in the data, the present sample consists primarily of the husbands born between 1675 and 1775. The start of the sample in 1675 is around 30 years after the fall of the Ming dynasty (1368 to 1644), so that the turbulence of

the dynastic transition from Ming to Qing have receded.[6] In addition, principles of data recording became more standardized during the Qing. We end the main sample with the 1775 birth cohort because the final year of compilation for some of the genealogies was in the 1860s; thus, using later birth cohorts would introduce censoring due to incompletely recorded lifespans.

These husbands are intergenerationally linked to their fathers and mothers for an analysis of intergenerational lifespan, while information on the status of the husband (as a son) together with the status of his father yields information on intergenerational status mobility. We also use information on the husband's wives to compute maternal mortality, as well as information on their male children to examine child mortality.

The main sample is restricted to observations for which vital statistics—birth year, birth month, year of death, month of death—for all individuals are recorded in the genealogies. In general, Chinese family genealogies vary in their completeness, and often some vital statistics are missing. While research at times estimates missing vitals using demographic tools such as period-relevant life tables, this paper relies exclusively on non-estimated vital statistics. In the appendix, we discuss how this affects empirical findings.

II.2.C Measures of Status, Health, and Summary Statistics

The genealogies record for each adult male his highest lifetime achievements in biographical sections that resemble obituaries. Our genealogies use more than 20 standardized descriptors of these achievements from which the social status of that man's household is derived. As is typical of data for the sample period, individual-level information is unavailable.

Participation in China's civil service examination, and based on that attaining official position, was key for a man's social status during the sample period (Ho 1962). One sign of the importance of the civil service exam for status is that high achievements in this dimension are virtually always recorded (Ho 1962). Other factors affecting social status in Qing China include occupation, education, making local investments, as well as other indicators of wealth and prestige. Table A1 presents the twenty-three social status levels employed in our data.

Since the number of men in some of these status levels is small, we aggregate the status levels up to four for robustness. These four status levels are described in Table S.

[6] See Shiue and Keller (2024) for more on consequences of the fall of the Ming dynasty.

Table S: Sample Composition in Terms of Four Social Status Levels

(1)	(2)	(3)	(4)
Status	Description	Share in All 3-Generation Sample (N)	Share in 1675-1775 sample with complete vital stats (N)
1	No Status	71.0% (6,318)	62.6% (1,062)
2	Village head, multiple consecutive marriages, relative of minor official	18.0% (1,600)	17.8% (302)
3	Wealthy farmer, landowner, or merchant; scholar; official student	5.1% (450)	8.5% (145)
4	Graduates of the civil service examination at local, provincial, and national level, high government officials	5.9% (522)	11.0% (187)
Total		100% (n = 8,890)	100% (n = 1,696)

Notes: Authors' calculation. Social status coding informed by Chang (1955, 1962), Ho (1962) and Telford (1986, 1992).

The four status levels distinguished in this paper are roughly no status (1), moderate status based on some honors and distinctions (2), high status based on wealth, education, or preparation for the civil service exam (3), and top status based on passing the civil service exam and the resulting attainment of high government positions. If a genealogical entry of a man does not indicate distinctions of any kind, he is coded as No Status. We also refer to these men as commoners. For more details on these status levels, see Table A1.

Column (3) of the table indicates in the last line that the seven clan genealogies allow compiling a dataset of n = 8,890 three-generation observations of father, husband, and male child. In 71.0% of these cases, the husband has no status. The fact that the majority of individuals in this data are commoners is a key advantage. Information about Europe for this period, in particular, comes mostly from convenience samples composed almost exclusively of elites such as barons and other noblemen (Black et al. 2024).

Instead of the status categories 1 to 4, the empirical analysis employs a man's status in terms of his percentile rank. This has the advantage of accounting for the size of each status class, and has also been shown to lead to more robust regression results (Dahl and DeLeire 2008, Chetty et al. 2014). The percentile rank of each of the commoners is equal to 0.355, which is equal to the midpoint between 0 and 0.71. The percentile ranks of status levels 2, 3, and 4 are similarly the midpoints of their brackets in the status distribution

(corresponding to 0.801, 0.916, and 0.971, respectively). For simplicity, we refer to these four status classes as those at the "35th", the "80th", the "90th", and the "97th" percentile.

Our analysis focuses on a subset of observations from the seven genealogies, namely the husband birth cohorts 1675 to 1775 that have complete vital statistics (on fathers, mothers, husbands, wives, and male children). Column (4) of Table S shows the consequences of this sample restriction. First, the sample size is smaller. One reason for this is that focusing on the century from 1675 to 1775 eliminates not only many earlier generations—the seven genealogies record up to 22 generations—but also many observations for the later period.[7] Despite the decrease in size, however, the sample still allows quantitative analysis as we will see below.

We also see that the share of commoners in the 1675-1775 sample, at 62.6%, is lower than before; and correspondingly, the share of high and top status men is higher. This is because vital statistics for commoners are less consistently recorded than vitals for higher-status men. We discuss implications of this in the appendix.

Next, we turn to summary statistics of key variables for the sample given in column (4) of Table S.

[7] The latest year of death recorded for a husband is 1925.

Table 1. Summary Statistics

	N	Mean	Std. Dev.	Min	Max
Panel A. Intergenerational Parent-Son (Husband) Sample					
Husband Lifespan	1,677	56.34	15.22	16	90
Father Lifespan	1,697	61.49	14.11	22	91
Mother Lifespan	1,698	63.22	15.88	18	95
Husband Status	1,697	0.550	0.257	0.355	0.971
Father Status	1,698	0.564	0.269	0.329	0.957
Husband Birth Year	1,698	1736.0	26.35	1675	1775
Father Birth Year	1,698	1704.7	27.21	1629	1759
Mother Birth Year	1,698	1707.4	26.63	1630	1759
Panel B. Child Mortality					
Death of a male child before age 10	2,192	0.0055	0.074		
Panel C. Maternal Mortality					
Mother dies within 1 year of giving birth to a son	4,322	0.012	0.107		

Notes: Panel A. for all husband birth cohorts 1675 to 1775 for which mother, father, and husband reach at least 15 years of age. Panel B. based on all male children born to husbands in birth cohorts 1675 to 1750. Panel C. based on all female partnerships of husband birth cohorts 1675 to 1775 for which mother, father, and husband reach at least 15 years of age.

Summary statistics for the main empirical analysis are given in Panel A. We see that the average male lifespan (i.e., for the husbands in the sample) was 56 years. This is conditional on reaching age 15 as well as conditional on being able to marry (it is estimated that between 10% and 20% of men remained unmarried due to lack of resources at the

prevailing sex ratio). Average lifespan of fathers and mothers is 61 and 63, respectively, and lifespan in the sample reaches into the 90s.

Average husband and father percentile status is by design close to 50 percent. Four different status levels are distinguished, with Table 1 giving the minimum and maximum (0.355 and 0.971 for husbands, respectively). The fact that average status of fathers somewhat exceeds that of sons is because it became harder for men to succeed in the civil service examination, given a constant number of official positions and rising population.

We also record information on husband, father, and mother birth year. Husband birth year varies from 1675 to 1775, defining the sample. The earliest birth years of father and mother are 1629 and 1630, respectively, which is still in the Ming dynasty.

Panels B. and C. give summary statistics on child and maternal mortality. The number of observations is different because the unit of observation is no longer a particular intergenerational father-husband pair but rather a birth.

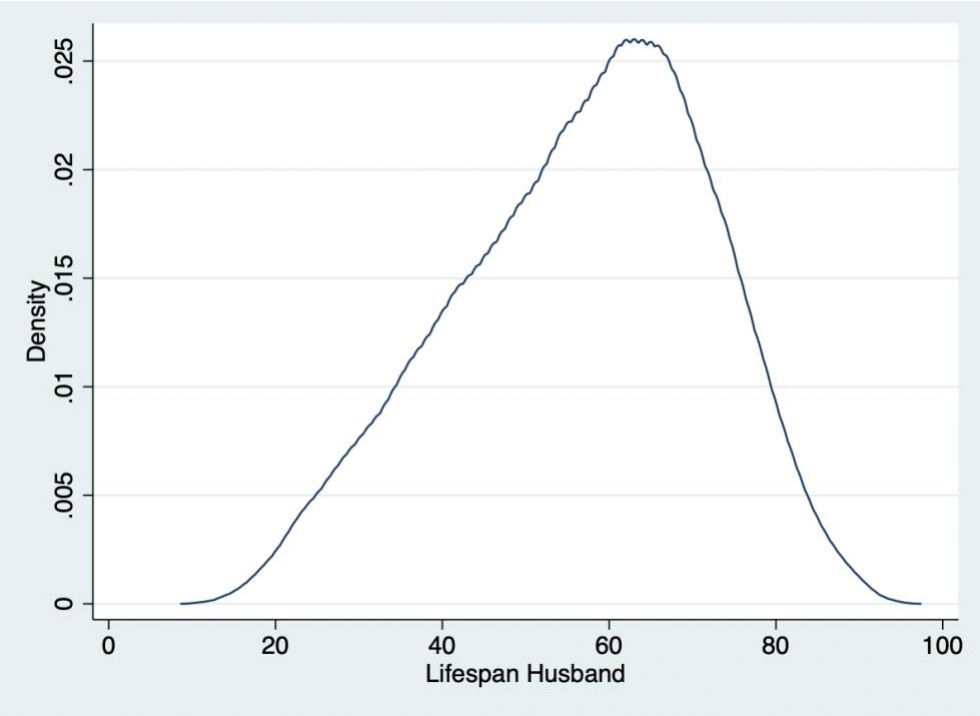
We now turn to the empirical results.

III. The Preston Curve in Tongcheng in the Early Qing Dynasty

III.1 The Preston Curve for Men

Figure 1 shows lifespan, defined as year of death minus year of birth, for all sample men that can be intergenerationally linked to the lifespan of their fathers and mothers for birth cohorts 1675-1775 that reached 15 or more years. We see that the average male adult lifespan is about 56.5 years, with wide tails, consistent with variation in health outcomes.

Figure 1: Completed Lifespan of Married Men



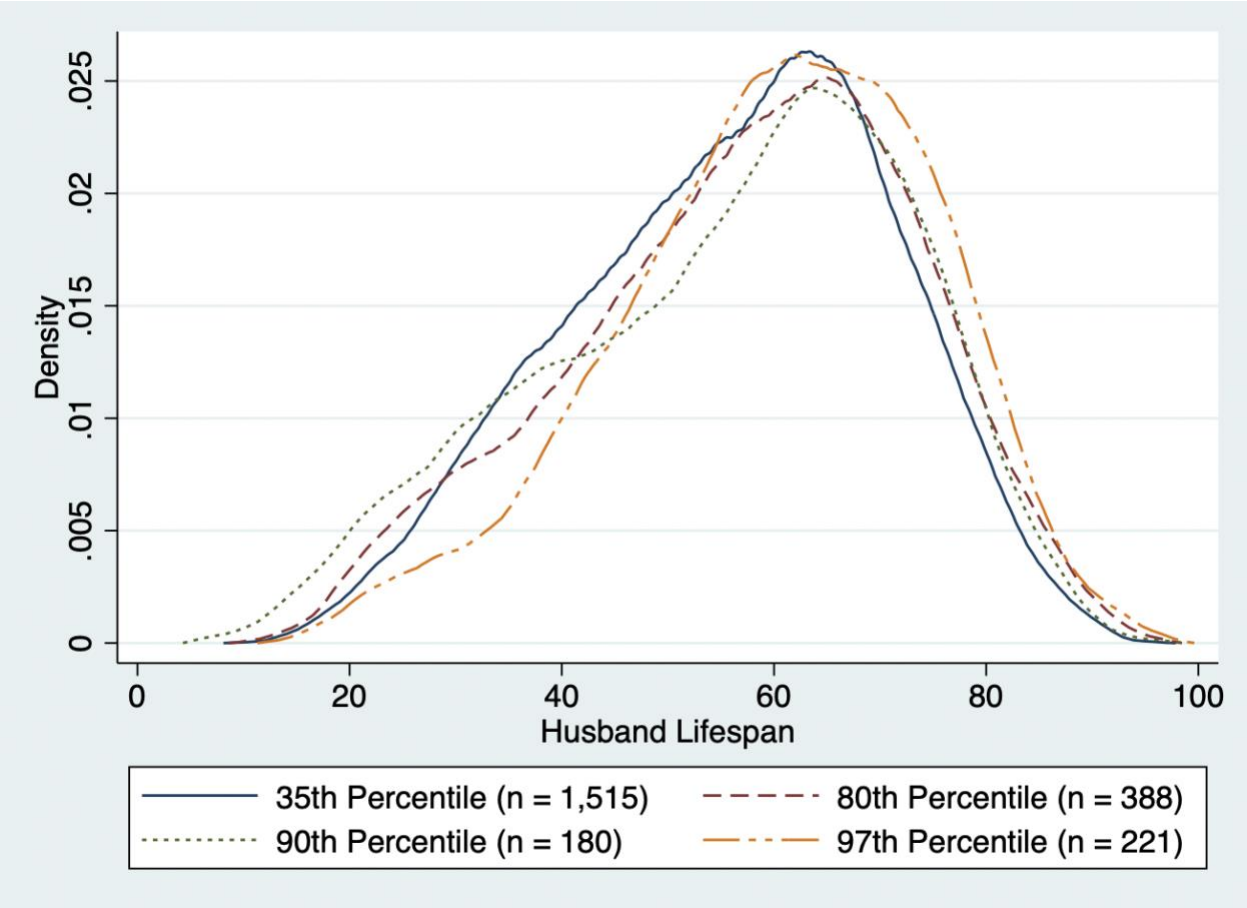
Notes: All husbands in birth cohorts 1675 to 1775 with 15 or more years of lifespan.

Employing birth cohorts 1675 to 1775 has the advantage that results are robust to the relatively smaller number of men of these families during the Yuan and Ming dynasties, and findings are also not affected by lifespans that extend beyond the latest genealogy update (for some genealogies the late 1860s) and were most impacted by the violence of the Taiping rebellion and dynastic decline.[8]

The analysis in this paper distinguishes four status levels (see the earlier section II). They are referred to in Figure 2 by the approximate percentile of the midpoint of their class. Thus, the 35th Percentile are men without signs of elevated status (commoners, or lowest class). Table S shows that these men account for 71% of all men (0.71/2 = 0.35). Similarly, the 97th Percentile men are approximately the top 6 percent of the status distribution; these are graduates of the national civil service examination, ranging from 94th to 100th of the status distribution. Figure 2 shows evidence for the Preston curve in Tongcheng for 1675-1775 birth cohorts in that the distribution is shifted to the right the higher is a man's status. At the mean, the difference in lifespan for the poorest and the richest men is 4.5 years.

[8] Demographers studying China have long known the problems with vital statistics in the 19th century; for example, Lee, Wang and Campbell (1994) focus on Qing nobility 1700 to 1840 precisely because "the quality of the data deteriorates significantly after 1840 (a reflection of dynastic decline)" (p.396).

Figure 2: Lifespan Densities by Status

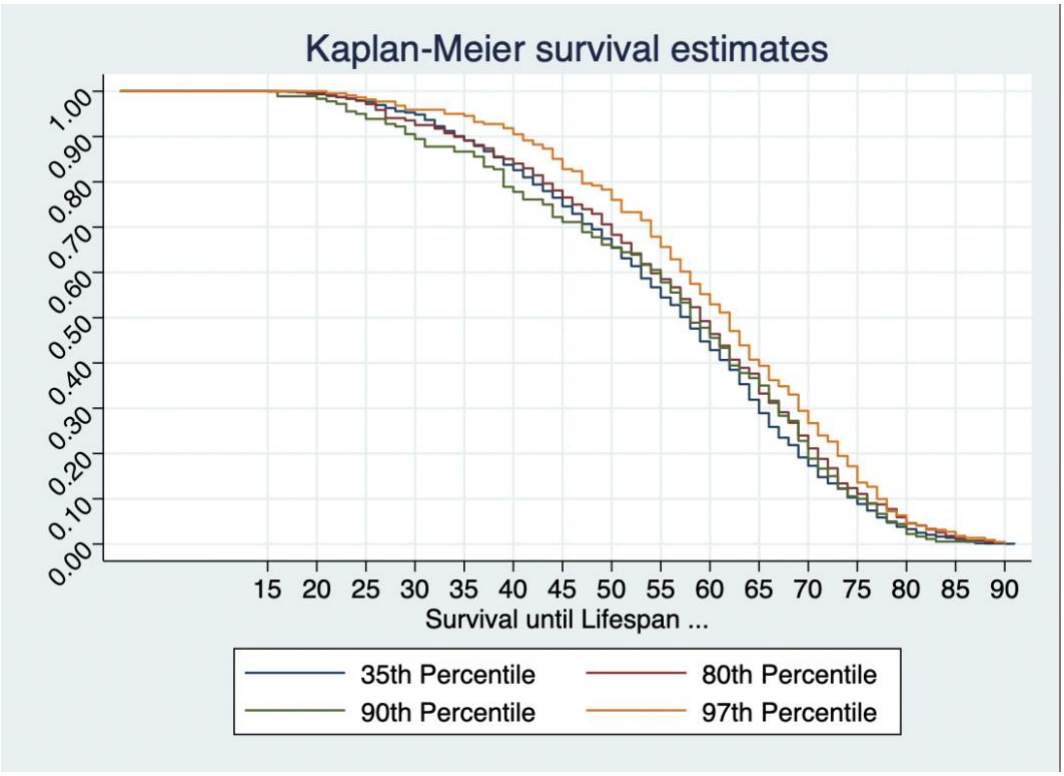


Notes: All husbands birth cohorts 1675-1775 with at least 15 years of lifespan.

At the same time, there is wide variation in lifespan for a given status, in line with results for more recent times and richer economies (e.g., Black et al. 2024).

Another well-known way of analyzing mortality differences are Kaplan-Meier survival estimates. Figure 3 shows these in the present case.

Figure 3: Kaplan-Meier Survival Estimates

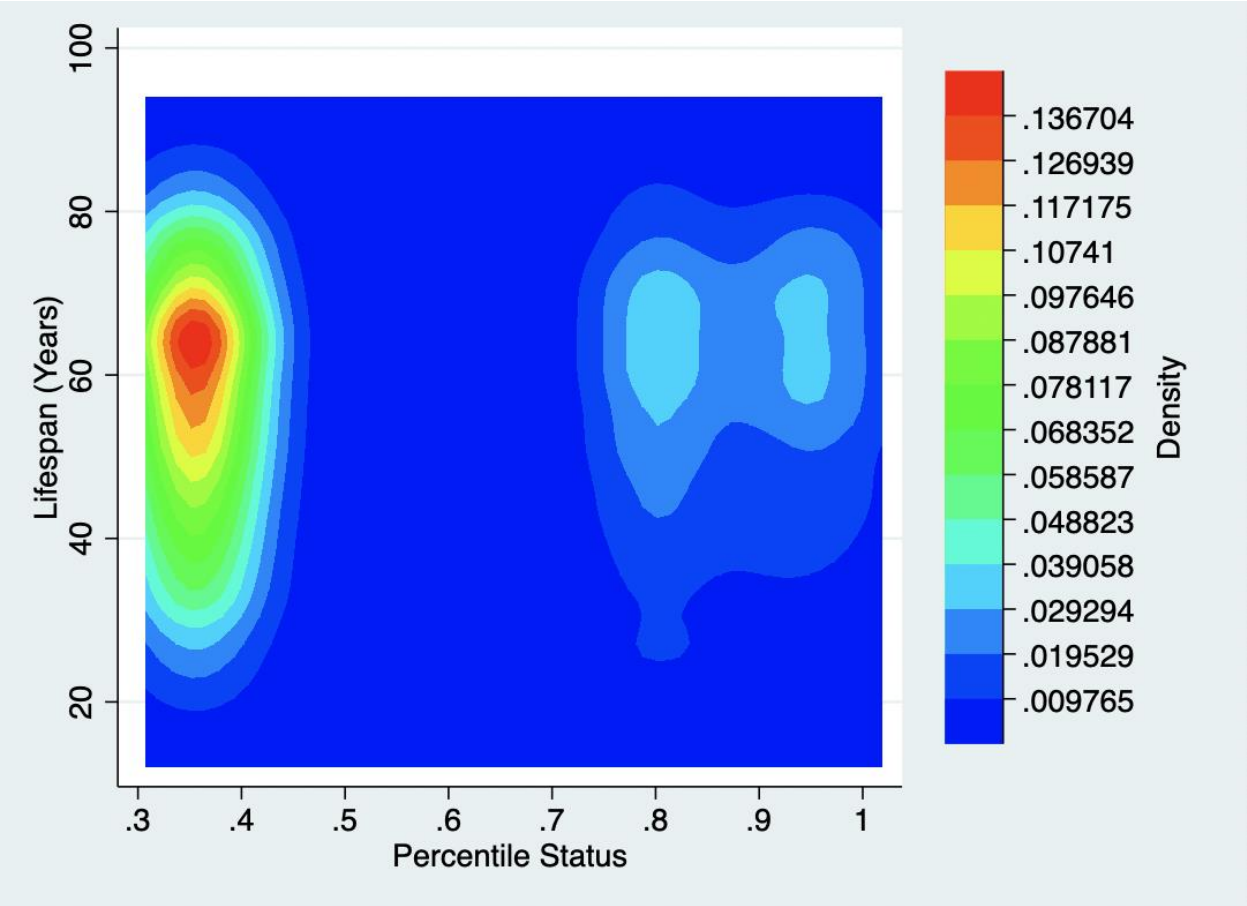


Notes: All husbands in birth cohorts 1675 to 1775 with at least 15 years of lifespan.

These survival estimates confirm that lifespans differed nontrivially by status. For example, the probability that a top status man (97th percentile) would reach 55 years or more was 67%, compared to the probability of 55% for a commoner (lowest status level). While this 12-percentage-point difference is non-neglible, it is clearly not the case that status determines a man's health destiny in this society. To provide some historical context, it is approximately the difference between men in Sweden in the 1750s and men in England almost a century later, in the 1840s (the earliest available robust data for both, based on the Human Mortality Database, HMD). That is, survival to 65 conditional on surviving to 15 (S(65|15)) was 35% for Swedish men born around 1751 and 43% for men born around 1841 in England and Wales, compared to our estimate of about 30% for commoners and about 40% for men achieving top status.

Figure 4 shows the densities of status and lifespan. The mass of sample are commoners (35th percentile). They have a typical lifespan just under 56 years, and experience considerably larger lifespan variability than those attaining higher status. At the same time, because we observe only four distinct status levels, the analysis depicted in Figure 4 is necessarily coarse.

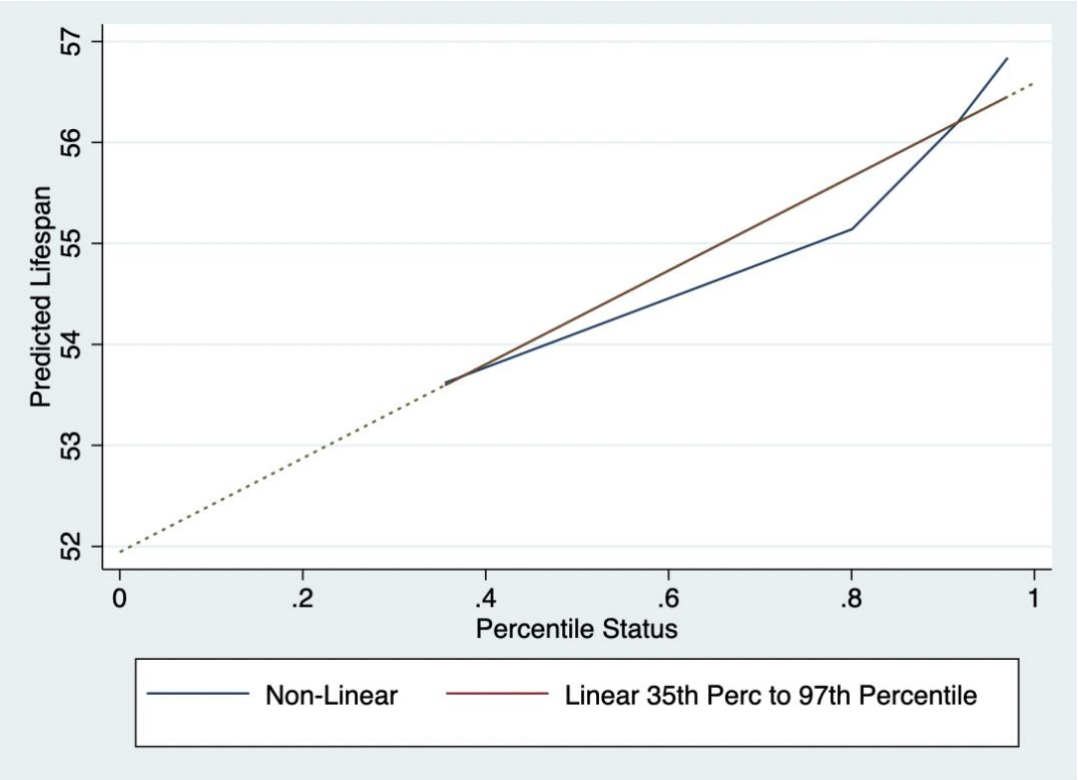
Figure 4: Densities of Status and Lifespan



Notes: All husbands in birth cohorts 1675-1775 with at least 15 years lifespan (n = 2,304).

For simplicity, the income-lifespan relationship is often described as linear. For large contemporaneous samples in the US for the central 90 percent, for example, it is approximately linear (Chetty et al. 2016). In the present case we cannot analyze potential non-linearities at the low end of the status distribution (because the lowest status level covers 70 percent and status cannot be further distinguished). However, as Figure 5 shows, at the top end of the status distribution, our results parallel findings of a steeper gradient in the US during the early 21st century (Chetty et al. 2016).

Figure 5: Status-Lifespan Gradients in Tongcheng, China, 1740 to 1760



Notes: All husbands in birth cohorts 1740-60 with at least 15 years lifespan (n = 558).

Figure 5 indicates that assuming a linear status-lifespan gradient when in fact it is non-linear underestimates the typical lifespan of a *jinshi* by about one year.

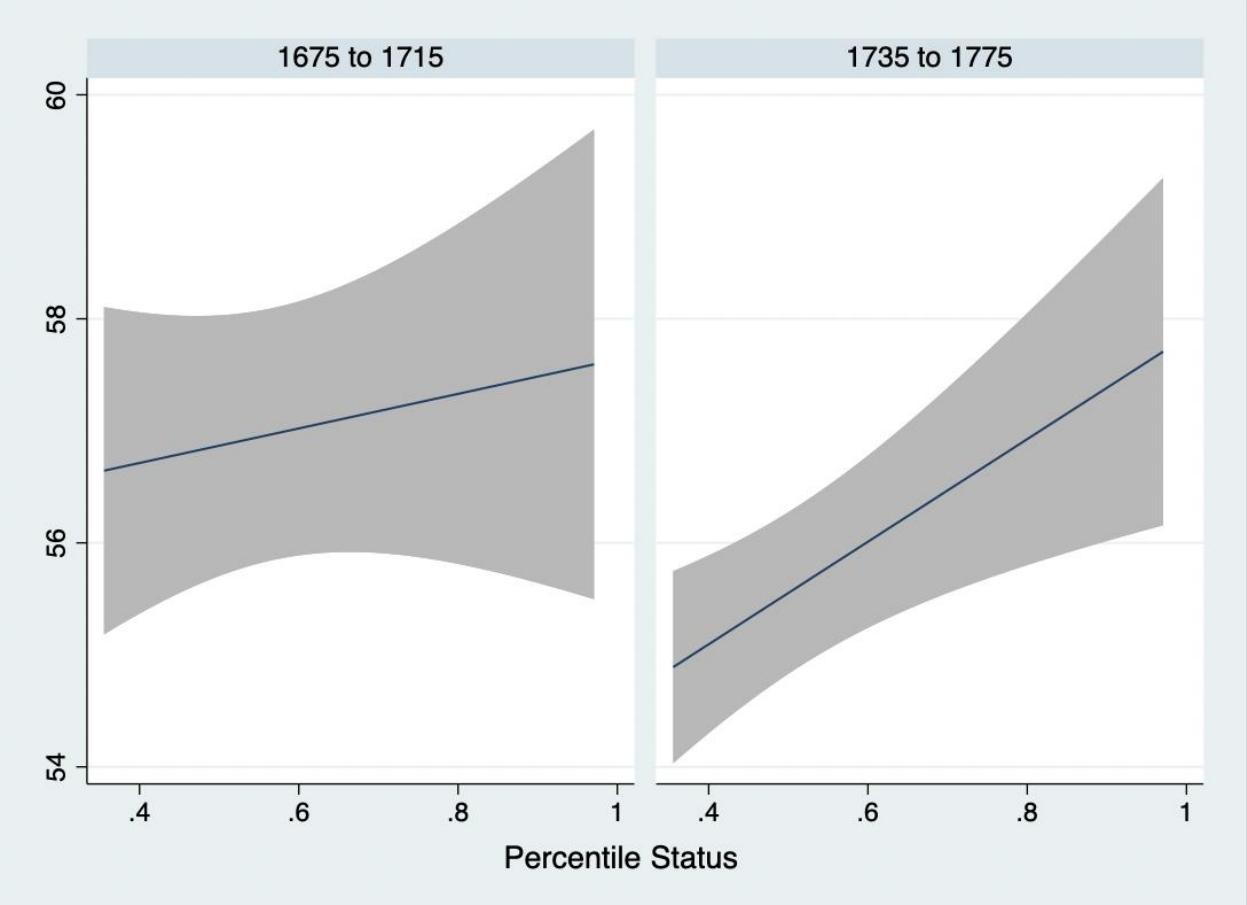
As noted earlier, lowest status is coded as the midpoint of the lowest status class which covers approximately the bottom 70 percent of the population. Our best estimate for the increase in lifespan associated with moving status from lowest (0 percent) to highest (100 percent) is roughly 4.5 years, as the green dashed line in Figure 5 indicates. The raw correlation of lifespan with status is 7 percent in this sample.

We have also explored the role of incomplete vital statistics in the genealogies for the evidence shown in Figure 5. Using a range of standard demographic tools such as period-specific life tables, this imputation of missing vital statistics strengthens the status-lifespan gradient shown in Figure 5. Specifically, while the difference in typical lifespan for men at the 100th versus the 1st percentile according to the linear specification in Figure 5 is 4.5 years, using the augmented dataset, the difference in lifespan predicted for a 100th versus 1st percentile status man is around 8 years (see Figure A1). This difference reflects the greater completeness of data for higher-status and longer-lived individuals.

It is also of interest to examine in our sample over more than a century any changes in the Preston curve over time. Figure 6 distinguishes two sets of cohorts, 1675 to 1715 and 1735 to 1775. For the earlier period the linear relationship between status and lifespan is positive; however, the curve is quite flat. In contrast, there is little doubt that by the middle

of the 18th century the typical lifespan of a *jinshi* is significantly higher than that of a commoner. While the larger sample size for the later cohorts favors finding a significant gradient, the slope coefficient is also higher in the later birth cohorts (slope coefficient of 4.6, versus 1.5 for the earlier cohorts).

Figure 6: Preston Curves by Birth Cohort



Notes: All husbands with lifespan 15 or higher; birth cohorts 1675 to 1715 on the left (n = 538), birth cohorts 1735 to 1775 on the right (n = 1,165).

From the evidence in Figure 6, the significant health-status gradient by the middle of the 18th century appears to be in part the result of higher young-adult mortality among those at the lower end of the status distribution.[9]

[9] This pattern might arise in part because of incomplete recording of commoners' premature mortality in earlier periods. Expanding the sample by estimating missing vital statistics (i.e., using other genealogical information, model life tables and other demographic methods) suggests that a significant Preston curve in Tongcheng might have been present already in the 1600s. Thus, Figure 6 is a conservative estimate of when a significant status-lifespan gradient emerged (or re-emerged) in Tongcheng.

III.2 Gradients beyond adult men: Child and maternal mortality by status

We have shown that a significant health-status gradient in Tongcheng emerged over the 18th century for married men. While the genealogical data for women and children is slightly more limited, this section documents mortality gradients for those population subgroups as well. This suggestive evidence supports our interpretation that the gradient among married men is a conservative estimate of the overall health-status gradient.

III.2A. Child Mortality

The gradient in adult lifespan may underestimate the overall gradient in health and life expectancy if there were significant differences in infant and child mortality by status. Moreover, the 'mortality tax' on parental investments in children may have constrained the quality-quantity trade-off that contributes to a health-status gradient. Therefore, in this section we examine data on child survival to document any gradient in the 'mortality tax' that might have played a role in the emergence of the health gradient and indicate an even larger gradient in life expectancy. However, death at an early age is often left unrecorded in genealogies, especially in earlier periods, among lower-status households, and for female infants; therefore we focus on male children with an understanding that the overall rates of child mortality are underestimates of absolute magnitudes and of the status gradient, so conclusions drawn from the following analysis have to remain tentative.

Table 2 provides information on the average age and mortality rates of boys by birth cohort of their fathers.

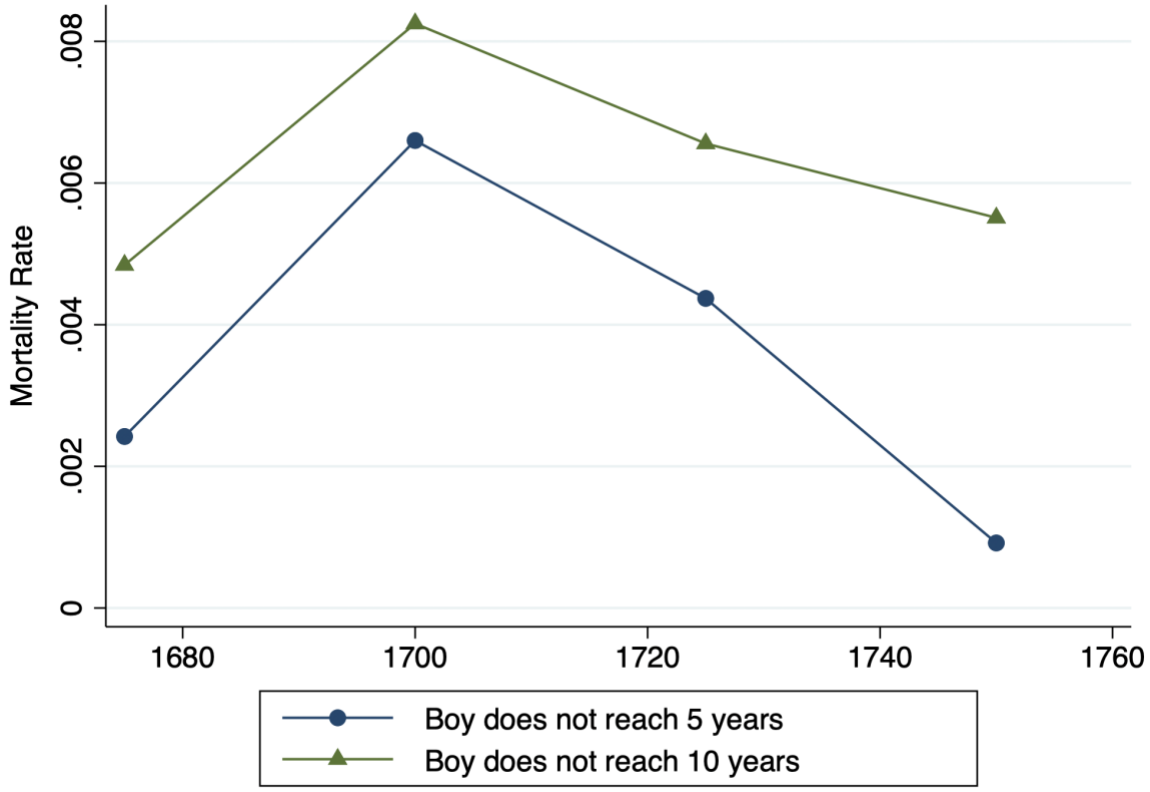
Table 2: Average Lifespan and Mortality Rates of Boys, by Father Birth Cohort

	N	Lifespan	Lifespan < 5 Years	Lifespan < 10 Years
1600-1650	352	59.88	0	0
1650-1675	397	55.97	0.0024	0.0048
1675-1700	576	55.94	0.0066	0.0083
1700-1725	844	53.02	0.0044	0.0066
1725-1750	978	52.34	0.0009	0.0055
1750-1775	1,191	49.58	0.0029	0.0044
1775-1800	940	43.83	0.0031	0.0138

Notes: All child vitals for specific father birth cohort.

Mortality rates of zero in the earliest birth cohort of Table 2 presumably indicate that at the time, early deaths of children were customarily not recorded yet. Moreover, from 1750 onwards, the data records a major reduction in lifespan (and increase in mortality). An important reason for that is likely censoring of the genealogy, i.e., the lack of recorded death dates for longer-lived individuals in the war-torn mid-1800s (Taiping rebellion and dynastic decline). The following analysis will therefore focus on the father birth cohorts from 1675 to 1750, who on average had children 20-30 years later.

Figure 7: Child Mortality Rates

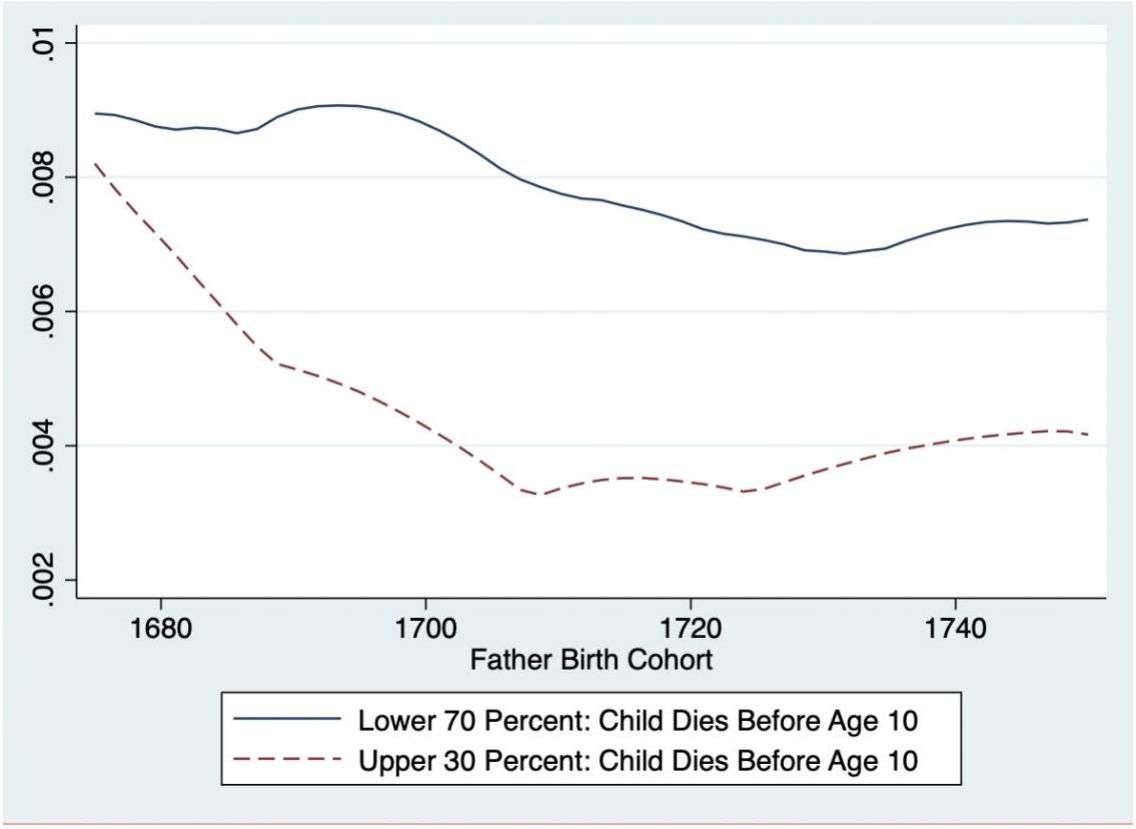


Notes: Child mortality rates, based on all boys born to fathers in birth cohorts 1675 to 1750.

Figure 7 shows how mortality rates below five and ten years of age evolved over this period. The decline in mortality rates shown in the figure for children born in the mid-to-late 1700s (to fathers of the 1700 to 1750 birth cohorts) is consistent with a lower 'mortality tax' on child investment at the time when the health-status gradient for their fathers strengthened (recall Figure 6).

The decline in child mortality for late 17th and early 18th century cohorts also exhibits a status gradient, as Figure 8 shows.

Figure 8: Child Mortality over Time by Status



Notes: Child mortality rates based on all boys born to fathers in birth cohorts 1675 to 1750.

Figure 8 indicates that the decline in mortality rates between 1670 and 1730 for higher status families was more substantial compared to that experienced by commoner families. This is consistent with the Preston curve diffusing through society from the top of the status distribution, see Figure 5.

Moreover, this decline in child mortality is strikingly consistent with (albeit slightly predating) the steep decline noted in the mid-1700s by previous studies (e.g. Lee et al. 1994). These patterns confirm that status conferred some protection from the heavy selective pressure of infectious disease, although the mortality rates in our data are clearly underestimated (presumably because, as noted, genealogies often did not record neonatal and infant mortality, the single largest cause of mortality in any population globally before the 20th century).

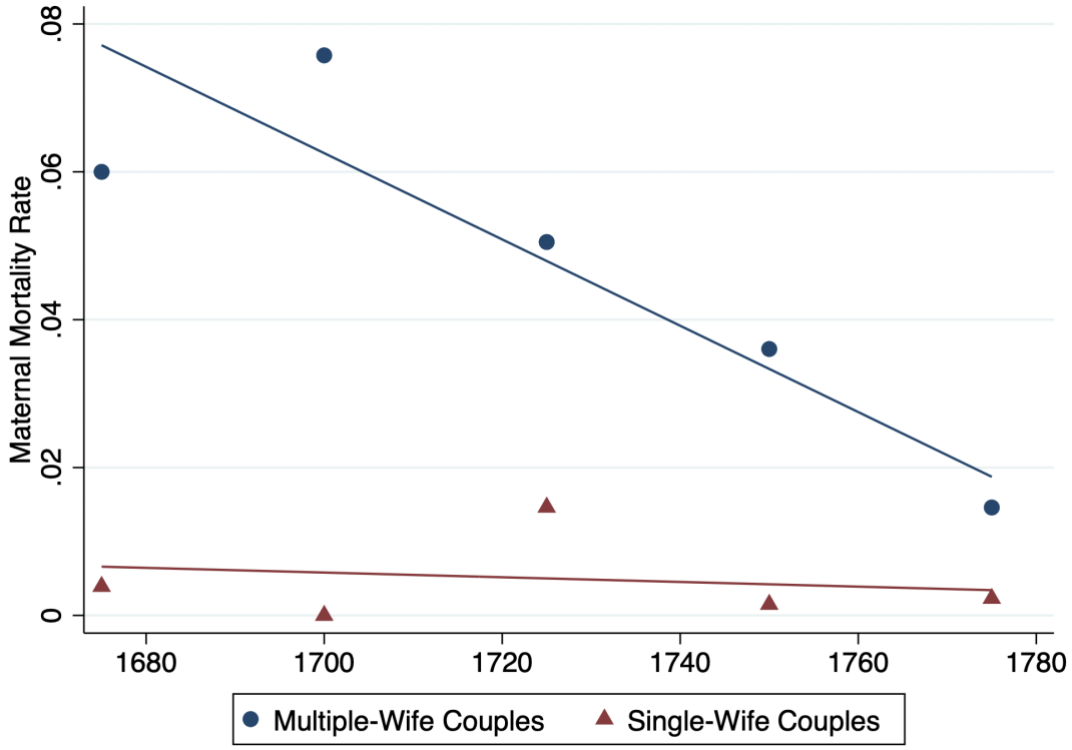
III.2B. Maternal Mortality

While the majority of our analyses focus on adult men, the patriarchal genealogical data can also provide insights into female health and survival. The mothers of surviving and married sons enjoyed some of the longest average lifespans in this period, varying between 62 and 66 years (Table 1). In this section, we also examine maternal mortality rates as an

important source of female premature mortality and a potential mechanism for the health-status gradient, since (as we will show) mother's health is highly predictive of son health. Maternal mortality is defined as the mother dying in the same or following year as the birth of one of her sons (there are too few cases of maternal mortality related to the birth of daughters in the 17th and 18th centuries in the data).

An important feature of the data is that it records possibly more than one wife for a given man, as a separate couple with their respective children. Typically, these are sequentially monogamous relationships that occur when a man's wife dies relatively early.[10] Given considerable wedding expenses, only men of higher status would have the financial means to marry multiple times. Because this complicates the relationship between status and maternal mortality, the analysis below will distinguish cases with a single lifetime wife from cases with multiple lifetime wives.

Figure 9: Maternal Mortality for Multiple vs Single Wife Relationships over Time



Notes: Maternal mortality rates for all marriages with husbands in birth cohorts 1675 to 1775.

Figure 9 shows that maternal mortality rates had a falling trend from the late 17th to the late 18th century in Tongcheng. As the figure indicates, maternal mortality in multiple-wife relationships was higher than in single-wife relationships. This is not surprising because the very reason that a man would marry another woman was typically that the first wife

[10] Polygamous relationships and non-married female partners (concubines) are recorded in the genealogies, but they are rare.

died early, and maternal mortality was a major cause of female premature mortality. Figure 9 indicates that the decline in maternal mortality among women in multiple-wife relationships was particularly large after the year 1700. Because this would be women married to relatively rich husbands, this is consistent with the Preston curve diffusing through society from the top during this period. At the same time, the number of multiple-wife cases is relatively small, and the following analysis of a possible status gradient will focus on single-wife cases.[11]

We ask whether maternal mortality is related to the status of the father (her husband) using simple OLS regressions. The sample is all births of boys to single-wife relationships during father birth years 1675 to 1775.[12]

Table 3: Maternal Mortality and Husband Status

	(1)	(2)	(3)	(4)	(5)
Husband Status	-0.006 (0.004)	-0.012** (0.006)	-0.011* (0.006)	-0.012* (0.006)	-0.015 (0.011)
Mother Birth Year FE		Y	Y	Y	Y
Father Birth Year FE		Y	Y	Y	Y
Son Birth Year FE			Y	Y	Y
Son Birth Order FE				Y	Y
Father Birth Order FE				Y	Y
Mother Birth Order FE					Y
N	2,829	2,826	2,821	2,820	1,193

Notes: Dep var is maternal mortality; sample is all single-wife births 1675 to 1775; robust s.e. in parentheses; */**/*** indicates sig at 10/5/1% level.

Table 3 indicates that the correlation of maternal mortality with husband status tends to be negative. Controlling for common shocks by including fixed effects for each father and mother birth year, the coefficient on husband status is -0.012 (column (2)). This estimate suggests that the predicted maternal mortality of a husband at the 85th percentile is 0.006 lower than a husband at the 35th percentile. With a mean mortality rate in the sample of 0.005, this is a large difference.

[11] Lee et al. (1994) also show the importance of examining monogamous versus polygamous nobility marriages for understanding fertility and infant mortality: "The coefficients for the period 1751-1820 imply that daughters born to monogamous members of the lower nobility were 2.36 times more likely to die during their first month of life than those born to polygamous parents in the higher nobility, and 2.85 times more if their own parents were polygamous." (p.406).

[12] We employ only cases in which neither mother nor child vital statistics are estimated.

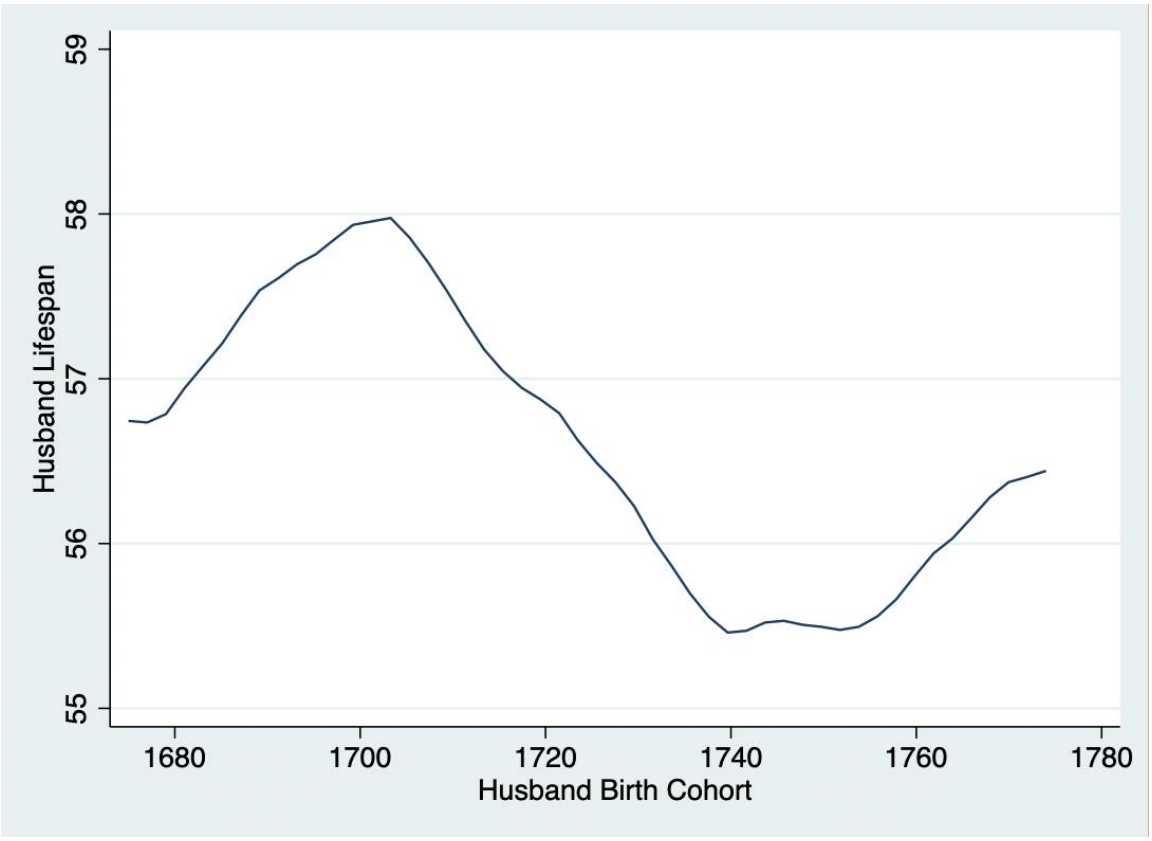
The remaining specifications in Table 3 show that the status-maternal mortality gradient is not strongly affected by controlling for birth order of son and father (column (4)), though precision goes down when mother birth order is accounted for due to missing data (column (5)).

Overall, the late 17th and earlier 18th century in Tongcheng was characterized by falling child and maternal mortality, which higher status families enjoyed to a somewhat greater extent than lower status families. These patterns suggest that status was somewhat protective of survival for children and mothers as well as for married men, and that the overall health-status gradient in life expectancy was larger than for adult lifespan.

IV. Intergenerational Transmission of Lifespan

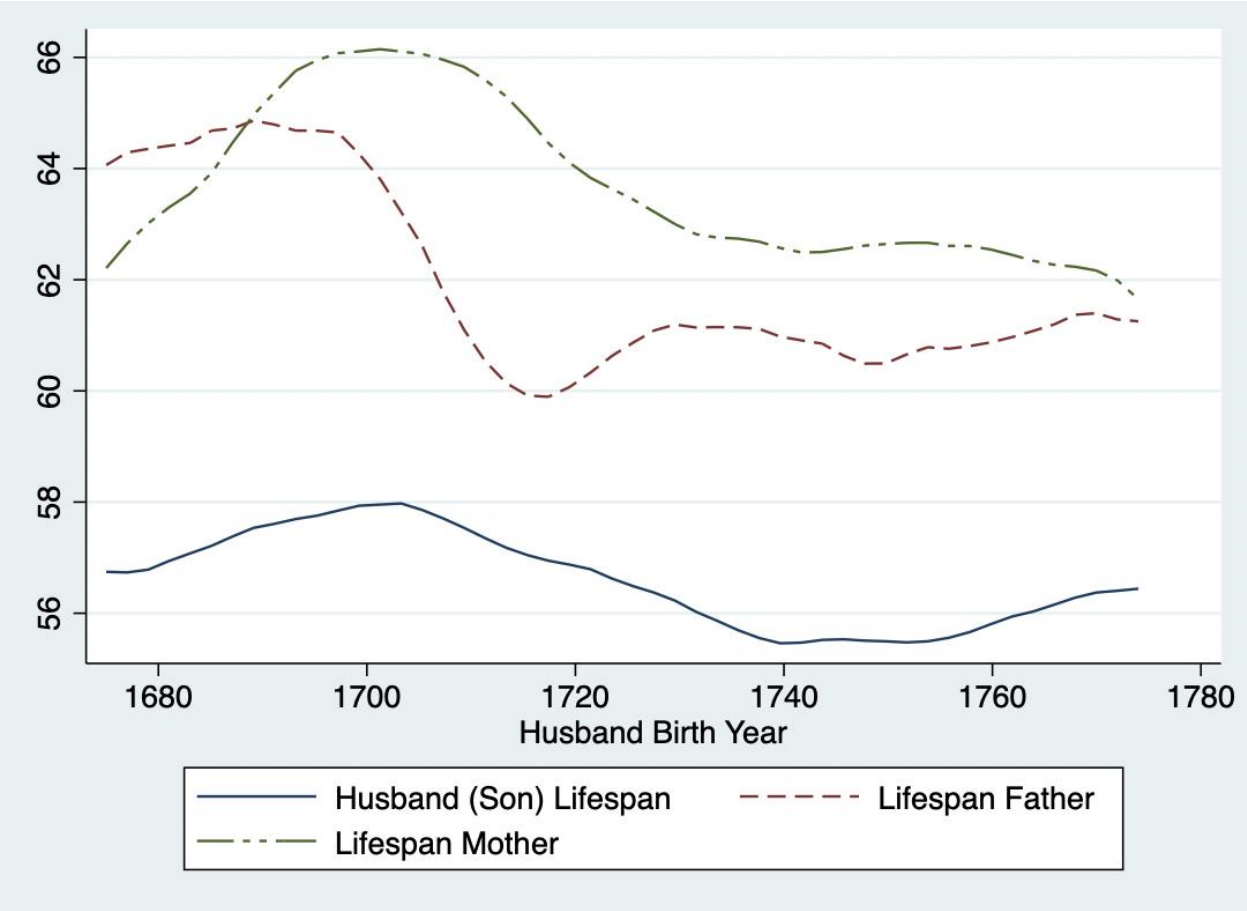
Figure 10 shows the evolution of male lifespan for birth cohorts between 1675 and 1775. There is a moderate downward trend over the century, with a mean around 56.5 years.

Figure 10: Male Lifespan (Husbands)



Notes: All husbands in birth cohorts 1675 to 1775 with at least 15 years of lifespan.

Figure 11: Lifespan of Father, Mother, and Son (Husband)



Notes: All husbands, fathers, and mothers with at least 15 years of lifespan for son birth cohorts 1675 to 1775.

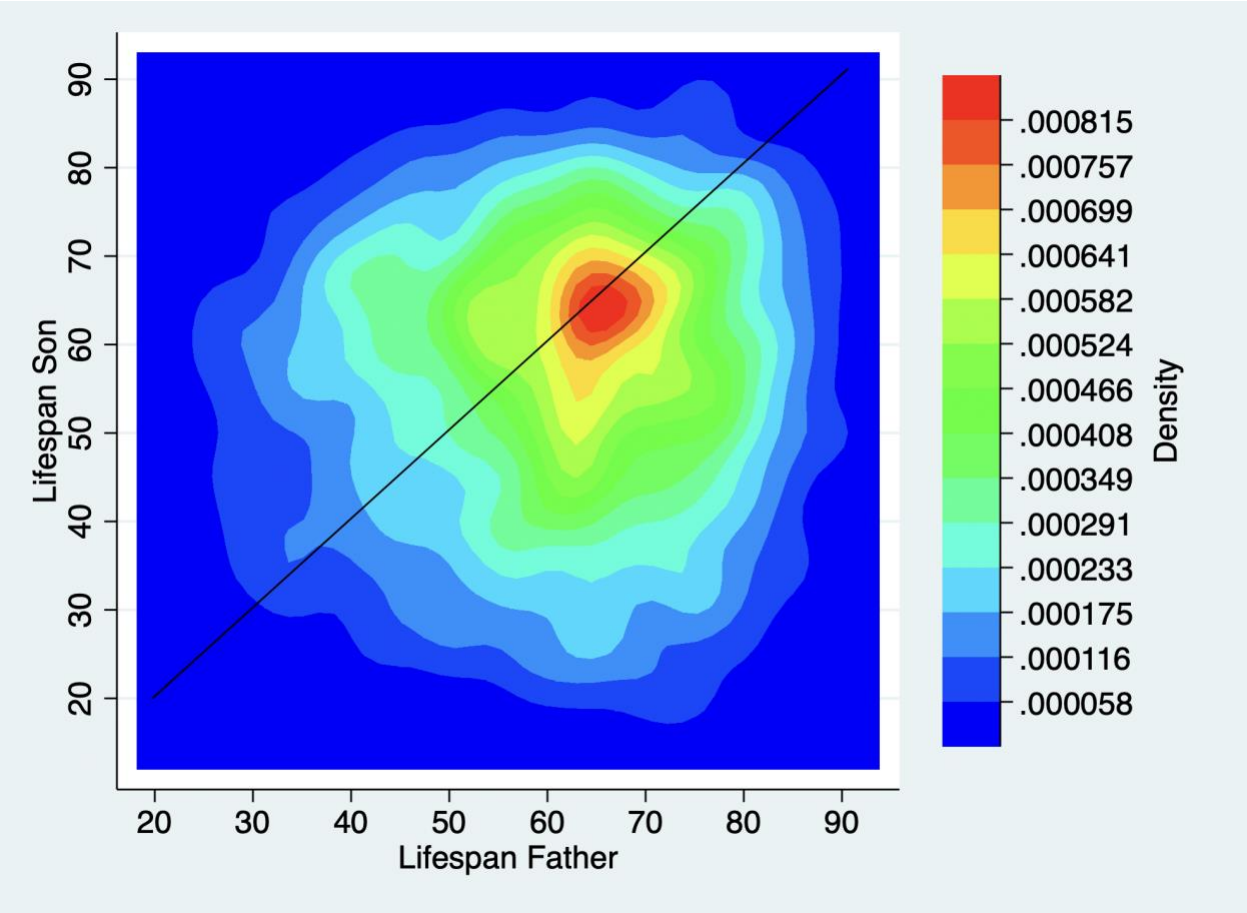
In Figure 11 we show the broad evolution of lifespans of fathers and mothers together with that of the sons, all as a function of the son's birth cohort. We see that sons have typically a shorter lifespan than their parents. For the earlier birth cohorts, lifespans across generations may have been affected by particular historical factors, such as the recovery from the Ming-Qing transition (the Qing dynasty began in 1644). The lifespan difference between generations is consistent with a downward trend in adult lifespan, conditional on (improved) survival to age 15. In other words, the parental generation would have been subject to stronger selective pressure in childhood, and this selective survival contributed to the generational gap in adult lifespan. The gap starts to stabilize with the 1740s birth cohorts, that is, after the decline in child mortality (see Figure 8).

We are now ready to examine various measures of intergenerational lifespan mobility. First, consider the lifespan densities of both father and son, shown in Figure 12.

The key finding is that there is tremendous variation in the way that health, proxied by completed lifespan, is transmitted intergenerationally. For example, for a father with a

relatively low lifespan of 40, the average son lifespan is 50, but the interquartile range is from a low of 38 years to a high of 67 years. Moreover, for fathers who survive to age 70, about 10% of their sons do not reach half that lifespan. The mode of the joint distribution is below the 45-degree line, which reflects the overall decline in lifespan over the sample period.

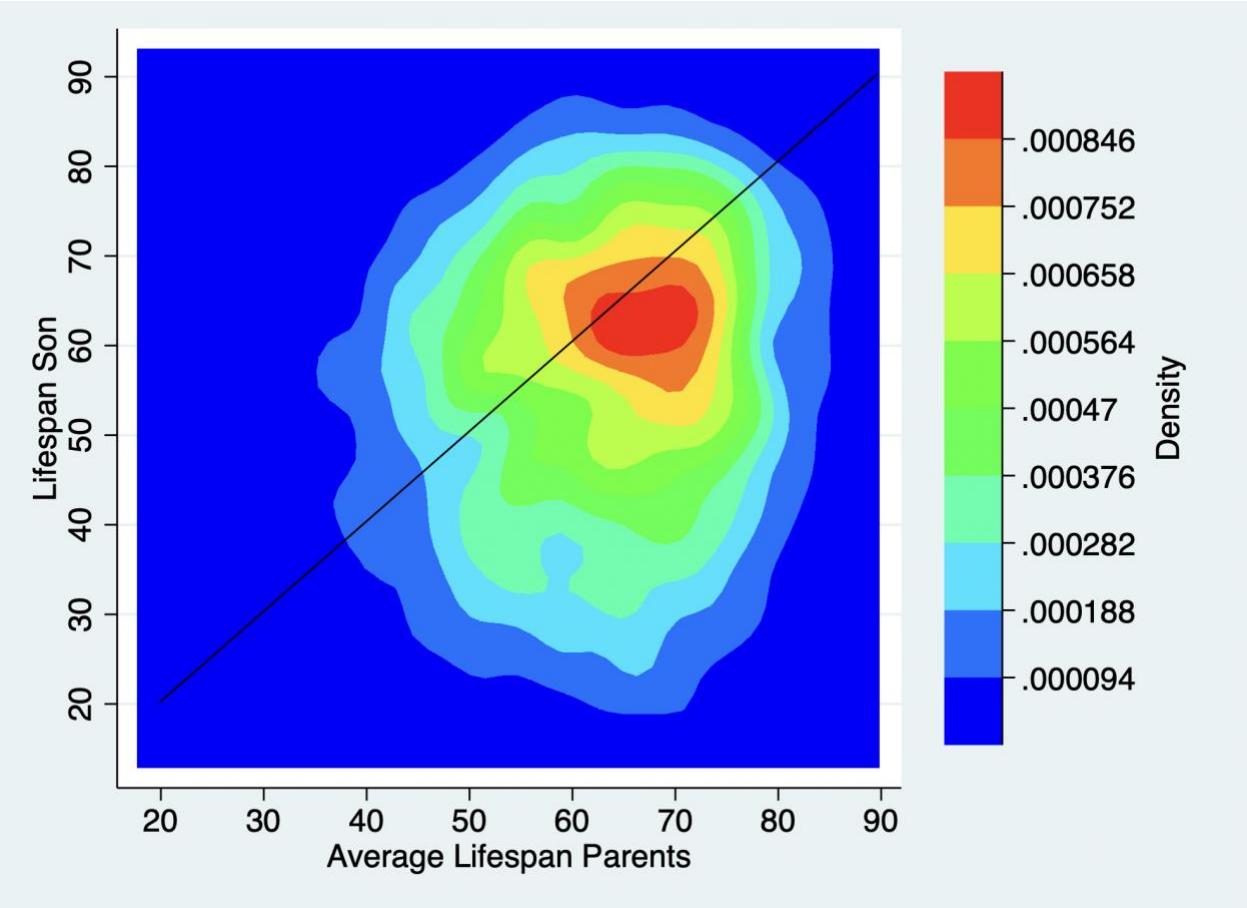
Figure 12: Lifespan Densities of Father and Son



Notes: All fathers and sons with at least 15 years of lifespan for son birth cohorts 1675 to 1775.

In addition, we can relate son lifespan to the lifespan of both his parents, father and mother. One way to think about this is that the second parent gives additional information on 'true' health endowment, thereby reducing measurement error. Figure 13 shows lifespan densities of son versus lifespan densities of his parents, measured as the average lifespan of father and mother.

Figure 13: Lifespan of Parents and Lifespan of Son



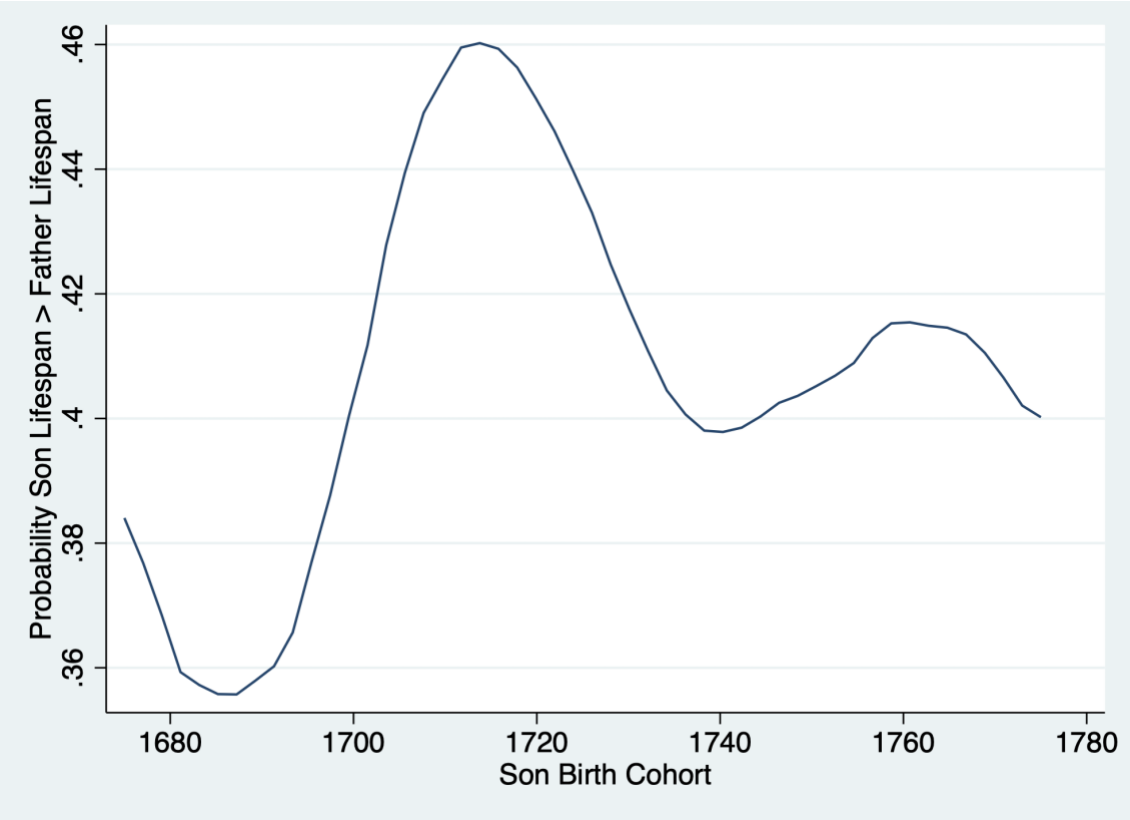
Notes: Average Lifespan Parents is mean of father's and mother's lifespan. All fathers, mothers, and sons with at least 15 years of lifespan for son birth cohorts 1675 to 1775.

Figure 13 indicates that the intergenerational transmission of lifespan from parents to sons is quite uncertain. On the upside, even a son whose mother *and* father do not reach age 40 has a 50% chance of surviving to age 57 or longer. At the same time, when a son has parents that both reach more than 70 years of age, there is still a 25% chance that he does not survive to age 50.

IV.1 Absolute Intergenerational Lifespan Mobility

We begin the analysis of intergenerational lifespan mobility by reporting the probability that the son's lifespan exceeds the lifespan of his father. For birth cohorts 1675 to 1775, this probability is on average 41%. Figure 14 indicates that it varies between 36% and 46% in the beginning of this period, before stabilizing around the average 41% of sons outliving their fathers for sons born between 1730 and 1775.

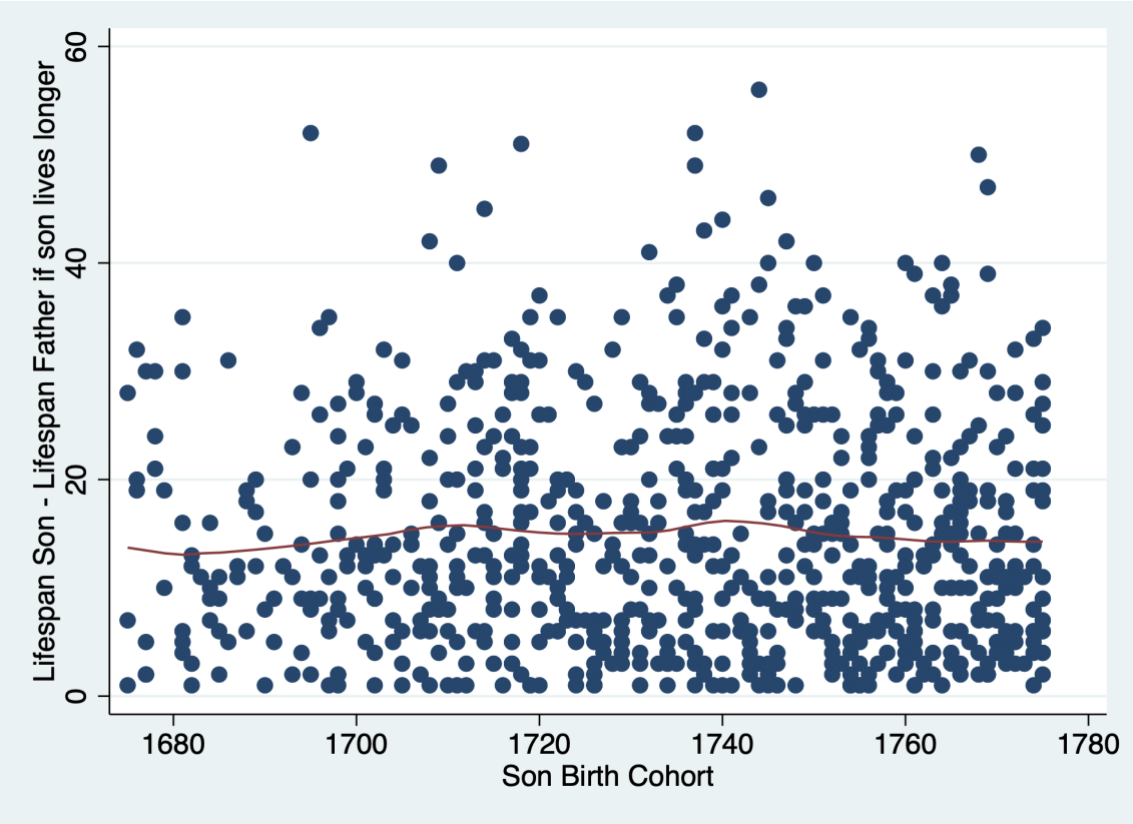
Figure 14: Absolute Mobility in Lifespan, 1675 to 1775



Notes: Figure shows probability that son lifespan exceeds his father's lifespan. All father-son pairs with lifespans for both 15 years or longer, all sons in birth cohorts 1675-1775 (n = 1,938).

Part of the explanation for this pattern is the smaller sample size in earlier years. This is also evident in Figure 15, which reports the extent to which the son outlives his father, conditional on upward lifespan mobility. We see that the extent of upward mobility is roughly constant over time at around 15 years.

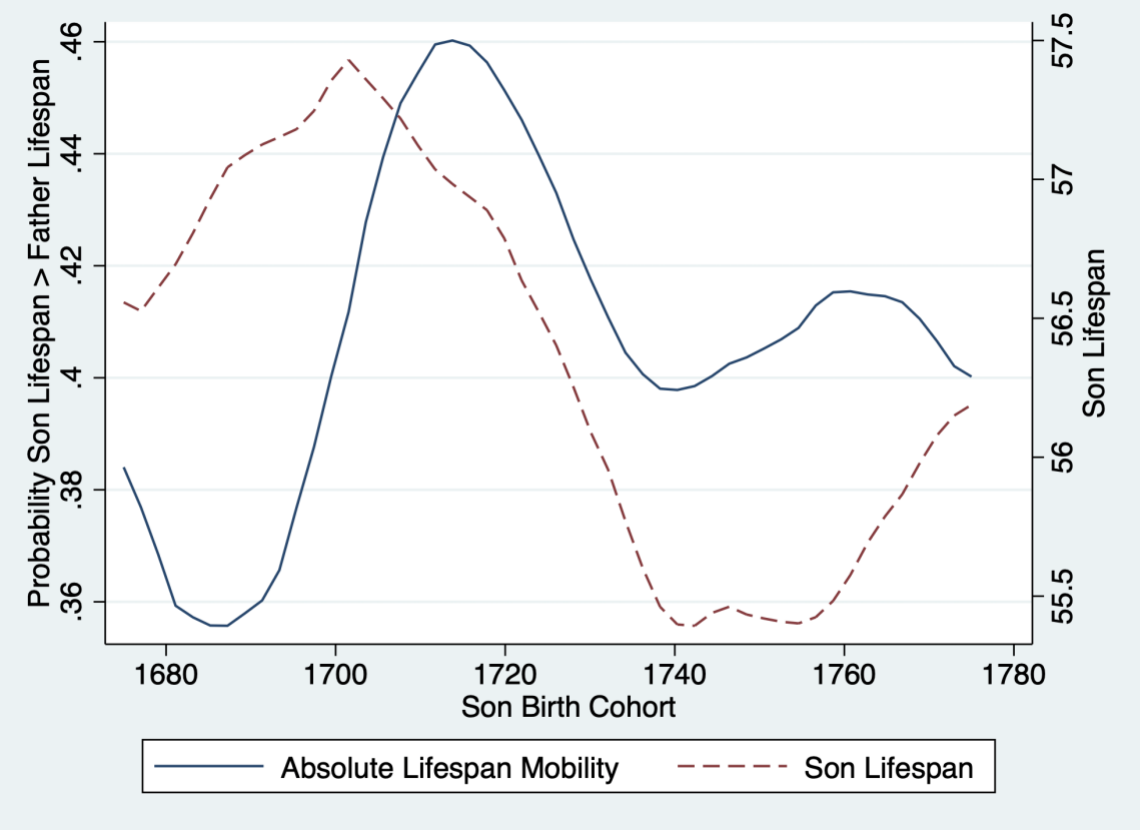
Figure 15: Son Lifespan Premium over Father, Conditional on Upward Lifespan Mobility



Notes: Figure shows difference in son and father lifespan conditional on son living longer than his father. All father-son pairs with lifespans for both 15 years or longer, all sons in birth cohorts 1675-1775.

Figure 16 shows the evolution of upward lifespan mobility together with the trend in son lifespan. From the definitions of these variables, as the typical adult lifespan of sons declines, upward lifespan mobility will become rarer. The finding that changes in absolute lifespan mobility mirror trends in average lifespan is also found for US cohorts of the late 19th and 20th centuries (Black et al. 2024). The correlation between the two series in Figure 16 is 0.51.

Figure 16: Upward Lifespan Mobility and Son Lifespan Trends



Notes: Sample is all father-son pairs with lifespans for both 15 years or longer, all sons in birth cohorts 1675-1775 (n = 1,938).

We have also examined upward lifespan mobility separately by status level. Table 4 shows the results.

Table 4: Upward Lifespan Mobility by Status Level, 1675-1775

Father Status	Prob Son LS > Father LS		Son Status	Prob Son LS > Father LS
35th Percentile	0.431		35th Percentile	0.404
80th Percentile	0.390		80th Percentile	0.381
90th Percentile	0.410		90th Percentile	0.387
97th Percentile	0.361		97th Percentile	0.531

Notes: LS stands for lifespan

We see conditional on the father achieving 97th percentile status, the chance that his son lives longer than the father is a relatively modest 36%. This is plausible, because to the extent that there is a status-lifespan gradient, such 97th percentile fathers will tend to have high lifespans. To attain the highest-ranked civil service exam level, the father would have had to have been healthy into his 30s, and there would likely be some regression to the mean among descendants. Conversely, despite a trend towards lower adult lifespan over

this period, if the son achieves the 97th percentile of the status distribution, the chance that such a high-status son would have a longer lifespan than his father is a relatively high 53%.

To abstract from such effects, we can focus on father-son pairs for which there is little intergenerational status mobility. For robustness, we also aggregate to two status groups, No Status and 'At least Some Status' (i.e., passed at least the lowest level of the civil service examination or has evidence of wealth). We obtain the results shown in Table 5.

Table 5: Upward Lifespan Mobility by Status

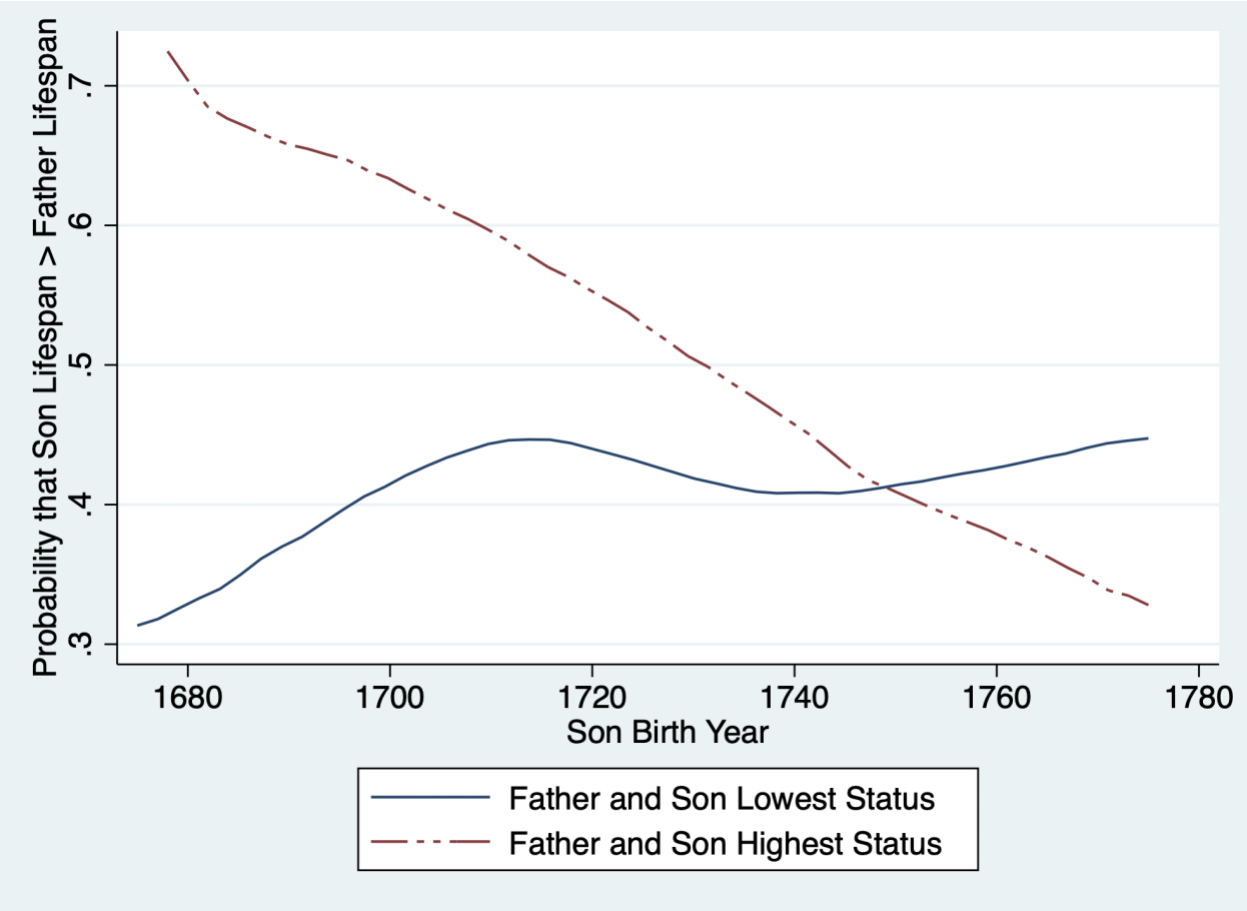
		Father	
		No Status	Status
Son	No Status	0.421	
	Status		0.405

Notes: Probability that son has higher lifespan than father for all sons in birth cohorts 1675 to 1775 and at least 15 years of lifespan. N = 1,000 for No Status/No Status, and N = 565 for Status/Status.

Focusing on the main diagonal of Table 5 are intergenerational health transitions in which status mobility is limited. We see that for these cases, upward lifespan mobility for richer families is somewhat lower than for poorer families, 40.5% versus 42.1%.

Consistent with these patterns in upward lifespan mobility, we depict in Figure 17 the absolute lifespan mobility trends for no-status-mobility father-son pairs at the top and the bottom of the status distribution.

Figure 17: Absolute Lifespan Mobility at the Extremes of the Status Distribution



Notes: Results are for transitions Father 35th Percentile Status and Son 35th Percentile Status, versus Father 97th Percentile Status and Son 97th Percentile Status; sample is all father-son pairs with at least 15 years of lifespan and sons in birth cohorts 1675-1775.

We see that upward lifespan mobility for intergenerationally stable rich families becomes increasingly rare, whereas upward lifespan mobility for stable poor families is rising from around 30% to more than 40%.

To sum up, upward lifespan mobility for birth cohorts 1675-1775 in Tongcheng is around 41%. Furthermore, consistent with some existing research, changes in upward lifespan mobility are related to the trend in average son lifespan.

IV.2 Relative Intergenerational Lifespan Mobility

The following presents intergenerational lifespan movements from a particular quartile in the lifespan distribution of the fathers to a particular quartile in the lifespan distribution of the sons. In this analysis, the overall number of observations is n= 1,989, and the median lifespan of the son in quartiles 1 to 4 is equal to 36, 51, 62, and 73 years, respectively.

Table 6: Lifespan Transition Matrix 1675-1775 (in percent)

		Lifespan Father				
		Q1	Q2	Q3	Q4	Total
Lifespan Son	Q1	26.95	23.52	20.04	21.10	22.88
	Q2	21.19	21.68	21.11	25.50	22.47
	Q3	25.10	27.40	30.06	23.49	26.40
	Q4	26.75	27.40	28.78	29.91	28.26
	Total	100	100	100	100	100

Table 6 indicates a relatively high degree of lifespan mobility in this population. Full lifespan persistence would mean that the cells on the main diagonal of Table 6 would equal 100%--however, they average only 27.15%. This is relatively close to the case of perfect mobility, in which each cell accounts for 25% of the column sum. A high degree of lifespan mobility has been documented in 20th century US populations as well (Black et al. 2024), and is consistent with the generally high intergenerational health mobility estimated for other metrics of adult health (Halliday et al 2024).

A closer look reveals several interesting findings. First, the highest level of lifespan persistence is obtained for father-son pairs in the highest lifespan diagonal cells (around 30%), with both genetic as well as socio-economic factors likely playing a role. Upward mobility, in contrast, is closer to random (e.g., Q1 father to Q4 son of 26.75%). Second, for a given father quartile, the transition probabilities are not necessarily monotonic by son quartile. For example, the chance that Q1 lifespan father has a Q3 lifespan son is higher than the probability that such a father has a Q2 lifespan son. In addition to sample size, this suggests that random health shocks play a significant role.

Third, a simple measure of persistence is the size of the main diagonal cells (that the entry exceeds 25%). Table 6 indicates that there is persistence in this sense for only 3 of the 4 quartiles, another indication of relatively high lifespan mobility.

Next, we relate son lifespan to the lifespan of his father or mother in an OLS regression. This familiar approach in research on mobility has the advantage that it is straightforward to include successively more control variables, **X**:

(1) Son_lifespan = b0 + b1*Parent_lifespan + b'**X** + e,

Where b0, b1, and b are parameters to be estimated, and e is the mean-zero but possibly heteroskedastic regression error. The higher is b1, the stronger is the intergenerational transmission of lifespan, or the stronger is intergenerational lifespan persistence. Quantitatively, the estimate b1 indicates the predicted increase in son lifespan for one more year of lifespan of the parent.

Table 7: Intergenerational Lifespan Mobility Regressions, 1675 - 1775

	(1)	(2)	(3)	(4)	(5)
Father Lifespan	0.062**	0.063**	0.060**		
Mother Lifespan				0.095***	
Parent Lifespan					0.176***
Son Birth Year FE		Y	Y	Y	Y
Father Birth Year FE		Y	Y		Y
Mother Birth Year FE				Y	Y
Clan FE			Y		
N	1,896	1,889	1,889	1,737	1,662

Notes: Dep var is son lifespan in years; parent lifespan defined as average of mother and father lifespan; sample is all sons in birth cohorts where mother, father, and son reach at least 15 years; */**/*** indicates sig at 10/5/1% level w/ robust s.e.

We see that the simple correlation of father and son lifespan is about five percent (column (1)). It means that 1 year more in the father generation translates into 1/20 years in the son generation. This estimate relatively close to zero confirms results from the transition matrices that indicate high levels of intergenerational lifespan mobility. For comparison, in a larger sample for the US at the turn of the 20th century, Black et al. (2024) estimate b1 at 0.089 (see Table 3 of Black et al. 2024). The difference between the two estimates may not only indicate higher levels of mobility in Qing China, however. Another factor is likely that measurement error in our data is higher than in the more recent US data. To the extent that there is classical measurement error in father lifespan, it will bias the regression coefficient towards zero.

Temporary shocks do not strongly affect the intergenerational transmission of lifespan. Specifically, including both son and father birth year fixed effects does not change the b1 estimate significantly (column (2)).

Another factor that may be playing a role is the lineage or clan that the parent-child pair belongs to (the sample includes 7 different clans). In the case of status mobility, Shiue (2025) shows that clans are a significant force towards status persistence, because they redistribute resources among their members. In the case of lifespan mobility, we do not find this effect. If clans were redistributing lifespan-relevant resources among their members, the b1 estimate in column (3) would be lower than in column (2); however, we see from Table 1 that the b1 estimates are similar.

Lifespan transmission from mother to son is estimated to have been higher, with b1 equal to 0.095, see column (4). It is interesting and consistent with other evidence that mother lifespan plays a larger role for son lifespan than that of his father. One possible explanation is that mothers are more important than fathers in teaching the son skills that are health

relevant (or that the health and survival of a son was more important for older women than for older men).

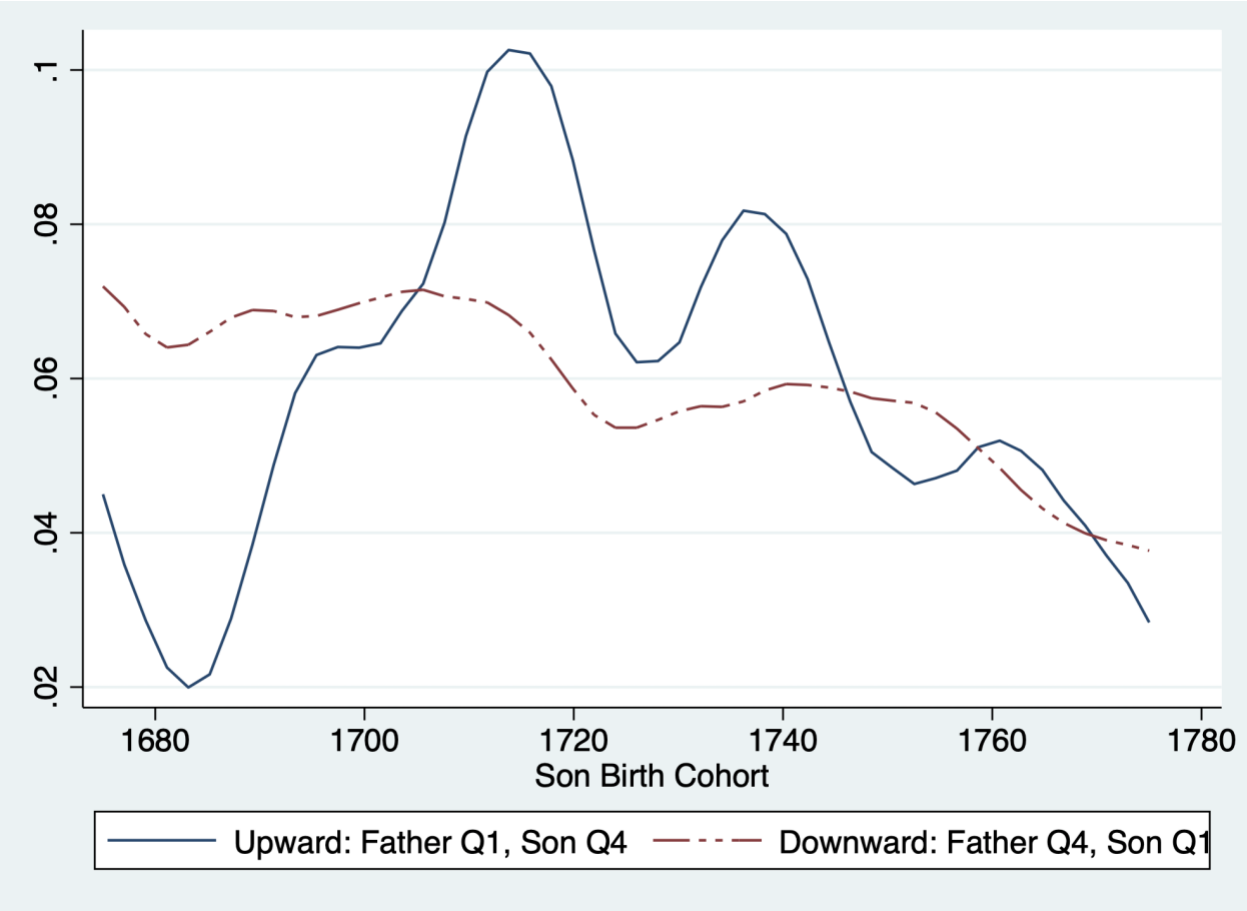
Finally, a regression of son lifespan on parent lifespan, as the average of father and mother lifespan, yields a b1 coefficient of 0.176, see column (5) of Table 7. Thus, one year longer parental lifespan translates typically into about 0.18 years longer lifespan for the son. The fact that the coefficient is higher than either father or mother lifespan alone is in line with attenuation bias. Using only one parent's lifespan does not characterize the health endowment of the parents as well as using information from both parents.

A higher level of intergenerational lifespan persistence when using data for both parents, as opposed to only data from one parent, is found in previous studies as well (e.g., Black et al. 2024). Furthermore, analyzing the US for the late 19th/early 20th century, these authors estimate a degree of intergenerational lifespan persistence quite similar to what is estimated here (0.141, versus 0.176 here; Black et al. 2024, and column (5), respectively). Given that the US economy by the early 20th century was richer than Tongcheng in the 17th and 18th century, we conclude that higher average income does not necessarily translate into more (or less) intergenerational lifespan mobility.

We are also interested in any changes in relative lifespan mobility over the sample period. A first approach is to consider the intergenerational lifespan transition matrix, as in Table 3, and consider variation across birth cohorts over time. We have done this separately for the 1675-1715 son birth cohorts versus the 1735-1775 son birth cohorts. As a simple measure of status persistence, we consider the average of the cells on the main diagonal. For the earlier cohorts, 1675-1715, the average is 25.3%, while for the later, 1735-1775 cohorts the average is 27.9%. This points to an increase in intergenerational lifespan persistence.

Second, we have computed measures of relative mobility for different birth cohorts based on the quartiles of the sons' and fathers' lifespan distributions. A measure of downward mobility is that the father is in the highest and the son in the lowest quartile of the lifespan distribution. The analogous measure of upward mobility is that the father is in the lowest while the son in the highest quartile of the lifespan distribution. Figure 18 shows both measures across cohorts.

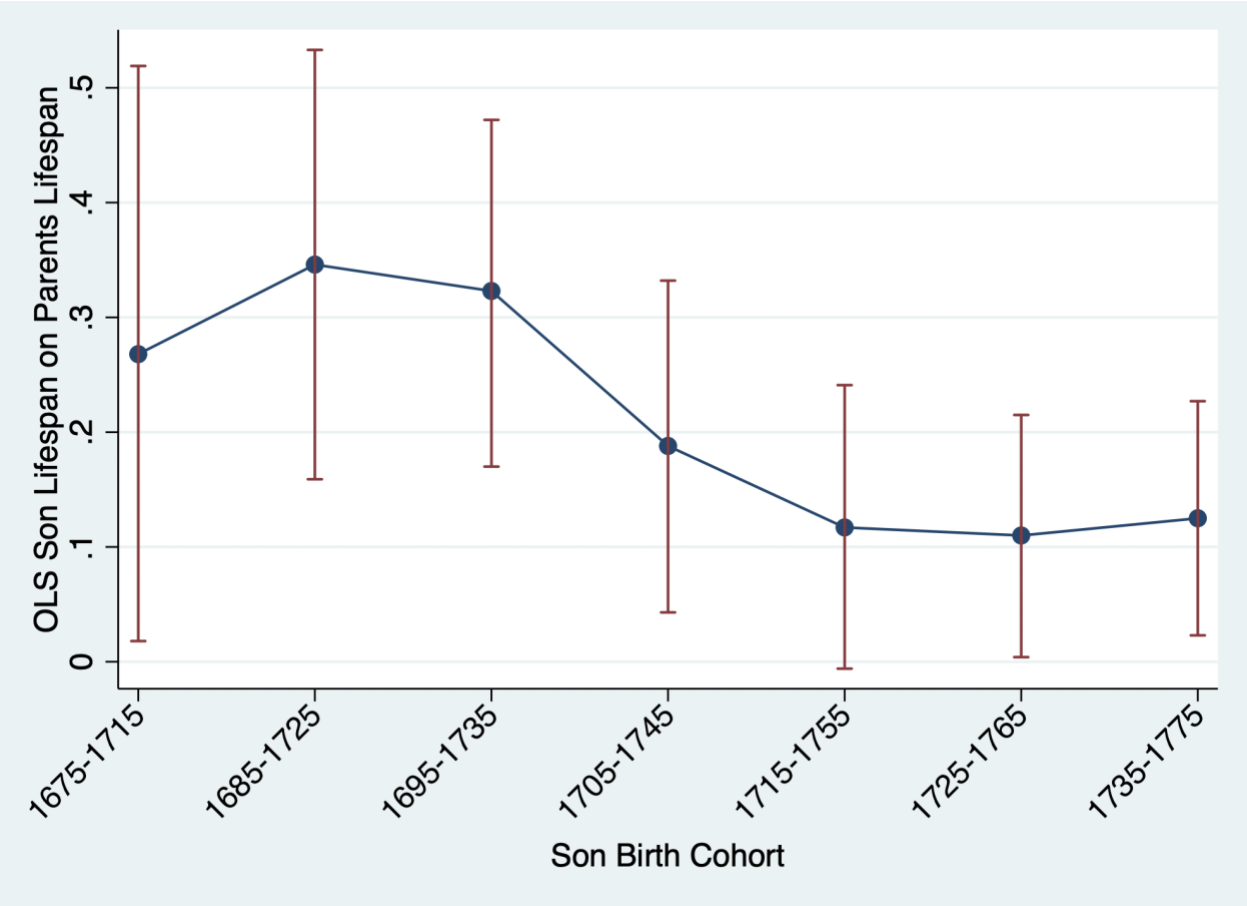
Figure 18: Trends in Upward and Downward Mobility based on Lifespan Quartiles



Across birth cohorts, there is a good amount of variation in the upward mobility measure, with relatively high probabilities in the 1710s, and a falling trend thereafter. The downward mobility measure changes relatively smoothly in comparison; it falls from around 7% in the late 17th century to about 4% by 1775. Figure 18 suggests that big intergenerational lifespan movements between father and son became rarer over the 18th century.

Third, we have employed the lifespan mobility regression, equation (1), with parent lifespan as the independent variable for 40-year rolling birth cohorts, see Figure 19.

Figure 19: Relative Lifespan Mobility: Regression Results for Rolling 40-Year Windows



Notes: Shown are b1 coefficients and robust 95% confidence intervals from running equation (1) on the sample defined by the son birth cohorts indicated.

We see that the b1 estimate of 0.147 for the entire period masks a substantial amount of heterogeneity across birth cohorts. Specifically, the b1 estimate falls from about 0.3 to around 0.11 from the first to the second and third of the birth cohorts shown in Table 3. This points to more lifespan mobility over time.

It is also interesting to examine relative lifespan mobility using regressions that distinguish levels of social status.

Table 8: Intergenerational Lifespan Mobility Regressions by Status

	(1)	(2)	(3)
	All Status Levels	Lower 70 Percent	Upper 30 Percent
Parent Lifespan	0.151***	0.095**	0.227***
	(0.033)	(0.044)	(0.048)
N	1,676	922	754

Notes: Dep var is son lifespan in years; all sons in birth cohorts 1675 to 1775; parent lifespan is average of father's and mother's lifespan; robust standard errors in parentheses; */**/*** indicates sig at 10/5/1% level.

Table 8 provides evidence that lifespan mobility for families of relatively high social status is lower than for families at the bottom of the status distribution.

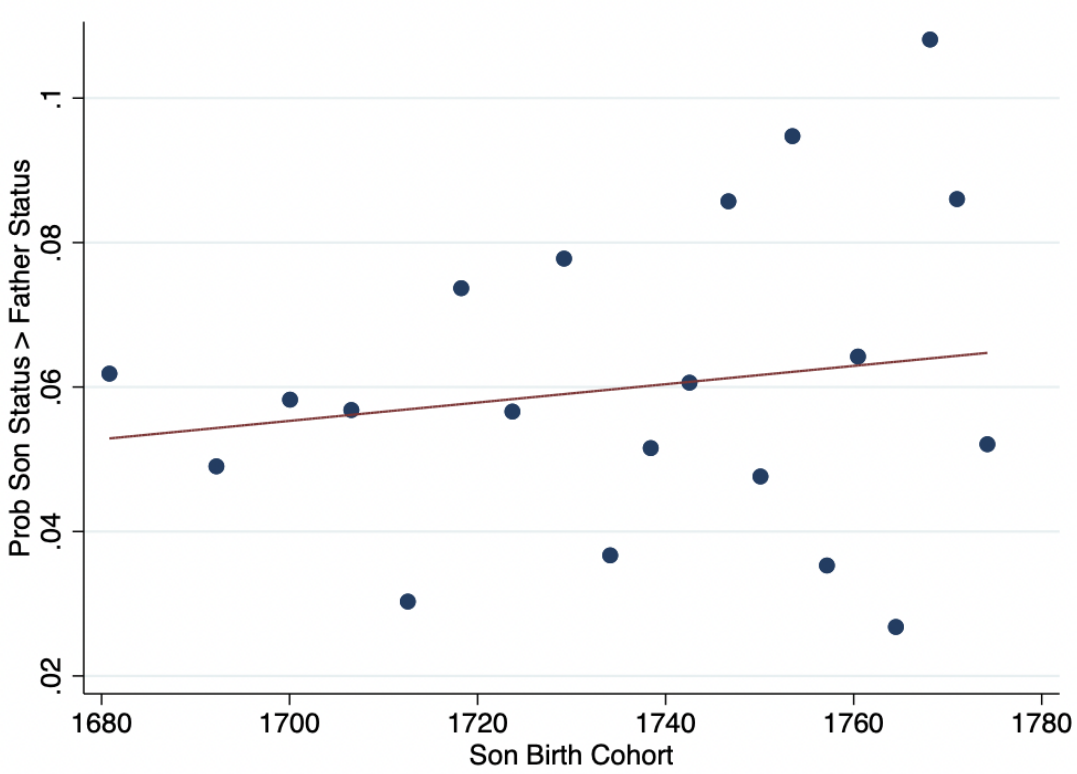
In sum, while intergenerational regressions point to more lifespan mobility over time in our sample, other measures suggest that mobility was flat or even decreased over the sample period. This highlights the importance of broadening the analysis to a range of mobility measures. Across different levels of social status, there is more agreement among different mobility measures. We typically find a relatively high level of lifespan persistence for men that belong to relatively high-status families.

V. Intergenerational Transmission of Social Status

This section provides results on status mobility in this society. This is the topic covered in the recent work by Shiue (2025) so we will be briefer than in our analysis of intergenerational lifespan mobility.

Beginning with absolute mobility, we distinguish the same four status groups as in section III.1 of the paper, referred to as the 35th, 80th, 90th, and 97th percentile, respectively. Figure 20 shows the probability of upward status mobility across birth cohorts, defined as the son having a higher status percentile compared to his father.

Figure 20: Absolute Status Mobility across Birth Cohorts



Notes: Shown is probability that son status exceeds father status for son birth cohorts 1675 to 1775, both father and son lifespans 15 or more years.

We see that the probability a son achieved higher status than his father was only around 6 percent during this period, albeit increasing over time. A key reason why it was difficult for sons to achieve a higher status of their fathers was that the number of official positions in the Qing government was kept roughly constant even though China's population was growing substantially, leading to a crowding out effect. Figure 20 also indicates that absolute status mobility fluctuated increasingly in the short run towards the end of the sample period.[13]

Turning to relative status mobility, the status transition matrix between father and son is shown in Table 9 (n = 1,851).

[13] For comparability with the lifespan mobility results, only observations for which full vital statistics are available are employed to create Figure 19. At the same time, there is also an increase in absolute mobility when we augment the sample with observations using estimated vital statistics.

Table 9. Transmission of Status between Father and Son, 1675 to 1775

		Status Father					
		35th Perc	80th Perc	90th Perc	97th Perc	N	Total
Status Son	35th Perc	89.61	87.46	3.90	0.00	1,257	67.91
	80th Perc	7.44	12.54	65.58	40.48	339	18.31
	90th Perc	1.79	0.00	30.52	23.13	135	7.29
	97th Perc	1.16	0.00	0.00	36.39	120	6.48
	N	1,116	287	154	294	1,851	
	Total	100	100	100	100		100

Comparing the status transition matrix with the lifespan transition matrix, it is important to keep in mind that in contrast to the lifespan quartiles, status classes are not of the same size. For example, the last column indicates that about 2/3 of all sons belong to the lowest class ("35th Percentile"), while 6.48% of sons belong to the highest status class ("97th Percentile").

Table 9 indicates a strong degree of status persistence. That is particularly true at the very top, where more than one third of the top status fathers have sons that also achieve top status (36.39%). If there were full status mobility, based on the sizes of top father and top son classes one would expect less than 1% of all transitions to be father 97th Percentile to son 97th Percentile, but in fact this cell of intergenerational status transition accounts for 5.78% of all transitions (not shown).

There is also strong evidence for status persistence for commoners ("35th Percentile"), with almost 90% of commoner fathers having a son who is also a commoner. Table 9 also shows that certain transition cells do not occur in our sample, for example the father is an official student (90th percentile) and the son is a *jinshi* (97th percentile). This reflects to some extent our moderate sample size. In the analysis below the issue will be addressed by adopting a more aggregate status definition (No Status versus At Least Some Status). The average of the main diagonal entry in the status transition matrix is 42.3%, compared to 26.1% in the case of the lifespan transition matrix. This also shows that there is less intergenerational status mobility than intergenerational lifespan mobility.

Table 10 shows results from status mobility regressions analogous to the lifespan mobility regressions of equation (1). For all birth cohorts from 1675 to 1775, we obtain a coefficient estimate of 0.665. Given that father status is measured in percentiles, varying from 0 to 100, column (1) indicates that a 10 percent higher father status translates on average into a 6.7 percent higher son status. Thus, in his intergenerational relationship, the son's social status is far from random. Instead, the son typically carries over the status of his father with a weight of two thirds, and only the remaining one third is determined by the son.

Table 10: Intergenerational Status Mobility Regressions

	(1)	(2)	(3)	(4)	**(5)**
		Father & Son Fixed Effects	Clan Fixed Effects	Years 1675 to 1715	Years 1735 to 1775
Father Status	0.665*** (0.014)	0.657*** (0.016)	0.606*** (0.018)	0.699*** (0.031)	0.639*** (0.021)
Son Birth Year FE		Y	Y	Y	Y
Father Birth Year FE		Y	Y	Y	Y
Clan Fixed Effects			Y		
N	1,936	1,928	1,928	482	1,094

Notes: Dep var is son percentile status; shown are slope coefficients from OLS regressions analogous to equation (1). Sample is all father-son pairs for sons in 1675 to 1775 birth cohorts, father and son 15 years of lifespan or more. Robust standard errors in parentheses; */**/*** indicates sig at 10/5/1% level.

Column (2) of Table 10 indicates that the intergenerational status relationship is not strongly affected by shocks in particular years. At the same time, we see that the intergenerational regression coefficient is somewhat smaller upon the inclusion of clan fixed effects, indicating that clan differences matter. In particular, membership of a particular father-son pair in a particular clan is a force towards status persistence. The reason is likely that clans redistribute resources from their richer to their poorer clan members.

The last two columns of Table 10 show that relative status mobility increased during the sample period. This confirms existing evidence for a related sample (Shiue 2025). In the case of social status, there is a trend towards more mobility during the sample period for both absolute and relative mobility (Figure 19 and Table 10, respectively).

Furthermore, the evidence for increased status mobility parallels our finding of greater relative lifespan mobility (see Table 8), although quantitatively, the change in status mobility appears to be more moderate compared to the halving of the coefficient in the relative lifespan mobility regression.

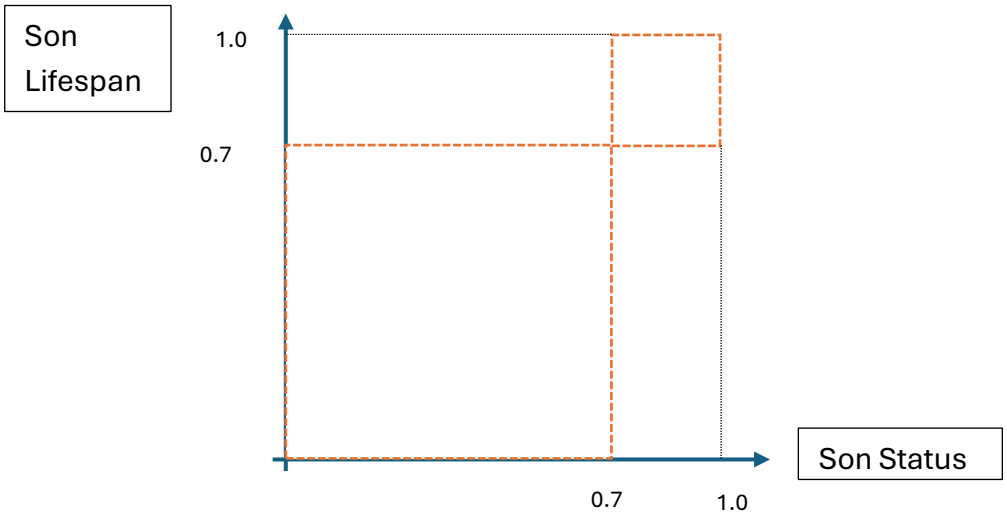
In sum, for the period 1675 to 1775 we find a higher degree of intergenerational status persistence than lifespan persistence, however there is evidence for more mobility over time in both aspects of well-being.

VI. The Origins of the Health-SES Gradient (aka Preston Curve)

The positive association between health and status in any given generation may arise from health (and its intergenerational transmission) enabling status achievement, or status (and its intergenerational persistence) bringing better health, or both. This section provides new evidence on the origins of the health-status gradient (aka Preston Curve) by exploiting information on lifespan and status in the father generation when asking under which conditions the status-lifespan gradient in the younger generation is strong.

Consider the simplified setting of the Preston Curve in Figure 21.

Figure 21: Two-Category Preston Curve in Son Generation



Suppose we consider two status levels, high and low, with a cutoff at the 70th percentile. *StatusLow_son* is equal to 1 if the son has status below the 70th percentile, and zero otherwise. For symmetry, let us consider also two lifespan levels, high and low, with a cutoff at the 70th percentile, and *LifespanLow_son* is equal to one for sons whose lifespan is below the 70th percentile. Then, combinations of son status and son lifespan that generally support the presence of a Preston Curve in this sample would be observations located in the two brightly orange-colored rectangles in Figure 21. They are low lifespan/low status, and high lifespan/high status. We code such observations with a one, and zero otherwise. This indicator variable is denoted by *PC_son*.

Turning to status as the first intergenerational relationship, consider a graph similar to Figure 21 but with the axes being Father Status and Son Status. For both generations, we have Low Status (below 70th percentile) and High Status (above 70th percentile). Analogous to the orange rectangles in Figure 21, a strong intergenerational status relationship between a given father and son is present if either both father and son have low status or if

both have high status. If so, our new indicator variable *IG_stat* is equal to one, while it takes the value zero otherwise.

We proceed analogously with the second intergenerational relationship, that between father and son lifespan. The indicator variable *IG_life* is equal to one if either both father and son have low lifespan or if both have high lifespan, and zero otherwise. The purpose of this exercise is to see whether *PC_son* can be explained by either *IG_stat* or *IG_life*, or by both, and if so to what extent. We also define a Preston Curve variable in the father generation, *PC_fath*, analogously to *PC_son* (and Figure 21).

Table 11: Summary Statistics for Stylized Preston Curve

	N	Mean	Std Dev
PC_son	1,697	0.573	0.495
PC_fath	1,697	0.538	0.499
IG_life	1,697	0.561	0.496
IG_stat	1,697	0.805	0.396

Notes: Variable definitions as described in text; sample is all observations for birth cohorts 1675 to 1775 for which son, mother, and father reach at least 15 years of age.

Notice that the means of three of these indicator variables are similar and around 0.55, whereas the mean of the intergenerational status indicator *IG_stat* is higher. This reflects the fact that father status is highly predictive for son status, as shown above. Bivariate correlations for these variables are shown in Table 12.

Table 12: Bivariate Correlations

	PC_son	*PC_fath*	*IG_stat*	*IG_life*
PC_son	1			
PC_fath	0.063	1		
IG_stat	-0.061	0.096	1	
IG_life	0.106	0.033	-0.001	1

Notes: sample is all observations for birth cohorts 1675 to 1775 for which son, mother, and father reach at least 15 years of age (N = 1,697).

We see that the strongest correlation of *PC_son* is with the intergenerational lifespan variable *IG_life*, and it is positive. It means that there tends to be a strong Preston curve in the son generation whenever father lifespan is highly predictive of son lifespan.

We know that father lifespan is not predictive of son lifespan all that often—from Table 11, it is so in 56% of all cases, while in 44% of cases it is not. But when it is, we tend to observe a Preston curve in the son generation, according to Table 12 (positive correlation of 0.106). In comparison, father status is highly predictive of son status in 80% of the cases, however, that is not positively correlated with the presence of a Preston curve in the son generation (correlation of -0.061, see Table 12).

Note that these findings are robust to conditioning on lifespans for father and son of at least 30 years (to account for attainment of top status being in effect conditional on survival to that age). They are also robust to using only data on fathers and sons, not mothers, which yields a roughly 15% larger sample.

We expand on these findings using regressions, see Table 13. This has the advantage of allowing inclusion of covariates, and it might also lend itself to a causal interpretation.

Table 13: Determinants of the Son Preston Curve

	(1)	(2)	(3)	(4)	(5)	(6)
IG_stat	-0.076**		-0.076***	-0.084***	-0.053	-0.060*
	(0.030)		(0.030)	(0.030)	(0.033)	(0.034)
IG_life		0.106***	0.106***	0.103***	0.095***	0.095***
		(0.024)	(0.024)	(0.024)	(0.026)	(0.027)
PC_fath				0.066***		0.046
				(0.025)		(0.028)
Son Birth Year FE					Y	Y
Father Birth Year FE					Y	Y
Clan FE					Y	Y
Adj R2	0.003	0.011	0.014	0.018	0.040	0.041
N	1,697	1,697	1,697	1,697	1,688	1,688

Notes: Dep var is *PC_son*; sample is all sons in birth cohorts 1675 to 1775 with father, mother, and son lifespan of at least 15 years; robust standard errors in parentheses; */**/*** indicates sig at 10/5/1% level.

We see that a strong intergenerational lifespan relationship is predictive of the Preston curve even in the presence of additional covariates. The same finding is obtained if intergenerational lifespan mobility is measured using both mother and father lifespan, as opposed to only father lifespan. For example, the coefficient of *IG_life* in a regression based on parent instead of father lifespan analogous to the specification underlying column (4) is equal to 0.107 (s.e. of 0.027), which is similar to the result obtained in Table 13. In contrast, there is no evidence to suggest that strong intergenerational status transmission is behind the Preston curve in the son generation. We see that the correlation between *IG_stat* and *PC_son* tends to be low and if anything it is negative.

We push this analysis further to think about the causal relationship between lifespan (as proxy for health) and status. If strong intergenerational lifespan transmission predicts the Preston curve in the younger generation, does high lifespan of the parents predict high son status? Mother and father lifespan may not be exogeneous, but they are somewhat pre-determined (although median survival among fathers overlaps with son survival beyond

age 15, and having a surviving healthy son may provide vital 'pension support' enabling longer parental lives). In the same spirit of using different generations to address the simultaneity issue between health and income, Case, Lubotsky, and Deaton (2005) relate child health to parent income (since children do not have income of their own yet).

A simple OLS regression of son percentile status on the average lifespan of his father and mother yields a coefficient of 0.142 (column (1), Table 14). Because we measure lifespan in units of 100 years from now on, the magnitude of this estimate indicates that a 10 year longer parent lifespan is typically associated with a 1.4 percentile point higher rank in the status distribution.

Table 14: Son Status and Parents Lifespan versus Son Lifespan and Father Status

	(1)	(2)	**(3)**	**(4)**
Dep. Var.	*Status Son*	*Lifespan Son*	*Status Son*	*Lifespan Son*
Lifespan Parents	0.142*** (0.054)		0.172*** (0.058)	
Status Father		-0.328 (1.399)		-3.260* (1.909)
Son Birth Year FE			Y	Y
Father Birth Year FE			Y	Y
Mother Birth Year FE			Y	Y
Clan FE			Y	Y
N	1,696	1,677	1,682	1,664

Notes: Dep var given at top of column; sample is all sons in birth cohorts 1675 to 1775 for which mother, father, and son live at least 15 years; lifespan is measured in units of 100 years; robust standard errors in parentheses; */**/*** indicates sig at 10/5/1% level

In contrast, higher father status is not associated with longer lifespan of the son (column (2)). These findings do not change with the inclusion of birth year and clan fixed effects, see columns (3) and (4) of Table 14.

We are now ready to run formal regressions with the goal of estimating the causal effect of health on status. Given the results presented so far, it is natural to consider mother and father lifespan as instrumental variables for son lifespan. The following presents instrumental-variable results based on this approach. The exclusion restriction that needs to hold to obtain the causal effect of son lifespan on son status is that parent lifespan only affects son status through its effect on son lifespan. This will be discussed further below. Notice that in a setting where mother and father lifespan instrument for son lifespan, column (1) of Table 14 presents the reduced-form results. The finding that higher lifespan of parents has a positive reduced-form impact on son status suggests that the instrumental-variables approach may work.

Table 15 presents the instrumental-variables results of regressing son status on son lifespan, which is instrumented by parent lifespan. Without additional covariates, we obtain the results shown in column (1). According to the first-stage results parent lifespan has strong predictive power for son lifespan. The coefficient is equal to 0.15, which confirms our earlier results (Table 7, column (1)). The robust Kleibergen-Paap F-statistic of 21.5 indicates that the first-stage correlation is strong. The second-stage coefficient is estimated at 0.97, suggesting that one year more of lifespan leads to roughly one percentile point higher son status.

Next we employ father and mother lifespan as separate instrumental variables. The advantage of that is that we can employ overidentification tests to shed light on the validity of the instruments. Notice that in the first-stage, both father and mother lifespan are positively correlated with son lifespan (although the mother point estimate is somewhat higher). The second-stage coefficient is quite similar to before. The main novel result is the J statistic from the Hansen overidentification test. This statistic essentially compares the second-stage results using different sets of instrumental variables. The larger the differences, the stronger is the evidence against instrument exogeneity. The p-value of the Hansen J overidentification test is equal to 0.215, indicating that second-stage results with either father or mother lifespan as the instrument are similar. This provides evidence in favor of instrument exogeneity.

The following columns report results for specifications with additional covariates, thereby reducing the likelihood that spurious correlation affects the findings. We begin with the inclusion of son birth year fixed effects. This reduces the second-stage coefficient somewhat, from 0.97 to 0.88, in line with the idea that these fixed effects reduce the influence of shocks (column (3)). At the same time, when both parent lifespan variables are employed, the second-stage coefficient is similar to before (compare Table 15, columns (4) and (2), respectively).

Next, we include an additional series of fixed effects, namely for father birth year, mother birth year, as well as for each clan. The impact of son lifespan on son status in the second stage is estimated with a coefficient of 1.1, see columns (5) and (6). The first-stage F-statistic is between 8 and 16, somewhat lower than before. This is not surprising, because some variation in the first stage is now attributed to these fixed effects. Interestingly, conditional on these covariates the p-value of the overidentification test is close to 1, that is, virtually no evidence against exogeneity of the instrumental variables (see column (6)).

It is well-known that the two-stage least squares approach can be unreliable in the case of a weak first stage, while a limited-information maximum likelihood (LIMIL) approach is more robust. We therefore present LIML results as well, see column (7). It turns out that they are similar to the corresponding two-stage least squares results, with second-stage estimates of 1.07 versus 0.97 (see columns (7) and (4)). In addition to strengthening the evidence for a positive effect of health on status, this result confirms that this IV estimation does not suffer from a weak first stage. The final IV specification in Table 15 shows that the efficient GMM approach (using the optimal variance-covariance weighing matrix) leads to broadly similar results (column (8)).

The results present new evidence for an effect of health on status during the period 1675 to 1775. How plausible is it that the IV estimates in Table 15 represent true causal effects? Statistical tests such as the first-stage F-test or the overidentification test have only so much power, and there are clearly numerous channels that could be at work to generate endogeneity between social status and health.

A useful starting point is the OLS relationship between son lifespan and son social status (the son Preston curve). In our sample, a regression of son status on son lifespan yields a coefficient of 0.108, see Table 15, column (9). While both the OLS and the IV estimate are positive, quantitatively, the OLS estimate is substantially lower than the corresponding IV estimate (of 0.88, see Table 15, column (3)). If the IV estimate gives the true causal effect, the OLS estimate is downward biased. What might explain a substantial downward bias of OLS in the present context?

One possibility is that the IV estimation solves an omitted variable problem. For the OLS estimate to be downward biased, the omitted variable must be negatively correlated with son lifespan and positively correlated with son status, or it must be positively correlated with son lifespan and negatively correlated with son status. It is challenging, however, to think of a variable whose correlation with status and lifespan is of a different sign. Instead, if something is positively correlated with status, it tends to be positively correlated with lifespan as well (and analogously for negative correlations).

Another explanation focuses on the fact that even under the best of circumstances, with strong first stage and quasi-random assignment, the IV estimate gives the local average treatment effect (LATE), whereas OLS gives an estimate based on the entire sample. We have seen above that the relationship between status and lifespan has non-linearities (Figure 5). At the same time, these non-linearities seem to be quite moderate compared to the sizable difference between OLS and IV estimates.

It could also be that a strong positive effect of status on health creates a complicated simultaneous relationship between lifespan and status in our setting. At the same time, we have seen above that father-son pairs with strong status transmission do not predict a Preston curve in the son generation, and furthermore, son lifespan is not positively correlated with the social status of the father. These findings do not suggest that the large IV estimate of lifespan on status masks causality that runs in the opposite direction.

Part of the explanation is likely that the IV approach reduces measurement error problems. We have seen above that lifespan of the son varies widely at any son status level, and that the intergenerational correlation between parent and son lifespan is relatively low. Under these circumstances, it is plausible that son lifespan is a relatively imprecise measure of the son's true health endowments; instead, son lifespan is potentially subject to major idiosyncratic health shocks. If so, lifespan of father and mother may have considerable power to better pin down the son's true health endowment.

Table 15: The Impact of Health on Status: Instrumental-Variables Results

	(1)	(2)	(3)	(4)	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Method	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	LIML	GMM	OLS
I. Second Stage									
Son Lifespan	0.970**	1.007**	0.880**	0.971**	1.098***	1.098***	1.077**	1.061**	0.108***
	(0.399)	(0.401)	(0.432)	(0.437)	(0.396)	(0.390)	(0.504)	(0.447)	(0.040)
Hansen OverID p-val		0.215		0.172		0.997	0.184	0.177	
II. First Stage									
Parent Lifespan	0.151***		0.137***		0.169***				
	(0.033)		(0.034)		(0.043)				
Father Lifespan		0.069***		0.056**		0.063*	0.056**	0.056**	
		(0.026)		(0.028)		(0.033)	(0.028)	(0.028)	
Mother Lifespan		0.081***		0.079***		0.102***	0.079***	0.079***	
		(0.023)		(0.024)		(0.029)	(0.024)	(0.024)	
First Stage Robust F	21.50	10.81	16.08	8.22	15.79	8.18	8.22	8.22	
Son BY FE			Y	Y	Y	Y	Y	Y	Y
Father BY FE					Y	Y			
Mother BY FE					Y	Y			
Clan FE					Y	Y			
N	1,675	1,675	1,675	1,675	1,675	1,675	1,675	1,675	1,675

Notes: Dep. Var. is percentile status son; sample is all observations of sons in birth cohorts for 1675 to 1775 for which mother, father, and son reach at least 15 years of age; robust standard errors in parentheses; */**/*** indicates sig at 10/5/1% level.

VII. Discussion

Utilizing unique panel data on intergenerational health and status persistence, we show that over a century of Malthusian imperial China, there was a nontrivial health-status gradient in adult lifespan, and health persistence was a far better predicter of this health-status gradient than status persistence was. Our findings point to a relatively strong portion of the adult health gradient flowing from health to status between 1675 and 1775 in central China.

These results contribute to an economic understanding of the importance of health for gaining and maintaining status based on distinctive intergenerational data over a relatively long timespan where other evidence is sparse, especially for commoners (representing more than 60% of our data). Our findings of modest adult lifespan gradients and the lack of evidence for status persistence explaining the lifespan gradient are consistent with recent causal evidence on the contemporary health gradient showing little evidence of a causal arrow from wealth to health among those who survive to adulthood (Cutler et al. 2011), as well as some historical studies of 18th century Europe. For example, DeWitte et al. (2015) study London in 1700-1853, concluding that lower SES households suffered from higher infant and child mortality, but there were no evident gradients in adult health and survival. Although selective migration of healthier young adults into the city could also partly explain the findings, their assessment of little gradient in adult lifespan is consistent with numerous studies of disparate parts of industrializing Europe.

We find decreasing adult lifespans across the status distribution in our sample, a trend consistent with evidence on stagnating or declining living standards in China during this period (Broadberry et al. 2018, Chen and Peng 2022), and a substantial decrease in real wages concentrated in the 1740s-1760s (Liu 2024). Baten et al. (2010) note that the anthropometrics of Chinese in Taiwan measured by Japan colonial authorities reveals stagnation and decline in living standards in China by the 1800s. Few other studies have documented the evolution of the health-status gradient during such a period of static or falling living standards, rather than the general increase in living standards and lifespan of the late 19th and early 20th century in Europe and the US. Since our genealogical data primarily covers married men, the health and survival of other population subgroups would presumably further strengthen the disparities in health by status. Indeed, we present some evidence of differences in under-5 and under-10 mortality by status, suggesting that differences in life expectancy at birth may have been even larger that the 5-year gap in adult lifespan between the highest and lowest status groups.

Note, however, that because reductions in child mortality may have led to stable or even increasing life expectancy at birth, our results are not inconsistent with assertions of Chen and Peng (2022) and others that overall human capital in China was remarkably high for low per capita income as well as quite stable despite exploding population. Indeed, Lee, Wang and Campbell (1994) estimate that improvement in child survival led to increasing

life expectancy at birth of the Qing nobility during this period (from low 20s to about age 40), similar to or slightly earlier than among elite Europeans (Lee et al. 1994).

Some observers might consider longer lives among high-status individuals to be obvious and omnipresent, but that has not always been the case. Long spans of history before the modern era appear to lack a significant survival advantage of the elite. The rich do not always indulge in healthy behaviors, and many health-promoting behaviors are relatively inexpensive but time-consuming (e.g., fetching clean water and consistently boiling it before drinking). Ruhm (2000) showed that in recent US history, physical activity increases and diets improve during recessions, with less smoking and reduced obesity. In an agricultural setting, child mortality *increases* in Colombia when world arabica coffee prices are high, a starkly procyclical pattern of child deaths linked to the opportunity cost of behaviors protective of child health (Miller and Urdinola 2010). These patterns are not surprising if we consider the time price of health investments.

While evidence for the mechanisms of child and adult mortality during our period of study is limited, we do know that elite lifestyles were not always conducive to longer lives: their food, though plentiful, did not necessarily reflect a balanced diet, and privileged behavior sometimes undermined health (e.g. overindulgence in alcohol). For example, although breastfeeding infants is protective of both infant and mother -- and poor households had few options but to carry on this tradition -- elite women in ducal families of pre-1750 England employed wet nurses and other infant-feeding practices that undermined maternal and child health (Kendall et al. 2021): high fertility of elite women and closer child spacing, lower immunity of infants and more susceptibility to early death from rampant infectious diseases.

Most previous studies of China investigate lifespans for a selective elite, with averages in the late 50s to 60s that are comparable to the highest-status men in our sample (Chen and Peng 2022, p.697; Lee and Li 2021). For example, Lee and Li (2021) utilize the recorded lifespans of 5000 extremely prominent Chinese elite over two millennia, from Confucius to high-profile officials, scholars, and generals as documented in the "Biographical Encyclopedia of Eminent Chinese in History (BEECH)"; their lifespans were relatively high (concentrated in the 60s and lower 70s) over the Ming and Qing dynasties. This lifespan range for Chinese elites of late imperial China is similar to what we document in our data for high-status men with top-quartile lifespans. The comparative advantage of our data is the inclusion of a robust sample of non-elites showing average lifespans declining with status, linked across generations to study persistence of health and status.

Studies documenting a status gradient in imperial China are even more limited. Analyzing genealogical data on the Qing dynasty nobility, Lee, Wang and Campbell (1994) find that the steepest decreases in infant and child mortality arose during the period between the 1740s and 1760s, similar to what we document. Moreover, they estimate that the overall mortality rates were similar in magnitude to those documented for European elites in the 18th century (e.g. 100 per 1,000 among Qing nobility, compared to Hollingsworth's data of

80 to 100 for British elite boys, and other estimates for continental Europe (Hollingsworth 1957, Lee et al. 1994). While the overall under-5 and under-10 mortality rates in our sample are lower, the pattern of decline is consistent with significant gains also in the mid-1700s, with our high-status households mirroring survival improvements of the Qing nobility and the low-status households in our sample filling in the gaps regarding non-elites and their later improvement in child survival. Indeed, even among the Qing elite, there was significant heterogeneity: the official allocations varied by more than 2 orders of magnitude between the top lineages to the lower nobility, according to Lee et al. (1994). Their evidence of a status gradient in child survival suggests that one mechanism underlying the health-status gradients we document may also have been differential infant and child mortality and selective survival of the least frail of the commoners.

Other potential mechanisms underpinning the differential survival by status and diffusion of health improvement from the elites to the commoners are unclear. Some candidate health production technologies include inoculation against smallpox and other traditional medical practices in China. Although the availability, cost, and uptake of variolation against smallpox in the populations we study is not well documented, this effective prevention strategy was widespread in many parts of China (and elsewhere) before spreading to Europe and North America. Evidence about smallpox inoculation during the 1700s indicates that differential uptake of variolation across the status distribution probably contributed to the gradient in infant and child survival, as well as to the significant improvement during the mid-1700s (Figure 8 father birth cohorts of the late 1600s and early 1700s). Strong negative externalities of smallpox epidemics presumably spurred the diffusion of variolation across the status spectrum, as evident in the government-run variolation clinics in other localities during the period (Lee et al. 1994). There is less evidence about how other practices in Traditional Chinese Medicine (TCM) shaped the health patterns that we document. Most herbal medications are not extremely expensive and some TCM is well known to be quite effective against certain conditions. For example, a TCM-based therapy—artemisinin combination therapy—is now the globally recommended treatment for malaria.

To assess the magnitudes of the seemingly modest adult lifespan gradient we find, we also compare to the earliest available cohort lifetables for European men. This comparison shows that adult lifespan in this relatively well-off Yangzi river region was comparable to contemporary Europe, and that the adult gradient in survival by status was of the magnitude of several decades of survival improvement in industrializing England. Consider our Kaplan-Meier survival estimates by status in Figure 3. They suggest an approximate ten percentage-point difference in survival to age 65 between the highest and lowest status men (conditional on survival to age 15). For comparison, we extract relevant survival percentiles from the cohort lifetables of European countries with good population-wide vital statistics for the 1700s (Sweden) and early 1800s (e.g. England & Wales), as compiled by the Human Mortality Database (HMD). Cohort lifetables represent the appropriate counterpart to our data since they follow birth cohorts over a century until the cohort is extinct, in contrast to period lifetables (which are based on age-specific deaths in

a given year). We use lifetables over five-year periods to help smooth across historical fluctuations during this period.

Specifically, HMD data for Sweden indicates that for men born in 1751-54, survival to 65 conditional on surviving to age 15 was 35%, rising to 41.2% by 1800-1804 and 52.6% for the 1825-1829 birth cohorts. Thus, the S(65|15) survival difference between Chinese commoners and elites was almost twice this 6pp survival improvement in Sweden for the birth cohorts spanning the half-century 1751-1800. For England and Wales, adult life expectancy (i.e., life expectancy conditional on survival to age 15) was 43 for men in the 1841-1845 birth cohorts, rising to 49 in 1870-1874 and 52.7 by the turn of the 20th century (1900-1904). The probability that a man survived to age 65 conditional on reaching 15 was 43% in 1841-44, rising 14 percentage points to 57% in 1870-74. This 14pp improvement in S(65|15) was slightly larger than the difference between the commoners and elite in Qing China.

Our empirical findings are strongly supportive of better health enabling attainment of status in 17th and 18th century China. This pattern is certainly consistent with hardship among orphans, as well as those growing up in families experiencing the loss or disabling of the primary breadwinner, and the role of clan support in such cases. Of course, we recognize that parental lifespan is not completely exogeneous relative to adult child's lifespan; most fathers will be still in their 20s or 30s when their sons are born, and their death is not predetermined relative to the son's adult health and status unless they die young (below the average of 56.5 years). Moreover, having a healthy, successful son probably contributed to parental lifespan by functioning as pension support; indeed, strong son preference arose in part from such filial piety as well as the desire to carry on the family line.

We explore the robustness of our results in various ways. For example, our analysis of intergenerational transmission as a potential determinant of the health gradient examines all married (adult) men in our sample for the most robust period of genealogical recording of vital statistics, to increase precision with larger sample size. One concern with this approach might be that our measure of male status is conditional on survival (and being healthy) to age 30 or so, to be able to take and pass the various levels of exams. (Figure 4 illustrates this pattern clearly.) We address this concern through a sensitivity analysis that defines intergenerational lifespan persistence (*IG_life*) conditional on survival to age 30. We find that lifespan persistence remains a stronger predictor of the health-status gradient than status persistence across generations, despite the latter magnitude far surpassing the former.

The extent to which the intergenerational lifespan correlation is "transmission" or spurious is hard to know. But several factors support our interpretation. First, the intergenerational health association is low (compared to intergenerational status association), and the magnitude closely matches that found by Black et al. (2024) in the US 1880-1920 cohorts. Second, our regression coefficient is almost the same whether one

introduces father and son birth year fixed effects or not. If the correlation were highly sensitive to random shocks, such as famines or epidemics in particular years, we would expect the inclusion of these about 200 birthyear fixed effects to change the coefficient more strongly. The fact that the association remains virtually identical with and without these year fixed effects suggests nontrivial persistence or "true transmission."

We find mothers are especially predictive of son lifespan. This pattern is consistent with much previous literature in health economics of the importance of mothers' knowledge and behaviors for infant and child health outcomes. Better maternal health and education accounted for a significant share of China's dramatic Mao-era increase in lifespans (Babiarz et al. 2015). Women's role in supporting policies prioritizing child health (Miller 2008) and fostering educational human capital has also been well documented, especially for the era before mass education. For example, Althoff et al. (2025) find that human capital mobility in the US increased significantly between 1850 and 1950, "driven by a declining reliance on maternal human capital, which had been most predictive of child outcomes before widespread schooling. Broadening schooling weakened this reliance on mothers, raising mobility in both human capital and income over time" (p.1). Although Althoff et al. focus on education and not health, their analysis reinforces the health economics literature on health human capital (Grossman 1972) and the many previous studies showing increasing economic mobility alongside dramatically declining infant mortality and a health gradient in early 20th century US, especially for some groups. Studying China over a century earlier, we contribute evidence of the importance of health transmission (especially from mothers) for the health gradient in a period of growing economic mobility albeit greater status persistence than health persistence.

Limitations and extensions. Despite the many strengths of our data for studying lifespan, status, and their intergenerational transmission, several limitations should be considered in interpreting our results and assessing their generalizability. The genealogies from Tongcheng represent the population in a small area of central China, and despite being more representative across status than many previously analyzed datasets, they need not be broadly representative of China's heterogeneous regions. We observe health and status along the male line, and only for those sons able to marry; although we can also parse the data to understand marital sorting, maternal mortality, and some aspects of child mortality, the poor recording of live births and neonatal mortality, especially for girls, forces a focus on adult lifespans (i.e., conditional on survival to age 15).

In ongoing work, we further probe the association between intergenerational health persistence and the strength of the health gradient by using additional instruments for adult male lifespan, from grandparental lifespan to the health shock of the Ming-Qing transition (Shiue and Keller 2024). We also explore robustness to accounting for multiple sons across generations, accounting for differential fertility (family size) by status, and examining transmission from parent to longest-lived son. In addition, we apply the statistical decomposition method developed by Althoff et al. (2025) as an intergenerational mobility metric for both parents (e.g., with status and lifespan rather than their income and

literacy, for grandparents as well as parents, and accounting for clan), and adopt their semiparametric latent variable method to study rank-rank relationships with only binary proxies (e.g., parent dead or alive during the son's formative years).

VII. Conclusions

Studying 17th and 18th century China, we document the evolution of the health-status gradient across multiple generations. Based on rich genealogical data that includes clear markers of status defined by China's national civil service examination system, we find that compared to married men without any evidence of wealth or status, those who achieved the highest rank lived on average five years longer. This gradient in adult lifespan by status, although seemingly modest, was of comparable magnitude to several decades of survival improvement in industrializing England. Some suggestive evidence also supports a gradient in maternal and child mortality over this period, indicating that life expectancy at birth may have exhibited an even stronger status gradient.

Leveraging the intergenerational linkages in our data, we find that lifespan persistence is a stronger predictor of the health-status gradient than status persistence across generations, despite the latter magnitude far surpassing the former. Wealth evidently could not buy significantly better health with any degree of certainty in this society. Rather, longer lives allowed time and opportunity to climb the social status hierarchy – in our setting, by enabling healthy men to pass successive levels of the civil service exam. Clan solidarity also meaningfully shaped economic outcomes including intergenerational transmission of status; but within-clan redistribution did not significantly reduce variability of lifespan.

Our findings suggest mobility of well-being -- proxied with 'full income' of years lived at a given per capita income – has long been higher than status persistence alone would suggest, as Black et al. (2024) also emphasize. Even in the era of modern medicine with many expensive therapies extending lives, health transmission appears to be relatively modest, at least in industrialized economies with universal health coverage. Contemporary Taiwan exhibits a rank-rank slope in parent-child health of 0.22 (Chang et al. 2024), only slightly higher than our estimate of 0.176 for lifespan in 18th century China, and comparable to studies of adult health in many modern settings (Halliday 2023) that document much lower health persistence than for income or SES. Any 21st-century breakthroughs that radically extend lifespan with expensive treatment might change this paradigm of considerable 'wellbeing mobility' through health mobility.

References

Althoff, Lukas, Harriet Brookes Gray, and Hugo Reichardt. “America’s Rise in Human Capital Mobility,” working paper, February 18, 2025.

Babiarz, Kimberly Singer, Karen Eggleston, Grant Miller, and Qiong Zhang. "An exploration of China's mortality decline under Mao: A provincial analysis, 1950–80." *Population Studies* 69, no. 1 (2015): 39-56.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey. “How Well Do Automated Linking Methods Perform in Historical Samples? Evidence from New Ground Truth.” *Journal of Economic Literature* 58, no. 4 (2020): 997-1044.

Baten, Joerg, Debin Ma, Stephen Morgan, and Qing Wang. "Evolution of living standards and human capital in China in the 18–20th centuries: Evidences from real wages, age-heaping, and anthropometrics." *Explorations in Economic History* 47, no. 3 (2010): 347-359. References

Beattie, Hilary. *Land and Lineage in China: A Study of T’ung-ch’eng County, Anhwei, in the Ming and Ch’ing Dynasties*. New York: Cambridge University Press, 1979.

Black, Sandra E., Neil Duzett, Adriana Lleras-Muney, Nolan G. Pope, and Joseph Price (2024), “Intergenerational Transmission of Lifespan in the US”, NBER wp 31034.

Brandt, Loren, Debin Ma, and Thomas G. Rawski. "From divergence to convergence: reevaluating the history behind China's economic boom." *Journal of Economic Literature* 52, no. 1 (2014): 45-123.

Broadberry, Stephen, Hanhui Guan, and David Daokui Li. "China, Europe, and the great divergence: a study in historical national accounting, 980–1850." *The Journal of Economic History* 78, no. 4 (2018): 955-1000.

Chang, Chung-li. *The Chinese Gentry*. Seattle: University of Washington Press, 1955.

Chang, Chung-li. *The Income of the Chinese Gentry.* Seattle: University of Washington Press, 1962.

Chang, Harrison, Timothy J. Halliday, Ming-Jen Lin, and Bhashkar Mazumder. "Estimating intergenerational health transmission in Taiwan with administrative health records." *Journal of Public Economics* 238 (2024): 105194.

Chen, Yuyu, Suresh Naidu, Tinghua Yu, and Noam Yuchtman. "Intergenerational mobility and institutional change in 20th century China." *Explorations in Economic History* 58 (2015): 44-73.

Chen, Zhiwu, and Kaixiang Peng, “Production, Consumption, and Living Standards,” Chapter 18, *The Cambridge Economic History of China, Volume 1: To 1800* Edited by Debin Ma and Richard von Glahn. Cambridge: Cambridge University Press, 2022, pp. 676-709.

Cutler, David, Adriana Lleras-Muney, and Tom Vogl, "Socioeconomic status and health: dimensions and mechanisms." Chapter 7 in *Oxford Handbook of Health Economics* 2011, pp124-163.

Dahl, Molly and Thomas DeLeire. “The association between children’s earnings and father’s lifetime earnings: estimates using administrative data.” IRP Discussion Paper 1324-08, 2008.

Deaton, Angus (2016), “On Death and Money: History, Facts, and Explanations”, Journal of American Medical Association, April 26, 2016.

Faure, David, and Xi He. “Family and Lineage in Late Imperial China”, Oxford Research Encyclopedias, Asian History, 2020. https://doi.org/10.1093/acrefore/9780190277727.013.394 (Accessed November 1, 2023).

Halliday, Timothy. "Intergenerational health mobility." In *Handbook of Labor, Human Resources and Population Economics*, pp. 1-22. Cham: Springer International Publishing, 2023.

Hollingsworth, T. (1957), “A demographic study of the British ducal families”, Population Studies 11(1), 4-26.

Ho, Ping-ti. *The Ladder of Success in Imperial China: Aspects of Social Mobility (1368-1911)*. New York: Columbia University Press, 1962.

HMD. Human Mortality Database. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at www.mortality.org (data downloaded on 7 June 2025).

Keller, Wolfgang, and Carol H. Shiue, Intergenerational Mobility of Daughters and Marital Sorting: New Evidence from Imperial China, NBER Working Paper No. 31695, September 2023.

Kendall, Ellen J., Alex T. Brown, Tim Doran, Rebecca Gowland, and Richard Cookson. "Health inequality in Britain before 1750." *SSM-Population Health* 16 (2021): 100957.

Lavely, William. Review of *Fate and Fortune in Rural China: Social Organization and Population Behavior in Liaoning, 1774-1873*, by James Z. Lee and Cameron D. Campbell, Cambridge University Press, 1997, in *American Journal of Sociology* Vol. 104, no. 1 (1998): 256-258.

Lee, James, Wang Feng, and Cameron Campbell. "Infant and child mortality among the Qing nobility: Implications for two types of positive check." *Population Studies* 48, no. 3 (1994): 395-411.

Lee, James Z., and Wang Feng. *One quarter of humanity: Malthusian mythology and Chinese realities, 1700–2000*. Harvard University Press, 2001.

Liu Ts’ui-jung. “Chinese genealogies as a source for the study of historical demography”, In *Studies and essays in commemoration of the golden jubilee of Academia Sinica*. Taipei: Academia Sinica, 1978.

Maddison, Angus, *Chinese Economic Performance in the Long Run, 960–2030 AD*, 2nd ed., revised and updated (Paris, OECD Publishing, 2007).

Marmot, Michael G., Geoffrey Rose, Martin Shipley, and Peter J. Hamilton. "Employment grade and coronary heart disease in British civil servants." *Journal of Epidemiology & Community Health* 32, no. 4 (1978): 244-249.

Maystadt, Jean-Francois, and Giuseppe Migali (2021), "The transmission of health across 7 generations in China, 1789-1906", Journal of Health Economics.

Miller, Grant. "Women's suffrage, political responsiveness, and child survival in American history." *Quarterly Journal of Economics* 123, no. 3 (2008): 1287-1327.

Miller, Grant, and B. Piedad Urdinola. "Cyclicality, mortality, and the value of time: The case of coffee price fluctuations and child survival in Colombia." *Journal of Political Economy* 118, no. 1 (2010): 113-155.

Perkins, Dwight H. *Agricultural Development in China, 1368-1968*. Chicago, Aldine publishing, 1969.

Pomeranz, Kenneth. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton University Press, 2000.

Preston, Samuel H. "The changing relation between mortality and level of economic development." *Population Studies* 29, no. 2 (1975): 231-248.

Ruhm, Christopher J. "Are recessions good for your health?" *Quarterly Journal of Economics* 115, no. 2 (2000): 617-650.

Shiue, Carol H. "Local granaries and central government disaster relief: moral hazard and intergovernmental finance in eighteenth-and nineteenth-century China." *The Journal of Economic History* 64, no. 1 (2004): 100-124.

Shiue, Carol H. "The political economy of famine relief in China, 1740-1820." *Journal of Interdisciplinary History* 36, no. 1 (2005): 33-55.

Shiue, Carol H. "A culture of kinship: Chinese genealogies as a source for research in demographic economics." *Journal of Demographic Economics* 82, no. 4 (2016): 459-482.

Shiue, Carol H. "Human capital and fertility in Chinese clans before modern growth." *Journal of Economic Growth* 22 (2017): 351-396.

Shiue, Carol H. (2025), "Social Mobility in the Long Run. An Analysis of Tongcheng, China, 1300 to 1900", *Journal of Economic History*.

Shiue, Carol H., and Wolfgang Keller. "Markets in China and Europe on the Eve of the Industrial Revolution." *American Economic Review* 97, no. 4 (2007): 1189-1216.

Shiue, Carol H., and Wolfgang Keller (2024), "Elite Capital Strategies for Big Shocks: The Case of the Fall of the Ming", NBER wp.

Telford, T. A. (1986). A survey of social demographic data in Chinese genealogies. Late Imperial China, 7, 118–48.

Telford, T. A. (1990). Patching the holes in Chinese genealogies. Late Imperial China, 11(2), 116–35.

Telford, Ted A. “Covariates of Men’s Age at First Marriage: The Historical Demography of Chinese Lineages”, *Population Studies*, 46, no. 1 (1992): 19-35.

Telford, T. A. (1995). Fertility and population growth in the lineages of Tongcheng county. In S. Harrell (Ed.), *Chinese historical microdemography* (Berkeley: University of California Press), pp. 1520–1661.

Wang, Feng, Cameron D. Campbell, and James Lee. "Agency, hierarchies, and reproduction in northeastern China, 1789 to 1840." In *Prudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900*, p. 287. 2010.

Weir, David R. "Family income, mortality, and fertility on the eve of the demographic transition: A case study of Rosny-Sous-Bois." *The Journal of Economic History* 55, no. 1 (1995): 1-26.

Appendix

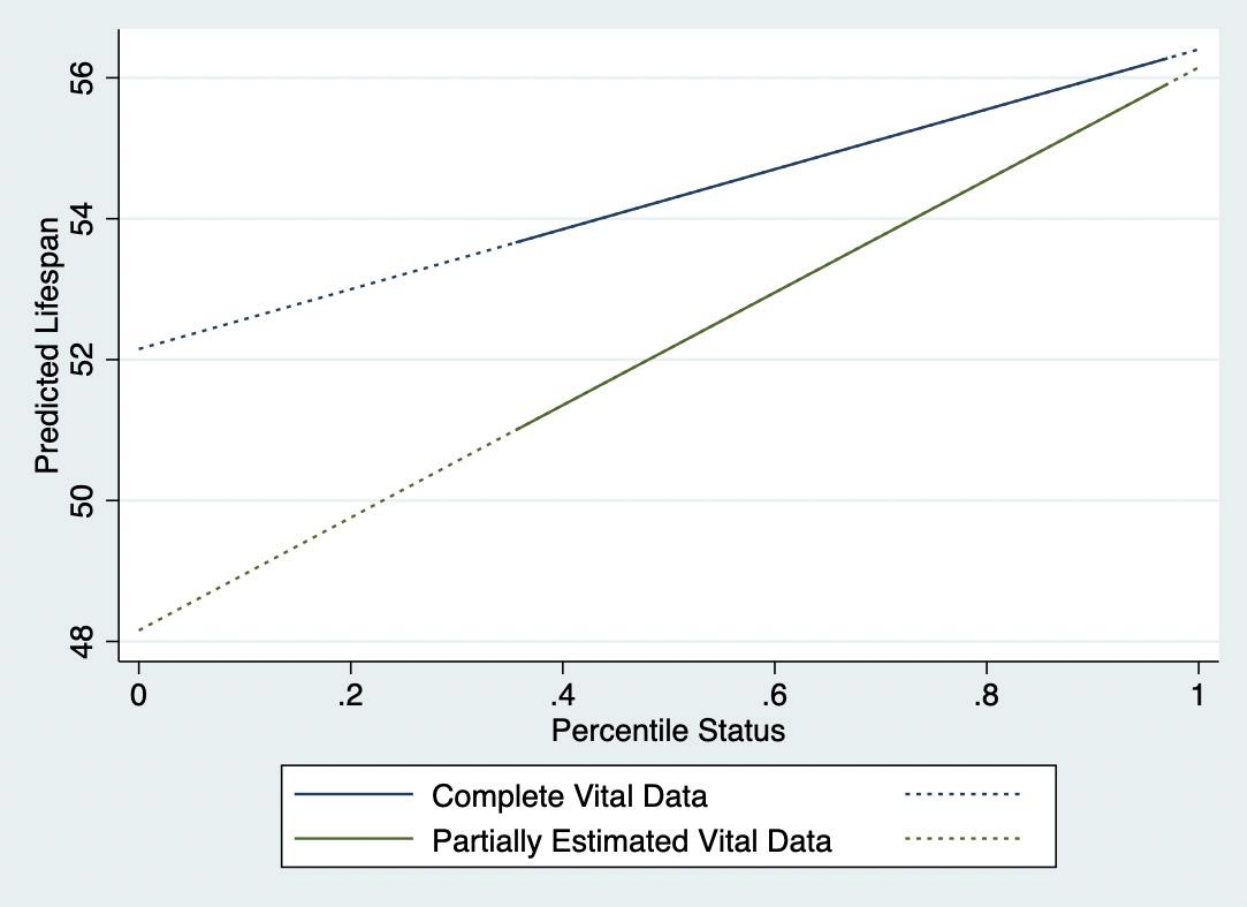
Table A1. Levels of Social Status and Descriptors in the Seven Genealogies

(1)	(2)	(3)
Status	Status Aggregated	Description
0	1	No title, degree, and evidence of status or wealth
1	2	Honorary or posthumous title; village head; other honors
2	2	Multiple wives in consecutive marriage (two or more not living at the same time)
3	2	Evidence of moderate wealth of 1st degree family, incl. minor and expectant official, lower-level degree *(shengyuan*, *jiansheng*), and official student
4	3	Wealthy family member 2nd degree, incl. official, *juren*, *gongsheng*, and *jinshi*
5	3	Wealthy family member 1st degree, incl. official, *juren*, *gongsheng*, and *jinshi*
6	3	Educated, scholar, no degrees or office; editor; refused office, or prepared but did not pass exam
7	3	Concubinage (two or more wives or concubines at the same time)
8	3	Substantial evidence of wealth and property; set up ancestral estates, large donations, philanthropy; wealthy farmer, landowner, or merchant
9	3	Official student
10	3	Military *shengyuan*, minor military office
11	4	Purchased *jiansheng* and/or purchased office
12	4	Student of the Imperial Academy (non-purchased)
13	4	Civil *shengyuan*; minor civil office
14	4	Expectant official; no degrees
15	4	Expectant official one of the lower degrees
16	4	Military *juren*, *jinshi*; major military office
17	4	Civil official with no degree, minor degree, or purchased degree
18	4	*juren*, *gongsheng*, with no office
19	4	*juren*, *gongsheng*; with expectant office
20	4	*jinshi*, no office
21	4	*jinshi* with official provincial post or expectant official
22	4	*jinshi* with top-level position in Imperial bureaucracy (Hanlin Academy, Grand Secretariat, Five Boards, Prime Minister)

Notes: Social status levels and descriptors in twenty-three levels with descriptors (columns (1) and (3), respectively. Column (2) gives status levels distinguished in this paper. Status ranking draws on Telford (1986, 1992), Chang (1955, 1962), and Ho (1962).

Important markers of status are those men who passed the civil service exam at the local level (*shengyuan*), at the provincial level (*juren*), and at the national level (*jinshi*).

Figure A1. The Role of Estimation of Vital Statistics



Notes: Figure shows linear Preston curves for husband birth cohorts 1740 to 1760 using Complete Vital Data (n = 502) and Partially Estimated Vital Data (n = 834).

We see that the Preston curve that is based not only on observations with complete vital data but that also includes observations for which vital data is estimated is steeper. In particular, the predicted lifespan premium of the top status percentile over the bottom status percentile (100th versus 1st) is about 8 years, whereas it is less than 4.5 years when one employs only observations with complete vital data.

There is a positive correlation of low lifespan and missing vital data, and furthermore, there is a status gradient in the missing data in that data is more likely missing for low status men than for high status men. As a result, using complete vital data misses part of the lower-left quadrant of Figure A1, compared to what is obtained when one includes partially estimated data. See Telford (1992) for more discussion.