NBER WORKING PAPER SERIES

NEW EVIDENCE ON THE HEALTH-STATUS GRADIENT FROM IMPERIAL CHINA

Wolfgang Keller
Carol H. Shiue
Karen Eggleston

Working Paper 35396
http://www.nber.org/papers/w35396

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2026

We thank Julie Shi, Joe Price, Yuli Xu, and audiences at the International Health Economics Association 2025 and American Economics Association 2026 conferences for comments and reactions that have improved this paper. Ted Telford initiated the data collection underlying this work. Support from National Institutes of Health (NICHD) through Award HD 042731-01 is gratefully acknowledged (Shiue). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Wolfgang Keller, Carol H. Shiue, and Karen Eggleston. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

New Evidence on the Health-Status Gradient from Imperial China
Wolfgang Keller, Carol H. Shiue, and Karen Eggleston
NBER Working Paper No. 35396
July 2026
JEL No. I14, J1, N35

ABSTRACT

We study the relationship between health and social status in China using genealogical data from the 14th to 19th centuries. The data covers broad segments of the population linking husbands, wives, and sons, allowing us to examine whether status was associated with survival not only for adult men but also across households and generations. In a sample centered on the early 1700s, the typical lifespan of married men is 52, but varies widely. High-status men have a 22 percent higher probability of living past 50 than low-status men, resulting in a lifespan advantage of seven years. High status is also associated with significant survival advantages for wives and children. The lifespan premium of the wife of a high-status man over the wife of a commoner exceeds 10 years, and the chance that the son of a high-status man dies before reaching 10 years is less than half that of a low-status man's son. Up to the 17th century the lifespan advantage is largely confined to those with top status, while the health-status gradient across all status levels emerges during the 18th century. The findings provide new evidence on health-status inequality in a less-studied setting, and suggest that health-status inequality is usefully analyzed from a family and intergenerational perspective.

Wolfgang Keller
University of Colorado Boulder
Department of Economics
and NBER
Wolfgang.Keller@colorado.edu

Karen Eggleston
Stanford University
and NBER
karene@stanford.edu

Carol H. Shiue
University of Colorado Boulder
Department of Economics
and NBER
carol.shiue@colorado.edu

1 Introduction

A large literature analyzes the relationship between health and dimensions of socio-economic status such as income and education in contemporary economies (see Cutler, Deaton, and Lleras-Muney 2006, Cutler, Lleras-Muney, and Vogl 2011). While a negative correlation of premature mortality with status is virtually ubiquitous – it has been found wherever and whenever it has been sought (Deaton 2016) – pre-20th century evidence for non-western settings remains limited. Precisely because the historical emergence of the health-status gradient is obscure, fully understanding it has been a challenge. This paper seeks to make progress by examining health and status in the 14th to 18th centuries using data on married men and women and their children from Chinese family histories.

Genealogical data from China provides several advantages for documenting the evolution of the health-status gradient. First, we record individual social status systematically based on more than 30 standardized descriptors, chief among which are educational attainment and official position in China's civil service examination system (*keju*). In contrast, widely-cited cohort data used to analyze historical mortality elsewhere, such as the Human Mortality Database, does not include status information. Historical studies that do include status mostly use occupation to measure status, and top status in Europe was largely hereditary. In contrast, the Chinese civil service exam is a distinctive setting where achievements directly convey social status, and thus-achieved status is in principle and in practice open to a wide swath of the population. Second, a central principle of Chinese genealogies is their inclusivity – all members of the clan, irrespective of status, are included. This principle is key to representativeness; more than 70 percent of our sample of men belong to the lowest social class. Other studies of health and status during the sample period, in contrast, tend to rely upon 'convenience samples', such as members of Europe's aristocracy, and they are not representative of society (Black et al. 2026).[1]

Our analysis reveals a status gradient in adult male lifespan of sizable magnitude. The positive health-status gradient is about seven years difference in adult lifespan between the highest- and the lowest-status men. Equivalently, the chance of survival to 50 conditional on surviving to 15 (i.e., S(50|15)) was 72% for top-status graduates of the civil service exam (*jinshi*) and 58% for commoners. That range includes the S(50|15) of 63% for Swedish men born around 1750 from the Human Mortality Database (HMD), though their status is unknown.

Generally, the health-status relationship is steeper for men of relatively high social status.

[1] In the literature, Chinese clans are also referred to as lineages or common descent groups. We use these terms interchangeably.

In fact, we find that men with lifespans ranging from 25 to 60 years typically do not vary much in terms of their status in the 16th and 17th centuries. This pattern starts to change in the 18th century. Consistent with the hypothesis that the positive health-status gradient emerges from the top of the distribution, we find that starting with post-1700 birth cohorts, lifespan is longer for intermediate status households than for commoners. While before 1700, lifespan advantages were the exclusive domain of top status families, by the 18th century a broader gradient emerged such that health and status were positively correlated for men across all classes of society.

We extend our health-status gradient analysis to two other family members: wives and children. Given family investments to find a suitable marriage partner, the relationship between wife lifespan and the status of her husband reflects not only the health-status gradient but also the extent of marriage matching.[2] We show, first, that the gradient between wife health (as proxied by lifespan) and the status of her husband is not only positive but it is stronger than the relationship between husband status and his own health. In particular, while highest-status husbands can expect to live seven years longer than low-status husbands, their wives typically differ in lifespan by eleven years. This reflects in part higher maternal mortality among commoners. The average mortality rate between age 20 and 40 of the highest-status married women was half that of the lowest-status married women.

The gradient between father (husband) status and child health reflects intergenerational factors including family endowments and norms. Because mothers provide some of these, marital sorting is also reflected in the gradient. We find strong evidence for a gradient between household status and child health. Mortality of a son at any age – e.g., 5, 10, or 15 years – born to a top-status father is only half of what it is for a son of a father ranked just ten percentiles lower, and it is almost one third of the mortality of a son born to a low-status father. Our results on lifespan of wives and sons shed light on the role of the health-status gradient for the formation of kinship networks as well as intergenerational mobility.

Our analyses contribute important historical perspective for understanding of the strength of the health-status gradient. A frequent starting point for studying the gradient in different populations and time periods is the cross-country relationship between health and per capita income and its shifting up over time, as 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

[2] A measure of a female's status based on her professional achievements or labor earnings is not available; this is typical for this historical period.

the well-being component of the 'great divergence' (Pomeranz 2001) between China and the west.

We also contribute to the historical and economic literature by analyzing data with a relatively time-invariant, multiple-category metric of socio-economic status (SES) – so that we are documenting a gradient, not merely a gap between a binary stratification of rich versus poor, or elite versus commoner. We provide evidence from information on lifespan of both husbands and wives that in our setting the step from the binary stratification of rich versus poor to the health-status gradient across more classes in society was taken in the 18th century. 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 – higher educational attainment leading to stable employment, assortative matching with wealthier marriages for himself and his sons, and bureaucratic or social power over consequential community resources. The significant gradients in lifespan for married men are supplemented with evidence on differential wife lifespan and child mortality by household status. Moreover, Qing dynasty China is an interesting period to study because of its dramatic population growth and trendless or declining living standards (Broadberry et al. 2018) ; by contrast, many previous studies focus on later periods, and especially cohorts living through improving living standards and overall rising life expectancy in the 20th century.

Relative to earlier studies of the health-status gradient in China, which for data availability reasons tend to focus on the elite and special settings (e.g. Lee et al. 1994, Maystadt and Migali 2021, Chen and Preston 2022, Bai et al. 2023), our sample includes commoners to a greater extent and thus captures a fuller span of the status distribution in society. Social mobility in Qing China starting in the late 18th century has recently been studied by Butaeva, Durlauf, and Shapoval (2026). Outside of China, Fogel (1986) has addressed questions of long-run economic status and health using genealogical data in North America, and further progress has been made recently with crowdsourced genealogies (e.g., Minardi, Corti, and Barban 2024, Black, Duzett, Lleras-Muney, Pope, and Price 2026). More generally, data from Chinese family genealogies have recently been employed to study issues such as Malthusian population responses, social mobility, and the child quantity-quality trade-off over the long-run (Hu 2023, 2025, Shiue 2017, 2025).

Due to the limits of data during the sample period, generality cannot be established in

the way one can for the more recent past (or today). The combination of the principle of including all clan members irrespective of status together with sampling poorer clans more yields a distribution of status in our sample – specifically, a small share of high-status individuals – that is similar to what is known to have been true across China as a whole (section 3). However, without a representative benchmark to compare to, sample selection cannot be ruled out. At the same time, the analysis of a small set of families for an early setting might prove useful because the evolution and reproduction of the Preston Curve is observed over multiple generations.

We also note that vital statistics are incompletely recorded in our data. It is common for vital statistics for this period to be incomplete. For example, the widely-respected Human Mortality Database compiled by expert demographers does not include England and Wales until the 1840s due to incomplete vital statistics, and warns users of the earliest data – Sweden starting in the 1750s – to be cautious in interpreting life tables before 1800 since they do not include unregistered individuals. In the present case, data reconstitution is performed by estimating missing vital information from auxiliary information in the genealogies or from model life tables, and our baseline includes observations with estimated vitals.[3] Compared to complete vitals, the estimation of vitals reduces variation and omits the demographic impact of any region- or period-specific violence, famine, and epidemics. However, because vital statistics are disproportionately missing for low-status individuals subject to high mortality, relying on unreconstituted data leads to sample bias (see Telford 1990). Section B in the appendix presents results from unreconstituted data; they also indicate a positive health-status gradient albeit its magnitude is smaller, which is related to the underrepresentation of lower-status men.

The paper is organized as follows. The next section describes the historical setting, followed by a description of data characteristics and representativeness. Section 4 presents our main empirical findings on the health-status gradient for men, while the following section extends them to married women and sons. The final section discusses our results. More information on the data and additional estimation results can be found in the Appendix.

[3]Model lifetables (e.g., Coale and Demeny, 1966, 1989; United Nations 1982) are compilations of average age of death regardless of status that have been "designed primarily for use in developing countries or for estimating historic populations, [and] are limited to mortality patterns for a life span from age 20 to 75," with the Human Mortality Database (HMD) listed as the standard reference (https://www.un.org/development/desa/pd/data/model-life-tables).

2 Historical Setting

2.1 Economy and governance

While this study covers (parts of) three imperial dynasties – the Yuan (1271-1368), Ming (1368-1644), and Qing (1644-1911) – the sample is centered on cohorts in the early 1700s, the relatively early Qing dynasty. 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 comparison with other states of the period. Moreover, the scope of markets for allocating scarce resources was limited by technology (e.g., transport technology, financial instruments) more than by government regulation, in part because effective enforcement would have required more state resources than were in fact allocated. Additional perspectives are presented in Brandt et al. (2014) and Keller and Shiue (2022).

In China's 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.

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 (*keju*). Throughout the Ming-Qing period, sucess in the tournament-style civil service exams was the most direct path towards status and income. Passing the exam and obtaining a government office brought substantial material and social rewards, and as a result, the *keju*, rather than wealth alone, became the principal institutional basis of social status (Ho 1962).

Despite the sample period extending over several centuries, there was substantial continuity in how China was governed. 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.

2.2 Health and medicine

Our study contributes novel results on the health-status gradient in this historical setting by including both commoners and elites. 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 5,000 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"; their lifespans were relatively high (concentrated in the 60s and lower 70s) over the Ming and Qing dynasties.

Studies documenting a status gradient in imperial China are even more limited, and focused mostly on infant and child survival to adulthood. For example, analyzing genealogical data on the Qing dynasty nobility, Lee, Wang and Campbell (1994) find that infant and child mortality was similar in magnitude to that 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 per 1,000 British elite boys, and other estimates for continental Europe; Hollingsworth 1957, Lee et al. 1994). Because reductions in child mortality may have led to stable or even increasing life expectancy at birth, a gradient in child survival and any improvement over time would be consistent with the Chen and Peng (2022) argument that overall human capital in China was remarkably high for low per capita income as well as quite stable despite exploding population in the late imperial period. 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). Analyzing a sample of 2,859 fathers who had records of non-surviving sons–a subsample from six lineages from Jiangsu and Zhejiang spanning 1350 to 1920–Hu (2023) also finds social status to be associated with statistically significant higher probability of sons surviving to adulthood.

Longer lives among high-status individuals may seem obvious and omnipresent, but that has not always been the case. Long spans of history before the modern era in other parts of the world appear to lack a significant survival advantage of the elite. For example, DeWitte et al. (2015) study London in 1700-1853, concluding that there were no evident gradients in adult health and survival, although lower SES households did suffer from higher infant and child mortality. Their assessment of little gradient in adult lifespan in London is consistent with numerous studies of disparate parts of industrializing Europe. Moreover, several mechanisms would tend to dampen the health-status gradient in China. For example, to the extent that long-lived fathers had more children, the early death of a high-status father might ensure a greater inheritance to each surviving son.[4]

[4]For example, Beattie (1979) writes of one famous Tongcheng lineage: "It was fortunate from this point of view that the affluent Chang P'eng had only one son...who inherited all his property undivided. [Then the son died] at the age of 36 when he had had time to father only two sons, ensured that it would be passed on

The comparative advantage of our data is the inclusion of a robust sample of non-elite adults with relatively consistent gradations of status across a long historical period. These features enable us to document the health-status gradient beyond childhood for a broad swath of the population. While evidence on cause of death among adults during our study period 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). Many health-promoting behaviors are relatively inexpensive but time-consuming and may have been considered menial, such as fetching clean water and consistently boiling it before drinking. For example, although breastfeeding infants is protective of both infant and mother -- and poor households had few options but to carry on this tradition -- married women of ducal families of pre-1750 England employed wet nurses (Kendall et al. 2021): these elite English women had high fertility and closer child spacing, leading to lower immunity of infants and more susceptibility to early death from infectious disease.[5]

However, there are many reasons to expect that a health gradient might become increasingly evident in China over the study period. Potential mechanisms underpinning differential survival by status and diffusion of health improvement from the elites to commoners include improved access to nutrition, shelter, and clean water, as well as literacy and effective medical practices such as variolation against smallpox. Certainly better nutrition and clean water, and the means to boil water, would have promoted better maternal, child, and adult health. Less strenuous caloric expenditures (e.g. from sedentary lifestyles rather than manual labor) would have reinforced the nutritional benefits of higher status, and reinforced their ability to survive bouts of infectious disease. Elite's investment in education in turn could contribute to diffusion of the these benefits by, for example, increasing the supply of local teachers, reducing the costs of schooling to attain basic literacy, and lifting overall literacy rates (Elman 2002; Shiue 2017), which no doubt contributed to improving human capital in both the health and educational dimensions. Female literacy in particular–initially confined only to high-status women in China–has been consistently shown to improve child health (Wolpin 1997; Schultz 2002; Miller 2008; Gakidou et al. 2010; Baker and Stabile 2011; Cutler et al. 2011).

for one more generation with a minimum of division" (Beattie 1979, p.89).

[5]More generally, the time price of health investments may lead to better health for those with lower opportunity costs of time, such as the poor. Ruhm (2000) shows that in recent US history, physical activity increases and diets improve during recessions, with less smoking and reduced obesity during periods of high unemployment with presumably lower opportunity cost of time. In an agricultural setting, Miller and Urdinola (2010) document higher child mortality 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.

Elites had incentives to help mitigate spread of infectious disease from epidemics, natural disasters, or famine, reinforced by the social role of clans. Historical records indicate high-status households contributed charitable granaries as well as land to support educational expenses of poor young men, helping to diffuse literacy even to some girls (Beattie 1979, p.50). Rawski (1979) estimates literacy rates by the late Qing dynasty of 30-45% for men and 2% of women. Clans also provided other forms of social protection to constituent households that presumably helped to break vicious cycles of illness-induced poverty.

Some additional 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, and could have promoted emergence of a health-status gradient. 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 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 Beijing and 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 of this period. 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 what extent might these factors have supported emergence and diffusion of a health-status gradient in late imperial China? The following section provides more information on the the source of the data employed to investigate this question: the lifespans of commoners and local elites as captured in genealogies.

3 Data from Genealogies

3.1 Chinese Genealogies as a Source of Information

This paper utilizes as its data source one of the main documents produced by clans: the genealogy. Genealogies allow for a more expanded definition of status compared to many other sources of socio-economic status (SES) proxies, because they include not only status information mirroring the ranking of exam degrees and official titles used by the state, but

they also record non-official types of status that clearly mattered, such as honors, education that did not translate in office, or wealth unrelated to the *keju*.

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, private genealogies had spread widely across the empire. As such, the Chinese genealogy is a well-known type of 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) compiles a list of approximately 50,000 publicly available genealogies, the bulk of which were produced in the Ming and Qing Dynasties.

A key reason why one might utilize family genealogies is that during the sample period they can be of much higher quality than official government records (this is the case here; Shiue 2025 presents some analysis).[6] At the same time, Chinese family genealogies follow certain principles that are important for interpreting any empirical results. For one, Chinese genealogies are organized patrilineally. The family tree follows the marriages of the males of the family, which generally implies less information on daughters than on sons, and less information on wives compared to husbands. In our genealogies, sons that survive to their first birth day are recorded consistently over the sample period (in contrast to daughters for whom norms changed); thus we examine information on sons when studying how child health relates to father status.

Another central principle of Chinese family genealogies is that for reasons of ritual, all males in the patrilineal line were included. This principle is key for our ability to create a sample that is roughly representative of China's society as a whole. While government data for the sample period in China (and elsewhere) tends to be more abundant and reliable for persons of higher wealth, for families it was important to include their less fortunate members in the family history, and as a result, the share of men of the lowest social status in our source is relatively high (above 70 percent). Another principle of Chinese genealogies worth keeping in mind in interpreting our results is that if a son does not marry, not only does his family line end but as a single adult male he will typically not be included in the clan's genealogy.

[6]The first modern census for China dates from 1953, while dates for the USA and France are 1793 and 1801, respectively.

The share of men who remained single during the sample period is estimated to be between 10 and 20 percent (Telford 1992).

Given these genealogical principles, Chinese family genealogies have generally found to be accurate in comparisons to population registry data once the latter is available (Harrell 1987). This confirms results for German and US family histories (Knodel and Shorter 1976 and Norton 1980, respectively). In the specific case of Tongcheng, systematic overstating of achievements is unlikely because the competition for social prestige in Tongcheng was high and any detected manipulation would have gravely undermined a family's reputation (Beattie 1979).

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 norms were not immutable, and they could adapt to new situations (on how family norms after the fall of the Ming dynasty changed, see Shiue and Keller 2026). However, 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, Chinese family genealogies include a considerable amount of information (including social status) that is rarely available in other sources for this sample period, and they are comparatively inclusive in covering relatively low-status individuals. 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.

3.2 Representativeness

The key representativeness concern is that genealogical data is disproportionately available for richer populations. Not every family had a genealogy, because the family needed to have the resources to compile one, at least some family members had to be literate, and survival chances of historical genealogies are also likely higher for richer families. Thus, employing all existing genealogies would yield an upward-biased sample. The extent of this can be quantified starting in the late 19th century. For example, remaining lifespan at age 30 in Familinx, a broad crowdsourced genealogical database, is about 7 years (23%) longer than in national life tables in Germany in the 1870s (Stelter and Gutierrez 2022). Thus, the sample can be selected in a major way.

To address this issue we select genealogies for a particular area of China from a relatively large set of available records so as to achieve a sample that is broadly representative of what we know about China during this period. Our sample is drawn from families that resided in Tongcheng County of Anhui Province, where genealogies exist for numerous clans (Beattie 1979 relies on more than 60 genealogies). 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. Tongcheng county is not the typical county in China. While neither the prefecture or province in which it was located (Anqing and Anhui, respectively) ranked high in producing top officials during Ming-Qing China, the county of Tongcheng gained some fame in the late Ming and early Qing for having produced a number of the highest officials of the empire (see Shiue and Keller 2026).

We select seven of the Tongcheng clans to be included in our sample, and to avoid over-sampling relatively high-status men we disproportionately sample genealogies from clans that were relatively less accomplished. The clan names are the Chen, Ma, Wang, Ye, Yin, Zhao, and Zhou (of Tongcheng). Representativeness of the sample can be assessed by comparing the share of high-status men in the sample with estimates of their share in China's population as a whole. Including all men from these seven clans in the analysis, the share of high-status men in the sample is somewhat below the share of high-status men according to estimates for the Ming and Qing dynasties as a whole, and somewhat above the share of high-status men for the early 19th century (Shiue and Keller 2026). This suggests that our sample is broadly representative of China during this period.

Given that sample representativeness is achieved by oversampling relatively low-status clans, one might ask what the role of within- versus between-clan differences is for our results. To address this concern we perform robustness analysis that discards between-clan

differences by including a fixed effect for each clan. It shows that the health-status gradient using within-clan variation is similar to that using also between variation (see Figure A.11). Of the more than 11,000 couples recorded in the data, the sample consists of the husbands born between 1298 and 1775. The majority of the sample lived during the late Ming dynasty (1368 to 1644) and the early-to-mid Qing dynasty. 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.

In general, Chinese family genealogies vary in their completeness, and often some vital statistics are missing.[7] The extent to which the recorded data is subject to selection bias, as well as how it can be addressed, is discussed in Telford (1990). The baseline estimates in the text use vital statistics that are estimated using auxiliary information in the genealogies, as well as model life tables, a standard demographic tool. Model lifetables generally used for such reconstitution (e.g. Coale-Demeny 1966; United Nations 1982) are compilations of average age of death regardless of status that have been "designed primarily for use in developing countries or for estimating historic populations, [and] are limited to mortality patterns for a life span from age 20 to 75," with the Human Mortality Database (HMD) listed as the standard reference as the model lifetables were updated (https://www.un.org/development/desa/pd/data/model-life-tables). Application of such model lifetables to fill in missing vital statistics obviously reduces variation and omits the demographic impact of any region- or period-specific violence, famine, and epidemics. In the appendix, we discuss how data reconstitution affects our findings by presenting several sets of results with samples that exclude estimated data.

3.3 Sample Data on Health and Status

The genealogies record for each adult male his highest lifetime achievements in biographical sections that resemble obituaries. Our genealogies use more than 40 standardized descriptors of achievements from which a man's social status is derived. The descriptors and the status ranking based on them is shown in Table A.1. Participation in China's civil service examination, and based on that attaining official position, was important for a man's status during the sample period (Ho 1962), and civil service exam achievements are virtually always recorded. Other factors affecting social status in Qing China include occupation, the size of the household, making investments, as well as other indicators of wealth and prestige.

Table 1 shows information on status and lifespan of the men in the sample. These are all

[7] See the survey by Telford (1986).

married men from seven Tongcheng lineages that are born by the year 1775 with a completed lifespan of 16 or more years.

Table 1: Status and Vital Statistics for Married Men from Tongcheng

Panel A: Social Status

Status Ranking and Descriptors	N	Sample Share	Status Percentile
Low: No Status	3,368	0.709	0.354
Intermediate: Village head, multiple consecutive marriages, relative of minor official	693	0.146	0.782
High: Wealthy farmer, landowner, or merchant; scholar; official student	363	0.076	0.893
Very High: Graduate of the civil service examination at local, provincial, and national level; high government official	329	0.069	0.965
Total	4,753	1.000	

Panel B: Lifespan

	N	Mean	Standard Deviation	Min	Max
Completed Lifespan	4,656	52.39	15.50	16	91
Year of Birth	4,753	1703.52	67.69	1298	1775
Year of Death	4,656	1755.66	69.82	1348	1856

Notes: Panel A: Percentile of status is mid-point of status group; e.g. with 70.9% of men in the lowest status class, the mid-point is $(0 + 0.709)/2 = 0.354$. Social status coding informed by Chang (1955, 1962), Ho (1962) and Telford (1986, 1992). Additional information on status classes and descriptors in Table A.1. Panel B: Sample is all married men in birth cohorts 1775 or before that reached age 15 or higher from seven Tongcheng clans.

The number of observations is N = 4,753. The analysis distinguishes four social status classes, namely Low, Intermediate, High, and Very High (see Table 1, Panel A). Men whose genealogical entry does not mention any aspects of achievement or wealth are coded as Low

status. These men should be considered as commoners, and they account for more than 70 percent of the sample. Men with intermediate status make up the next 15 percent of the sample. They include men who became village head, who had multiple marriages, or who were related to a minor government official. Examples of descriptors for a particular social class are given in Table 1, with more details given in the appendix (Table A.1). Wealthy merchants as well as official students of the civil service exam had high status, which accounted for almost eight percent of the sample, while graduates of the civil service exam at the provincial or national level (*juren* and *jinshi*, respectively) account for most of the remainder, which is men of very high social status. As is typical of data for the sample period, additional individual-level information that would allow us to rank men within each of these social classes is unavailable.

For robustness and to account for the size of each status class, we measure status in percentile rank following Chetty et al. (2024) and others. For example, given that the lowest status class covers 70.9 percent of the sample, each of these men is assigned the percentile mid-point of this class, which is 0.354 (Table 1, Panel A, last column). Given these status percentile midpoints, one can think of low, intermediate, high, and very high status men as roughly those at the 35th, 80th, 90th, and 97th percentile of the status distribution.

Information on vital statistics of these men is shown in Panel B of Table 1. Information on birth is somewhat more common than information on death, but there is information on social status for all men for whom we can calculate completed lifespan (year of death minus year of birth). Recall that a part of the vital statistics are reconstituted using auxiliary information in the genealogies, or they are estimated using lifetables. The sample of unreconstituted data underrepresents low-status men (Telford 1990).[8]

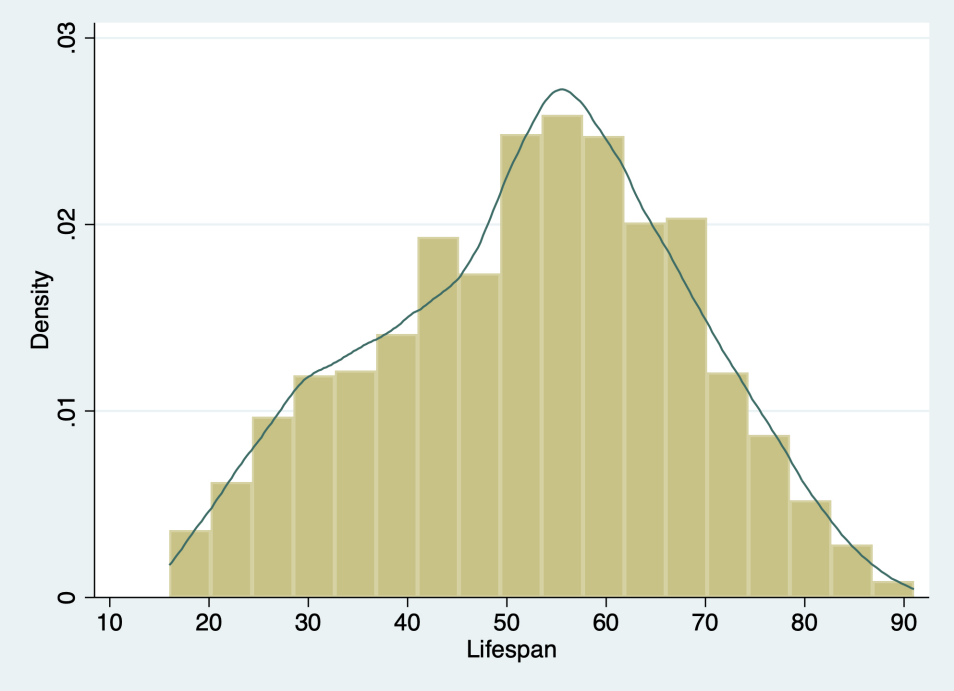
We restrict the sample to birth chorts that were born by the year 1775 in order to avoid the issue of censoring due to incomplete lifespan by the time the genealogy was updated for the final time (this varies by clan but can be in the 1860s). The earliest year a sample man was born is 1298, and the earliest observed death is in 1348. Relatively few men are born before 1500, and the average birth year of 1704 indicates that the sample sheds primarily light on the era of the Qing dynasty (1644 to 1911). The last observed death in the sample occurred in the year 1856.[9]

Figure 1 is a histogram of completed lifespan in the sample. The average lifespan in the sample is 52 years. This is higher than what we know about the life expectancy at birth in China during this time for two main reasons. First, the sample is conditional on reaching 16 or more years of age, that is, child mortality is excluded. Some information on child mortality

[8]See section B for results based on unreconstituted data.

[9]We provide some robustness analysis that assesses the role of the clans' progenitors for our results. See Figure A.10. Also, as shown in Fig. A.10, results are similar if we focus on birth cohorts >1600.

Figure 1: Lifespan of Married Men in Seven Tongcheng Clans



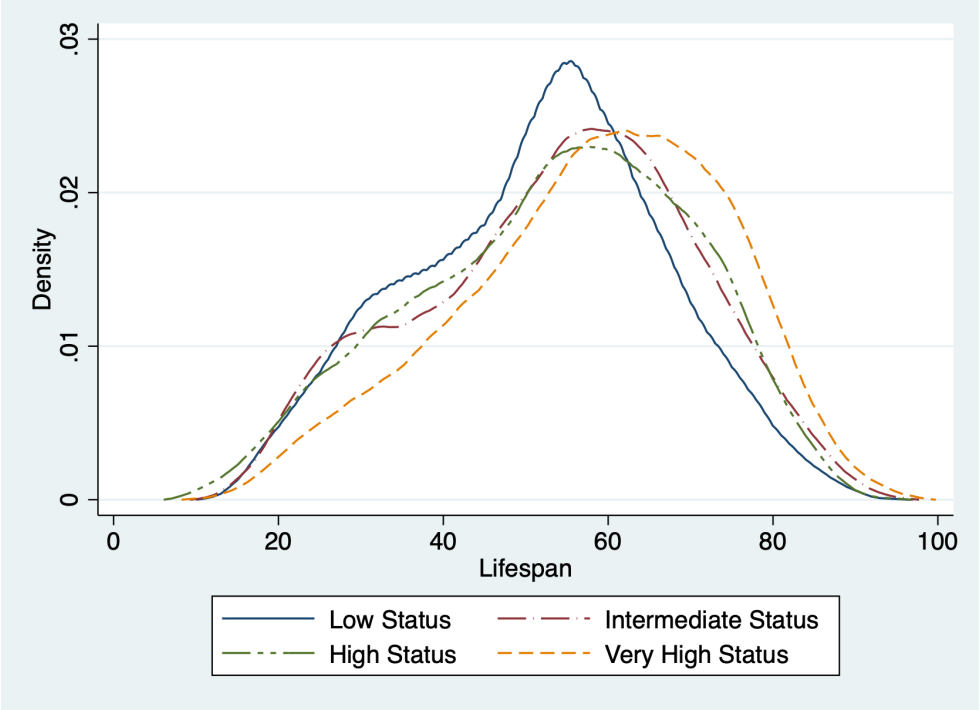
Notes: Completed lifespan of all married men from seven Tongcheng lineages with birth on or before 1775; n = 4,656.

will be presented below. Second, only married men are in the sample. This tends to raise average lifespan because lack of resources is both an important reason for not being able to marry and having a relatively short lifespan. Figure 1 indicates that there is wide variation in completed lifespan in the sample. The interquartile range (25th to 75th percentile) is 41 to 63 years, and the 90/10 range in lifespan is 72 to 30 years.

4 Health-Status Gradient for Married Men

This section describes the health-status gradient among married men. Figure 2 provides densities of lifespan by the four status levels shown in Table 1. We see that as status rises, the lifespan distribution shifts to the right. This indicates a positive health-status gradient. The difference is most clear between the lowest and highest status group. For men with very high status, there is a substantially fatter high tail in the lifespan distribution, while on the low lifespan side low-status men have a disproportionately high incidence of mortality in their 20s and 30s. Average lifespan for the four classes by increasing status is 51.4, 53.8, 53.4, and 58.2 years. Thus, while there may not be a clear lifespan difference between men at the 80th and 90th percentile (intermediate versus high status, respectively, the typical difference between commoner and elite lifespan is about 7 years.

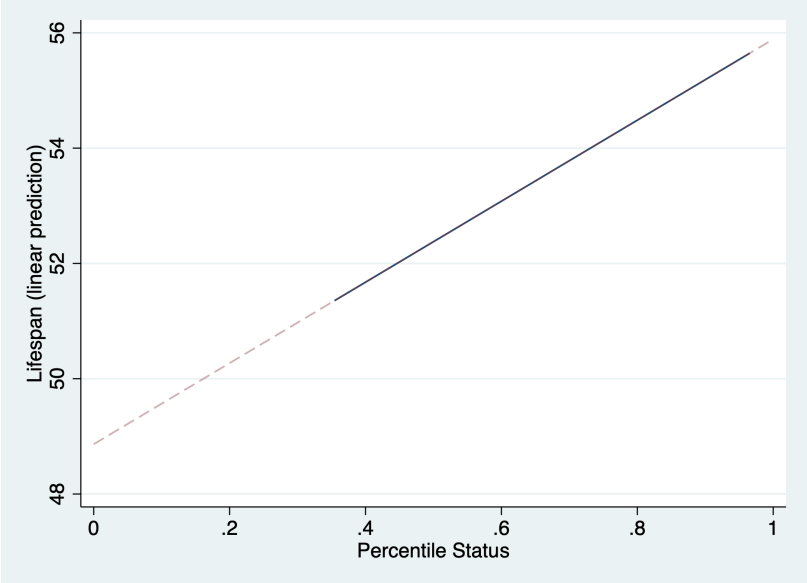
Figure 2: Lifespan Distribution by Status



Notes: Kernel estimates of completed lifespan by four social status levels given in Table 1.

If one imposes a linear health-status relationship, the result is seen in Figure 3.

Figure 3: Status - Male Health Relationship: Linear



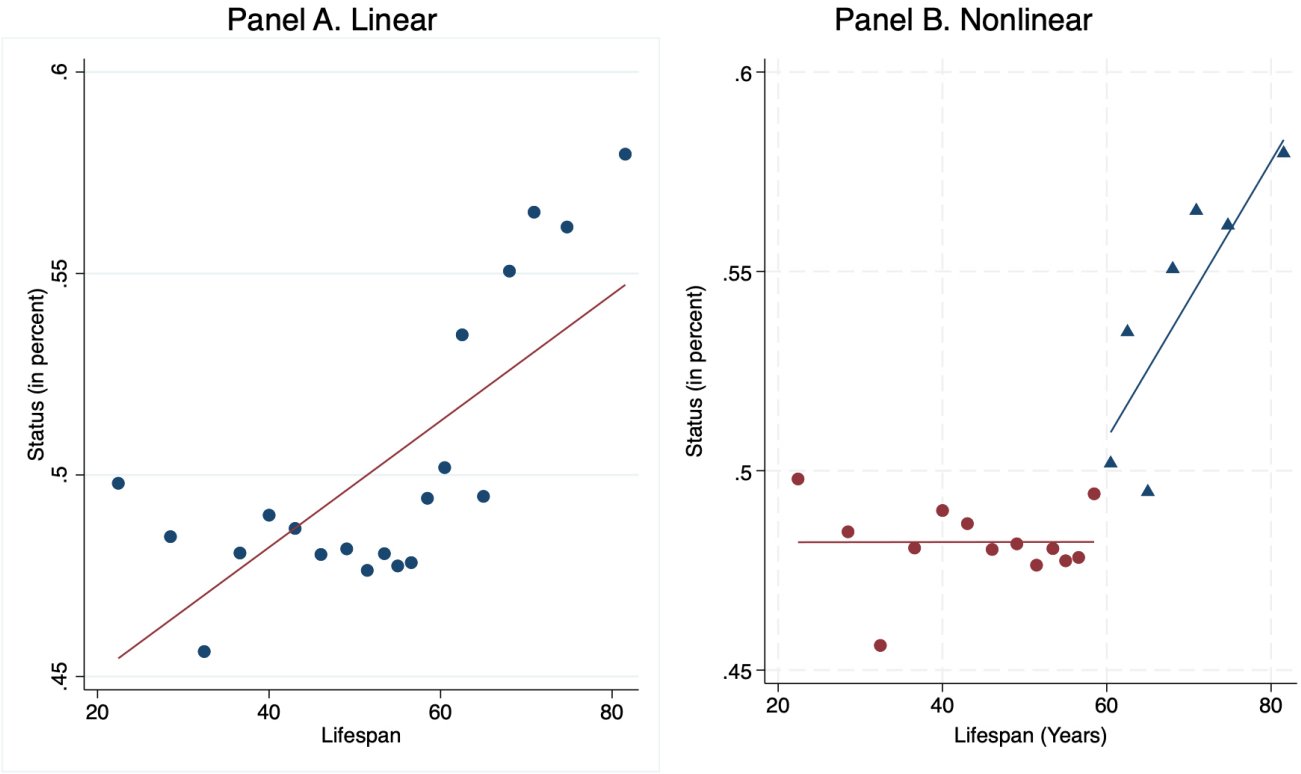
Notes: Linear relationship between percentile status and completed lifespan; n = 4,656.

We see the predicted change in lifespan for a man moving from lowest to highest status, 0

to 100th percentile, is about seven years. A seven-year difference in adult lifespan is equivalent to half a century of longevity improvement in industrializing England and Wales, according to demographic data from the Human Mortality Database (HMD): the difference between the birth cohorts of 1841-44 and 1890-94 in England and Wales was 7.07 years lifespan (cohort life expectancies conditional on survival to age 15).

A limitation with the gradient analysis so far is that the number of data points that pins down the gradient is with four status classes relatively small. An alternative to analyzing the health-status gradient is to ask whether men with different lifespans differ systematically in terms of their typical social status. This amounts to flipping the horizontal and vertical axes compared to the previous analysis.

Figure 4: Status-Health Relationship



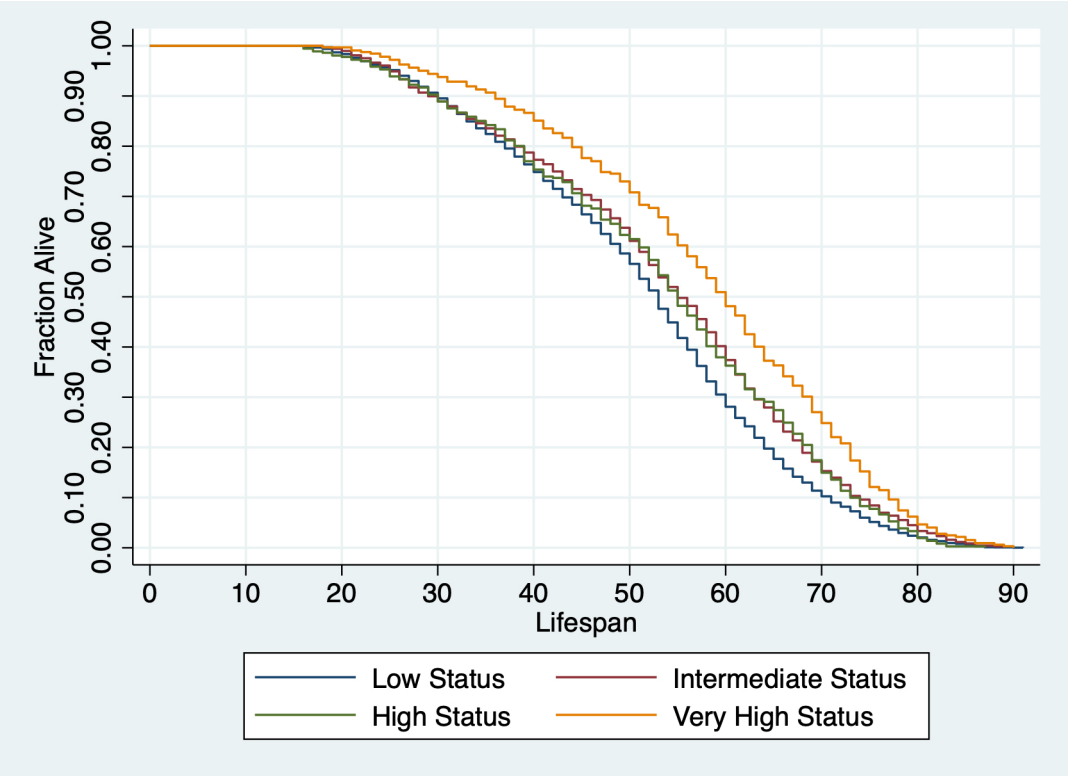
Notes: Scatter plot showing average percentile status by lifespan bin; Panel A imposes an overall linear relationship, while Panel B allows for different linear slopes below and above 60 years of lifespan; n = 4,656.

Panel A of Figure 4 indicates that as men have higher lifespan, their typical status level goes up as well. At the same time, there is evidence that the extent to this is not constant for all lifespan levels. In particular, if one allows for a change in slope at a lifespan of 60, the result is shown in Figure 4, Panel B. We see that while for lifespans above 60 years the

gradient between lifespan and status is positive and sizable, for lifespans between 20 and 60 years higher lifespan is not associated with systematically higher status. This finding parallels results of a steeper gradient at the top end of the status distribution in the US during the early 21st century (Chetty et al. 2016).

A large literature examines the health-status relationship in terms of survival to a particular age Figure 5 shows Kaplan-Meier survival estimates by social status.

Figure 5: Social Status and Male Survival

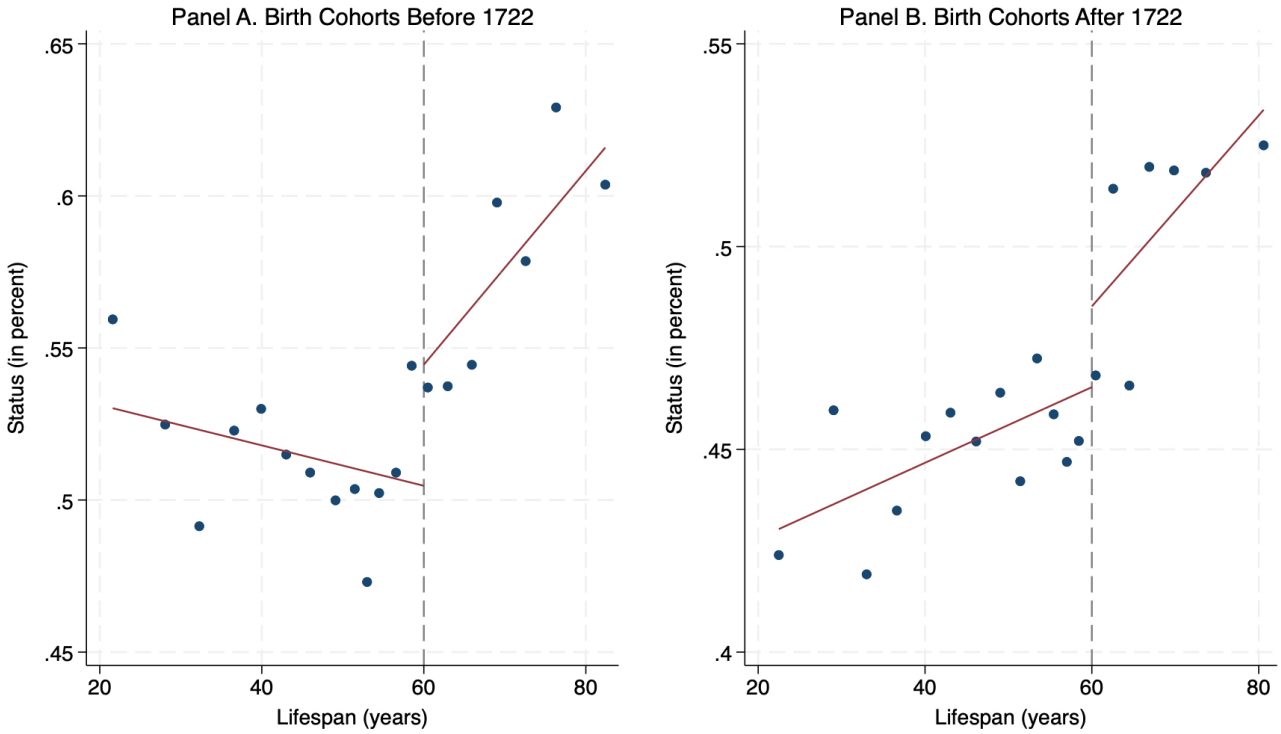


Notes: Kaplan-Meier survival estimates by status group; sample is all birth cohorts 1775 or before and men that reached 16 years of more of lifespan; n = 4,656.

There is a sizable lifespan gap between men of low versus very high status. For example, men of the highest status have a 72% chance of surviving to age 50, whereas only 58% of low-status men live longer than 50 years. Survival hazards for men at the 80th and 90th percentiles lie in between these extremes, as one would expect. This estimate is a plausible gradient that spans the 67.4% chance of survival to age 50 (conditional on surviving to age 15) for men born in England and Wales in 1841-1844, according to cohort lifetables compiled by the expert demographers of the Human Mortality Database.

An interesting question is whether the health-status relationship has been changing over time. To address this question, we compare birth cohorts before and after the year 1722, which splits our sample into two equal-sized groups (referred to as early and late). In these subsamples, average lifespan is similar at about 52 years, however, typical status levels in the late sample were lower than in the early sample, in large part because of the drying up of *keju* degrees relative to China's larger population (average status of 0.47 versus 0.53, respectively).[10] Figure 6 shows the status-health relationship separately for the early versus late birth cohorts of the sample. Given the increase of clan population in our data, this sheds light on a change in the status-health relationship over three quarter of a century (median birth year in the early sample is 1677, compared to 1752 in the late sample).

Figure 6: Status-Health Relationship over Time



Notes: Scatter plot showing average percentile status by lifespan bin in the early (Panel A) versus late (Panel B) part of the sample; N = 2,317 in Panel A, N = 2,338 in Panel B.

Figure 6 shows that the health-status gradient is present and largely unchanging over time for men of relatively strong health (and status). In contrast, for men with lower levels of health as measured by lifespans below 60, there is no positive health-status gradient in the early part of the sample, while later–in essence, seventy-five years later–, there is.

[10]Shiue (2017, 2025) studies the implications of this for fertility and intergenerational mobility.

This is consistent with Deaton's (2016) observation that the negative correlation of premature mortality with status is virtually ubiquitous — we find evidence for this even in the early part of the sample, centered on the 1670s, among the healthiest 1/3 of men. However, to the best of our knowledge, the evidence in Figure 6 is among the first that documents the emergence of a positive health-status gradient for the population at large in Asia.

5 Male Status and the Health of Wife and Children

The health and status of one member of the household might for numerous reasons affect health and status of other household members, particularly in the case of the head of household. Extending the analysis of how the man's status is related to other household members can thus be viewed as examining elements of a generalized health-status relationship. The following analysis begins by relating husband status to the lifespan of his wife. Higher-status men might make investments towards the health of their wives, or a healthy wife might support her status-seeking husband more than a less healthy wife. Furthermore, the relationship between husband status and wife health is also likely affected by the extent of marriage matching by the two families.[11] We then turn to the relation of man status with the health of his children, which will also provide some intergenerational perspectives.

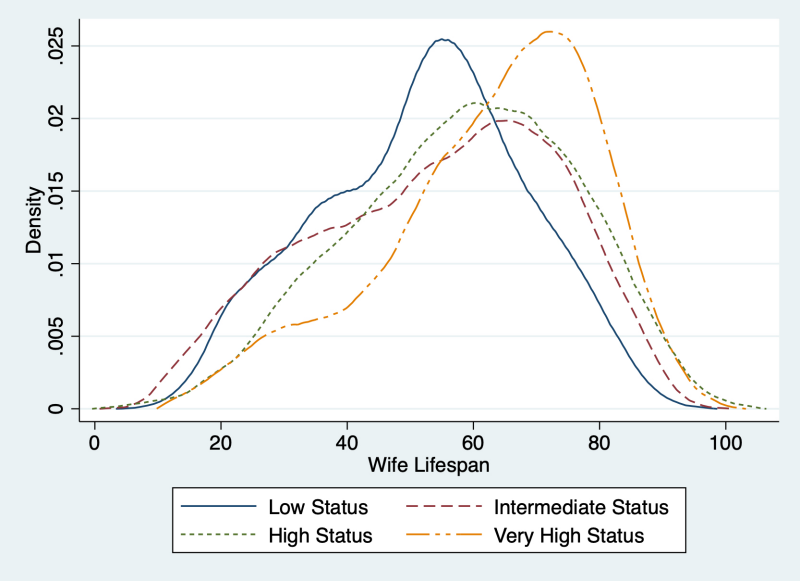
5.1 Male Status and Wife Health

Key information on the relationship between husband status and wife health is summarized in Table 2. We consider the same four groups of male status as before, from low to very high (for descriptors, see Table 1). The sample is smaller because we focus on women who are the single lifetime wife of their husbands. This avoids the mechanical correlation between high male status and short wife lifespan because only rich men have the resources to marry multiple times, and they do so typically because the first wife dies early (that is, low lifespan).[12] Compared to the analysis of male health, this sample restriction increases the share of commoners in the sample; it is now 83%, compared to 71% before.

[11] Results on marriage matching for Ming-Qing China include Keller and Shiue (2022).

[12] The average lifespan of all women who are spouses of men with more than one lifetime wife is 35 years, far below the sample average.

Figure 7: Wife Lifespan by Social Status of Husband



Notes: Distribution of wives lifespan by percentile status of their husbands; n = 3,706.

Table 2: Husband Status and Wife Lifespan

Husband Status Group	N	Status (in percent)	Wife Lifespan (average)
Low	2,870	0.393	51.95
Intermediate	389	0.836	54.14
High	241	0.917	58.51
Very High	205	0.974	62.72
Total	3,706		53.20

Notes: Sample is all single-spouse wives of men of birth cohorts before 1775 that lived 16 or more years. Status in percent is mid-point of status group in percentiles; social status coding informed by Chang (1955, 1962), Ho (1962) and Telford (1986, 1992). Additional information on status classes and descriptors in Table A.1.

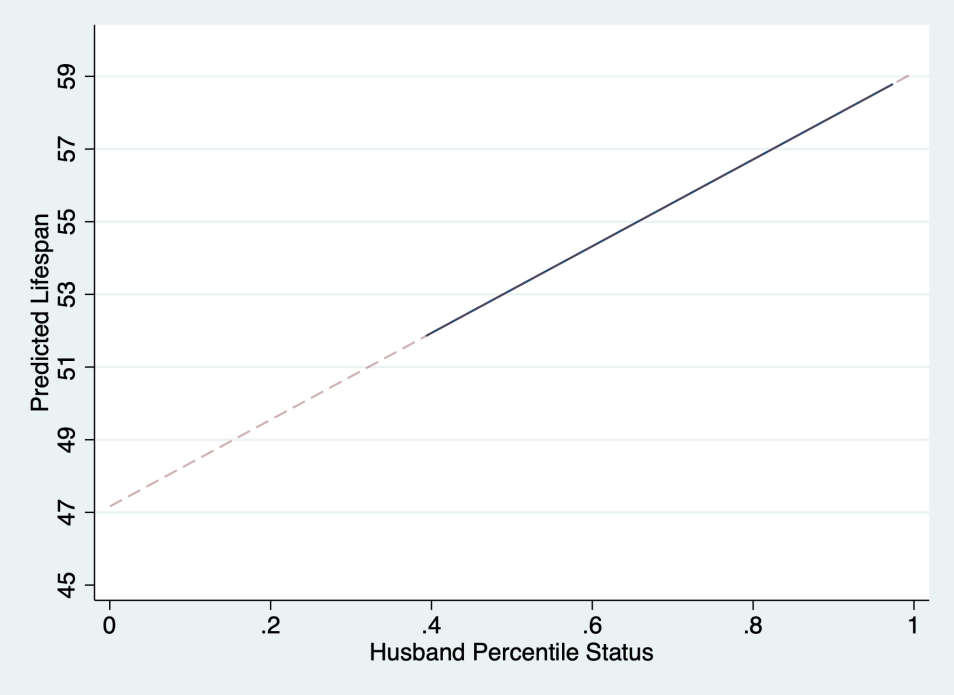
We see further evidence for a positive health-status gradient during the sample period in that lifespan of the wife is typically higher the higher the status of her husband (Table 2, last column). The lifespan difference of a wife of a low-status man versus of a wife of a very-high status man is around 11 years, which exceeds the lifespan difference for the husband.

Figure 7 shows the distribution of wife lifespan by the status of their husband. Figure 7 presents more evidence for a positive health-status gradient. Mortality in the child-bearing

years is clearly higher among lower-status women. In fact, at any given age between 20 and 40, high-status wives are twice as likely to survive as their commoner counterparts. This survival advantage reflects differential fertility and exposure to maternal mortality, as well as some protection against other health threats to which men were also exposed (e.g., infectious disease or injuries).

A strong female gradient is confirmed by Figure 8, which shows the linear relationship between husband status and wife lifespan.

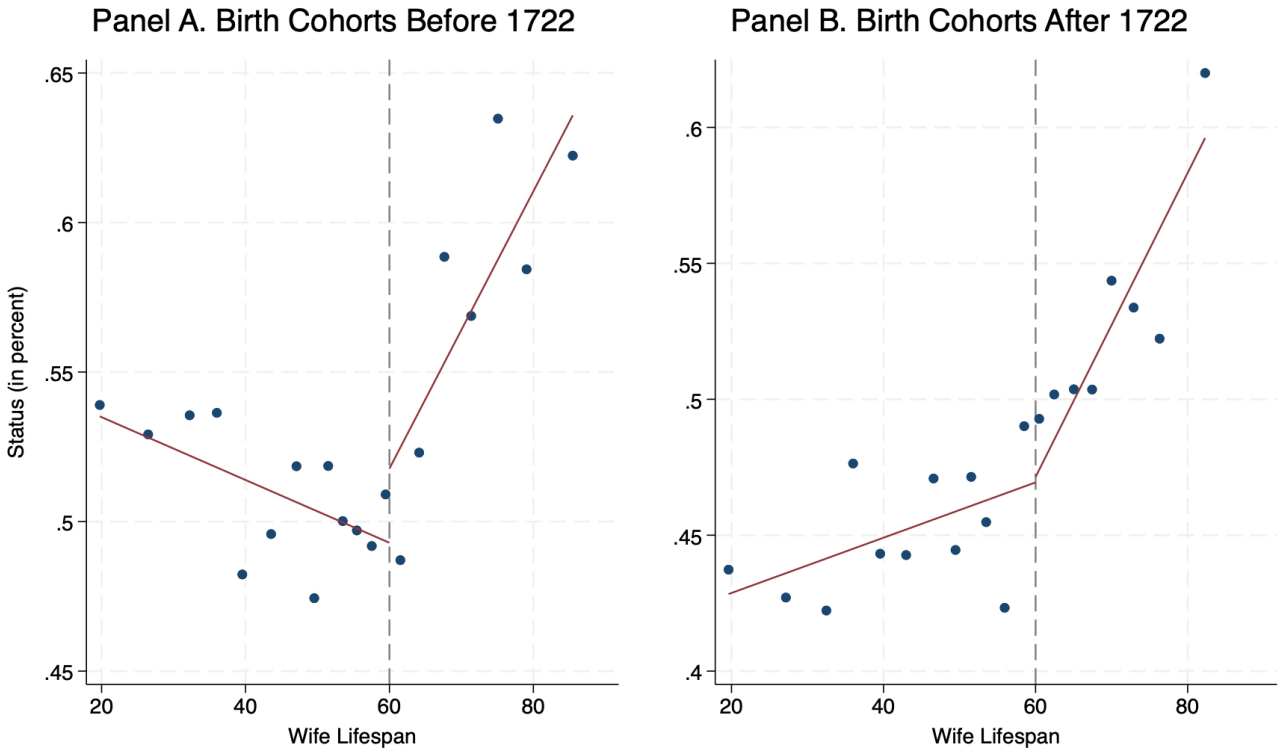
Figure 8: Linear Relationship between Husband Status and Wife Lifespan



Notes: Linear prediction of lifespan of wives married to men in 1775 or before based on the husbands' social status; n = 3,706.

The linearly-predicted difference between a wife of a husband of lowest versus highest status, percentile 0 to 100, is about 11 years (Figure 8). Also in the case of wife lifespan, it is the case that the gradient is steeper at higher levels of lifespan and status, which parallels our findings for the male health-status relationship (not shown). Furthermore, while in the early years of the sample a positive gradient was present only for wives of good health, over time a positive gradient emerges that holds for wives of either bad or good health. This can be seen in Figure 9.

Figure 9: Status - Female Health Relationship over Time



Notes: Scatter plot showing average percentile status by wife lifespan bin in the early (Panel A) versus late (Panel B) part of the sample; N = 1,847 in Panel A, N = 1,858 in Panel B.

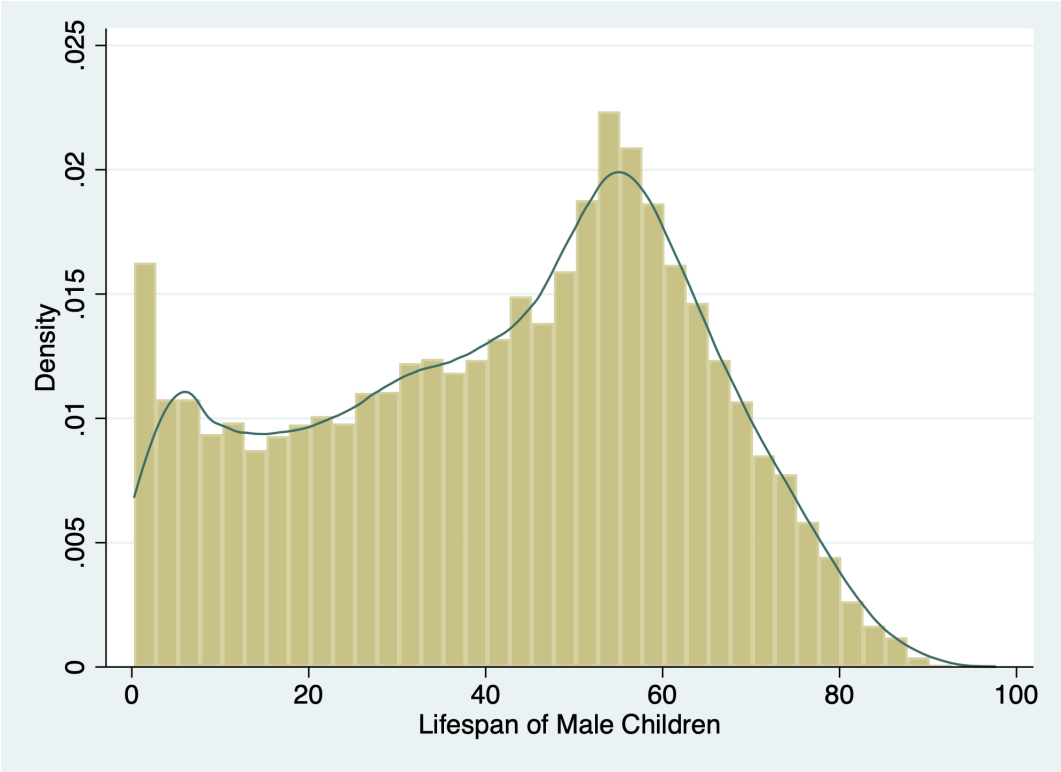
The diffusion of the positive wife health-husband status gradient over the sample period reinforces our earlier finding of the emergence of a positive male health-status gradient from the late 17th to the mid-18th century, see Figure 6.

5.2 Child Mortality and Father Status

This section asks whether there is a health-status gradient by examining the mortality of the children of our Tongcheng men. The sample is all male children of married men born by 1775 in our data who have a completed lifespan of 16 years or more (N = 9,406). Figure 10 shows the distribution of the childrens' lifespans.

Several salient patterns emerge, related to our genealogical data. First, average lifespan of these sons is 41.5 years, which is more than 10 years less than the average lifespan of their fathers (see Table 1). This is because in the father generation, the analysis is conditional on survival to age 16 and marriage, whereas analysis of son lifespan includes individuals who will not survive to adulthood, as well as those men who stay single. In this analysis, for example, the fraction of sons who do not live past their fifth birthday (60 months) is 6.4 percent.

Figure 10: Lifespan of Sons



Notes: Distribution of sons' lifespan of married men belonging to seven Tongcheng clans born by1775 and reaching a lifespan of at least 16 years; n = 9,406.

Second, even this formidable child mortality is likely an underestimate, given the moderate spike in deaths soon after birth compared to most historical records of premodern societies as well as for the Chinese elite. For example, infant and child mortality was about 10% among Qing nobility (Lee, Wang and Campbell 1994) and 8-10% for British elite boys (Hollingsworth 1957). The lower infant mortality rates (death within first 12 months of birth) in our data compared to estimates for this historical period are expected, given incomplete vital statistics especially for neonatal deaths as well as changing norms over the study period about when a son and his premature death should be recorded in the genealogy. (We know that recording of a child's birth is subject to considerable discretion, since baby girls were rarely recorded in the earliest period of our data, but are to some extent towards to the end.)

Moreover, the incompleteness of infant mortality likely differs by status, given differences in fertility and in the mortality rates in youth. Accordingly, we will interpret the data conservatively as representing mortality after survival to age 1 (i.e., net of infant mortality). Any protection from infant mortality conveyed by high status (as hinted at in the data) will increase the gradient relative to what we document. Third, the lifespan difference between generations is consistent with a stagnant or downward trend in adult lifespan in late imperial China (Broadberry et al. 2018, Chen and Peng 2022), albeit conditional on (potentially 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.

In the following we examine the extent to which son mortality varies with the social status of their fathers. Table 3 presents results for mortality rates at different ages. We focus on survival to age 5, 10, or 15 in our exploration of the gradient and in comparing these Chinese patterns of child survival to those available for historical European populations.

We see that consistent with a positive health-status gradient, mortality of these male children tends to be inversely related to the status of their fathers. For example, mortality by the age of 15 for sons of low status fathers is 18 percent, while it is only 16 percent for sons of intermediate-status fathers. It is also interesting to note that sons of fathers with very high status have the largest mortality advantage, as is true for the mortality of the wives, as well as for the own mortality of these very-high status men. This means that whatever channels are behind this mortality premium, they seem to operate both intra- and intergenerationally.

The analysis of child mortality also yields new insights on the evolution of mortality differences across status levels. In particular, while mortality by age 5 is similar for all sons except those of very high-status fathers (around 6-7 percent), by the age of 15 mortality for sons of low-status fathers is about 50 percent higher than the mortality of sons of high-status fathers (17.76 versus 12.55, respectively). This indicates that the health-status gradient

Table 3: Mortality Rates of Sons

	N	One Year or less (12 months)	Five Years or less	Ten Years or less	Fifteen Years or less
Low Father Status	6,460	0.84	6.72	12.40	17.76
Intermediate Father Status	1,447	1.11	7.32	12.65	16.10
High Father Status	749	0.53	6.01	9.88	12.55
Very High Status	745	0.13	2.95	4.56	6.58
All	9,401	0.80	6.46	11.62	16.20

Notes: Low status is father status percentile of 0.34, Intermediate is 0.76, High is 0.88, and Very High is status percentile of 0.96. Infant mortality (mortality for one year or less) underestimated due to genealogical reporting norms; mortality rates given in percent. N = 9,401.

doesn't only emerge because it diffuses from higher to lower status levels over time (recall Figures 6, 9), but it also strengthens in relation to life-cycle factors as sons get older.

5.3 Robustness

This section summarizes several important robustness checks. More information is given in section C.

Structural Factors Accounts of the Ming-Qing period indicate that achieving high status became harder over time. To account for this we have examined the health-status gradient allowing for a time-varying status definition. Figure A.9 shows that this does not change our main findings.

Progenitors and Early Birth Cohorts One might be concerned that the individuals in the first generation of each genealogy, the progenitors, have an outsized influence on the empirical findings, perhaps because they are unlikely to be commoners. Similarly, early birth cohorts might be special and results based on them misleading for a number of reasons. However, we show that dropping birth cohorts before the year 1600, which includes progenitors, does not greatly change our findings (see Figure A.10).

Status and Lifespan Differences between Clans Time-invariant differences between clans, in particular in the typical status achievements of members, can in principle affect

our findings on the relationship between health and status (e.g., Shiue 2025 finds that long-run clan differences are a source of social mobility). However, below we show through the inclusion of clan fixed effects that the health-status gradient is similar whether the analysis accounts for such differences across clans or not (see Figure A.11).

6 Discussion

Analyzing unique data on lifespan and status across multiple generations for a sample centered on the 1700s, we show that there was a substantial health-status gradient in adult lifespan over centuries of Malthusian imperial China. Married men of the highest status lived on average 7 years longer than married commoners, and their wives' average lifespan exceeded that of commoners' wives by 11 years. These survival advantages represent 13% and 20% longer lives than the average lifespan for married men and women, respectively, during this historical period. Our findings contribute to the understanding of the positive correlation between health and status over a relatively long span of history with sparse previous evidence, especially for commoners (representing more than 70% of our data). We also document that for most of the sample period, the lifespan advantage of status was strongest at the top of the status distribution, with longer lifespans diffusing to lower-status households in the 18th century.

Few other studies document a pre-20th century health-status gradient in a non-Western setting, much less during a period of rising and then static or falling living standards; the general increase in living standards and lifespan of the late 19th and early 20th centuries in Europe and the US are the subject of most previous studies (Preston 1975; Marmot et al. 1978; Cutler et al. 2011; Chetty et al. 2016; Black et al. 2024). Since our genealogical data primarily covers married men, the health and survival of other population subgroups would presumably demonstrate even larger disparities in health by status. Indeed, we present evidence of differences in child mortality by status that suggest differences in life expectancy at birth were probably larger that the 7-year gap in male lifespan between the highest- and lowest-status households that we obtain.

To assess the magnitude of the lifespan gradient we find for imperial China, it is useful to compare it to the earliest available and reliable cohort lifetables for Europe. We extract relevant survival percentiles from the earliest available cohort lifetables of European countries with good population-wide vital statistics, as compiled by the Human Mortality Database (HMD): for the 1700s (Sweden) and early 1800s (e.g. England & Wales). Cohort lifetables represent the appropriate counterpart to our data since they follow individuals born in a

given year 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 HMD lifetables over five-year periods to help smooth across historical fluctuations during this period.

This comparison shows that adult lifespan in this sample of central China was comparable to several parts of Europe during the 1750s through 1840s, and that the adult gradient in survival by status was of the magnitude of several decades of European survival improvement. Specifically, a seven-year difference in adult lifespan is equivalent to half a century of longevity improvement in industrializing England and Wales. The difference between England and Wales male birth cohorts of 1841-44 (the earliest available) and 1890-94 was 7.07 years lifespan (i.e., cohort life expectancy conditional on survival to age 15). Adult life expectancy was 58.04 for men born in England and Wales in 1841-1844, almost exactly the same as the average adult lifespan of the highest-status married men in our sample (58.22). The average lifespan of commoners (51) was slightly less than the 54.9 years on average that a man born in Sweden in the 1750s would live, conditional on surviving to age 15 (HMD, earliest cohort lifetable: Sweden 1751-1754). The chance of survival to age 50 conditional on surviving to age 15, S(50|15), was 63.5% for Swedish men born in the 1750s and 67.4% for the England and Wales male birth cohort of 1841-1844. We find that S(50|15) averaged 58% for low-status Chinese men and 72% for those of the highest status (conditional on survival to age 15 and marriage).

An alternative comparison of the strength of the gradient might be to represent survival differences by SES as a share of average lifespans in contemporary economies. Recall that our analyses suggest survival advantages of the highest-status individuals amounts to 13% of the average male lifespan and 20% of the average female lifespan, conditional on survival to adulthood and marriage. For the contemporary US, Chetty et al. (2016) find the gap in life expectancy between the richest and poorest 1% (measured by income in adulthood and thus conditional on survival to adulthood but not on marriage) to be 14.6 years for men and 10.1 years for women. These represent 19.1% of average male life expectancy (14.6/76.4) and 12.4% of average female life expectancy (10.1/81.2) in 2014, respectively. Interestingly, Chetty et al. (2016) find the average lifespan difference between Americans at the 97th and 35th percentiles of the national income distribution – the same percentiles of the status distribution that we analyze – to be 6.5 years for men and 4.4 years for women. Thus the health-status gradient as a fraction of typical lifespan for men is comparable or larger, while that for women is significantly smaller, than what we find for imperial China.[13]

[13]These life expectancy figures come from CDC analysis of the National Vital Statistics System by Murphy et al., available at https://www.cdc.gov/nchs/products/databriefs/db229.htm. The analysis of specific percentiles of the income distribution is based on the Chetty et al. (2016) estimates of adult US

We also explore the genealogical data for evidence of differences in survival to adulthood by household status, given the high infant and child mortality in pre-modern populations. Analyzing genealogical data for six lineages in Ming-Qing China, Hu (2023) finds that lower status is associated with lower reproduction, as measured by surviving sons. Studying the Qing nobility, Lee, Wang and Campbell (1994) estimate life expectancy at birth during this period was below 40. In the earliest available HMD data, Swedish life expectancy at birth was also below 40 throughout the 1700s, with at least 40% of babies dying before age 15. The chance of a baby boy surviving to age 10 conditional on surviving to age 1, S(10|1)), was 79.1% in Sweden 1751-54 and 82.2% in England and Wales 1841-44; these are equivalent to under-ten mortality rates of 20.9% and 17.8% respectively when infant mortality is removed (since we lack accurate data for infant mortality). These rates are a similar order of magnitude to the 11.6% under-ten mortality rate that we find. In England and Wales, male S(10|1) declined from 17.8% in 1841-44 to 8.8% in 1900-1904, comparable to the difference in under-ten mortality between a commoner (12%) and a high-status son (5%).

These results for child mortality are suggestive that health may have played an important role intergenerationally, especially for higher-status households endeavoring to protect or improve their status. Although Hu (2030) documents higher net reproduction among higher-status households, this arises in part from lower child mortality and greater probability of marriage. Controlling for marital fertility to invest more in each son – a quantity-quality trade-off for male heirs – is less risky if child mortality can be kept low (and the mother would be more likely to survive the childbearing years to invest in the son's early education, enhancing cultural transmission). Other research has explored aspects of this mechanism. Shiue (2017) shows a negative correlation between fertility and education when the return to education was high; Bai et al. (2023) find that households belonging to the Eight Banner System in Liaoning in the mid-Qing dynasty did use fertility control to achieve a quantity-quality trade-off; and Shiue and Keller (2026) explore the role of the quantity-quality trade-off in elite families' recovery from the devastation of the Ming-Qing transition. In further research, we study intergenerational persistence of lifespan and its relationship to status mobility in imperial China.

life expectancy conditional on survival to age 40 for every percentile of the US national income distribution (pooling 2001-2014). Although not all percentiles are reported in the main text, these estimates are graphed in "Figure 3. Changes in Race- and Ethnicity-Adjusted Life Expectancy by Income Group, 2001-2014," and the authors make available underlying data files at various levels of aggregation. See posted data files at Opportunity Insights or https://www.healthinequality.org/data/. The first-listed dataset (health_ineq_online_table_1.dta) provides the dataset for "Online Data Table 1: National life expectancy estimates (pooling 2001-14) for men and women, by income percentile." We use this dataset to extract the numbers for male and female race-adjusted life expectancy at given percentiles.

Despite the many strengths of our data for studying lifespan and status, 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 representative of China as a whole. We observe health and status along the male line, and generally only for those sons able to marry; although we can also parse the data to understand marital sorting, lifespans of married women, and some aspects of child mortality, the incomplete recording of live births and neonatal mortality, especially for girls, forces a focus on adult lifespans (i.e., conditional on survival to age 15). Nevertheless, the patterns we document are robust to limiting the sample to the most complete vital statistics and to accounting for changes in the cost of achieving status. 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. In ongoing work we are documenting intergenerational persistence in lifespan and its relationship with status mobility.

In sum, we document the evolution of the health-status gradient across multiple generations in imperial China, with a sample concentrated in the 17th and 18th centuries. Based on rich genealogical data that includes clear markers of status defined by China's national civil service examination system as well as occupation and wealth, we find that compared to married men without any evidence of wealth or status, those who achieved the highest rank lived on average seven years longer. This difference in adult lifespan by status was of comparable magnitude to several decades of survival improvement in industrializing England. Some suggestive evidence also supports a gradient in female and child mortality over this period, indicating that life expectancy at birth may have exhibited an even stronger status gradient. These findings provide new evidence on the health-status gradient for a less-studied, non-western setting, as well as highlight the usefulness of an intergenerational household perspective for understanding inequality in health, socioeconomic status, and their inter-linkage.

References

[1] Bai, Yu, Yanjun Li, and Pak Hung Lam (2023), "Quantity-quality trade-off in Northeast China during the Qing dynasty." Journal of Population Economics 36, no. 3: 1657-1694.

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

[3] Baker, Michael, and Mark Stabile (2011), "Determinants of Health in Childhood." In The Oxford Handbook of Health Economics, edited by Sherry Glied and Peter C. Smith, 164–187. Oxford: Oxford University Press.

[4] Baten, Joerg, Debin Ma, Stephen Morgan, and Qing Wang (2010), "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: 347-359.

[5] Beattie, Hilary (1979), 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.

[6] Black, Sandra E., Neil Duzett, Adriana Lleras-Muney, Nolan G. Pope, and Joseph Price (2026), "Intergenerational Transmission of Lifespan in the US", Columbia University.

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

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

[9] Butaeva, Kristina, Steven N. Durlauf, and Alexander Shapoval (2026), "Intergenerational Mobility in Late Qing Dynasty (1786-1906): Evidence from Northeast China", University of Chicago working paper.

[10] Chang, Chung-li (1955), The Chinese Gentry. Seattle: University of Washington Press.

[11] Chang, Chung-li (1962), The Income of the Chinese Gentry. Seattle: University of Washington Press.

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

[13] Chen, Zhiwu, and Kaixiang Peng (2022), "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, pp. 676-709.

[14] Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler (2016), "The association between income and life expectancy in the United States, 2001-2014." Journal of American Medical Association 315, no. 16: 1750-1766.

[15] Coale, A. J. and P. Demeny (1966), Regional Model Life Tables and Stable Populations. Princeton University Press.

[16] Cutler, David, Angus Deaton, and Adriana Lleras-Muney (2006), "The determinants of mortality." Journal of economic perspectives 20, no. 3: 97-120.

[17] Cutler, David, Adriana Lleras-Muney, and Tom Vogl (2011), "Socioeconomic status and health: dimensions and mechanisms." Chapter 7 in Oxford Handbook of Health Economics, pp. 124-163.

[18] Dahl, Molly and Thomas DeLeire (2008), "The association between children's earnings and father's lifetime earnings: estimates using administrative data." IRP Discussion Paper 1324-08.

[19] Deaton, Angus (2016), "On Death and Money: History, Facts, and Explanations", Journal of American Medical Association, April 26, 2016.

[20] Elman, Benjamin A. (2002), "The social roles of literati in early to Mid-Ch'ing."" The Cambridge History of China 9, no. Part 1: 360-427.

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

[22] Gakidou, Emmanuela, Krycia Cowling, Rafael Lozano, and Christopher JL Murray (2010), "Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis." The lancet 376, no. 9745: 959-974.

[23] Harrell, Stevan (1987), "On the Holes in Chinese Genealogies", Late Imperial China, Vol. 8(2), pp. 53-79.

[24] Hollingsworth, T. (1957), "A demographic study of the British ducal families", Population Studies 11(1), 4-26.

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

[26] Hu, Sijie (2025), "Evolutionary advantage of moderate fertility during Ming–Qing China: a unified growth perspective", Journal of Economic Growth Vol. 30(4): 497-519.

[27] Hu, Sijie (2023), "Survival of the literati: Social status and reproduction in Ming–Qing China." Journal of Population Economics 36, no. 4 (2023): 2025-2070.

[28] 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).

[29] Keller, Wolfgang, and Carol H. Shiue (2023), "Intergenerational Mobility of Daughters and Marital Sorting: New Evidence from Imperial China", NBER Working Paper No. 31695.

[30] Keller, Wolfgang, and Carol H. Shiue (2022), "Foreign Trade and Investment" in *The Cambridge Economic History of China*, 1800 to Present, Debin Ma and Richard Van Glahn (eds.), Cambridge University Press.

[31] Knodel, John, and Edward Shorter (1976), "The Reliability of Family Reconstitution Data in German Village Genealogies (Ortssippenbucher)", Annales de Demographie Historique, pp. 115-54.

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

[33] Lavely, William (1998), 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: 256-258.

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

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

[36] Liu Ts'ui-jung (1978), "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.

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

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

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

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

[41] Miller, Grant, and B. Piedad Urdinola (2010), "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: 113-155.

[42] Minardi, Saverio, Giulia Corti, and Nicola Barban (2024), "Historical Patterns in the Intergenerational Transmission of Lifespan and Longevity: A Research Note on U.S. Cohorts Born Between 1700 and 1900", Demography 61 (4): 979–994.

[43] Norton, Susan L. (1980), "The Vital Question: Are Reconstructed Families Representative of the General Population?", in Genealogical Demography, Bennett Dyke and Warren T. Monill (eds.), New York: Academic Press.

[44] Perkins, Dwight H. (1969), Agricultural Development in China, 1368-1968. Chicago, Aldine publishing.

[45] Pomeranz, Kenneth (2001), The Great Divergence: China, Europe, and the Making of the Modern World Economy. Princeton University Press.

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

[47] Rawski, Evelyn S. (1979), "Education and popular literacy in Ch'ing China." Ann Arbor: University of Michigan Press.

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

[49] Shiue, Carol H. (2004), "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: 100-124.

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

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

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

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

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

[55] Shiue, Carol H., and Wolfgang Keller (2026), "People versus Places: Elite Persistence After the Fall of the Ming", NBER # 34451.

[56] Schultz, T. Paul (2002), "Why Governments Should Invest More to Educate Girls." World Development 30 (2): 207–225.

[57] Stelter, Robert, and Diego Alburez-Gutierrez (2022), "Representativeness is crucial for inferring demographic processes from online genealogies: Evidence from lifespan dynamics", Proceedings of National Academy of Sciences 119 (10) e212045511, https://doi.org/10.1073/pnas.2120455119

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

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

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

[61] 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.

[62] United Nations (1982), Model life tables (building on Coale-Demeny 1966). Described and available at https://www.un.org/development/desa/pd/data/model-life-tables.

[63] Wang, Feng, Cameron D. Campbell, and James Lee (2010), "Agency, hierarchies, and reproduction in northeastern China, 1789 to 1840", in P*rudence and Pressure: Reproduction and Human Agency in Europe and Asia, 1700-1900,* edited by Noriko O. Tsuya, Feng Wang, George Alter, and James Z. Lee, MIT Press.

[64] Weir, David R. (1995), "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: 1-26.

[65] Wolpin, Kenneth I. (1997), “Determinants and Consequences of the Mortality and Health of Infants and Children”, in Handbook of Population and Family Economics, vol. 1A, edited by Mark R. Rosenzweig and Oded Stark, 483–557. Amsterdam: Elsevier.

Online Appendix

A Data

A.1 Social Status Categories

The sample comes from Tongcheng county in Anhui province. The status coding is given in Table A.1, column (1). It is related to a broader set of standardized descriptors which classify a man's social status into 23 ranked categories. The 23 categories are shown in column (2) of Table A.1, with descriptors shown in column (3). All measures in the genealogies are lifetime measures, and what is given would be the highest lifetime achievement.

Table A.1: Elite Status and Genealogy Descriptors

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

Notes: Table gives information on a man's status using four groups (column 1), based on 23 subgroups (column 2) using standardized descriptors in the genealogies (column 3). Analysis based on Telford (1986a, 1992), Chang (1955, 1962), Ho (1962), and Eberhard (1962).

There were multiple pathways to elevated social status in imperial China. Monetary wealth was one, but wealthy farmers, landowners, and merchants did not occupy the highest social ranks.[14] There is broad agreement that *keju* achievement represented the apex of social prestige in imperial China (Ho 1962; Chen, Kung, and Ma 2020). Within the *keju* hierarchy, *shengyuan* was the entry-level degree, awarded upon passing local (county) examinations. Exceptional *shengyuan* could receive the *gongsheng* ("imperial student") title, while higher levels included *juren* (provincial graduates) and *jinshi* (national graduates). The structure

[14]Unless they also held civil service examination (*keju*) degrees, these individuals belonged to class 8 (Table A.1).

was tournament-style: candidates had to earn lower degrees before attempting higher ones. Degrees and offices were occasionally sold, and our baseline specification keeps these cases separately. There were no formal age restrictions, but the examinations required extensive study, so successful *jinshi* were typically in their twenties or older, and many men remained *shengyuan* well into middle age. Those who failed to advance often pursued alternative careers, such as assisting officials, managing local affairs, settling disputes, organizing public works, or providing education (Chang 1962). Thus, *keju* preparation itself yielded social and economic returns even for unsuccessful candidates.

High status translated into differences in income and wealth. Prior studies document substantial monetary returns to *keju* achievement (Ho 1962; Chen, Kung, and Ma 2020). Chang (1962) provides detailed evidence on the earnings of *jinshi* and officials, noting that degree level strongly correlated with rank in the nine-tier official hierarchy of the late nineteenth century. A county magistrate held seventh-rank status, while a provincial governor ranked second. Although the mapping between degree and position was not deterministic, higher degrees generally led to higher-ranking and better-remunerated posts. Reaching top administrative positions without a *juren* or *jinshi* degree was exceedingly rare, while most *jinshi* held posts above the district level.

Degree level is a valuable proxy for socioeconomic position because it is consistently recorded in the genealogies. While official salaries are known for certain periods (see Chang 1962), they constituted only a fraction of total compensation. Informal income sources—such as local "contributions" and gifts—often exceeded formal pay, though systematic data is scarce.

B The Influence of Data Reconstitution

To avoid selection bias, our baseline sample includes reconstituted vital statistics. This is a standard practice because historical data for this period tends to be incomplete. In the present case, we rely on auxiliary information in the genealogies, as well as on life tables, see Telford (1990). As noted in the text, without data reconstitution the sample would suffer from selection bias because information for poorer men tends to be less reliably reported and is thus more frequently missing. This section assesses the role of reconstitution for the results, which is feasible because we know which data is reconstituted and which is not. The following presents results based on unreconstituted data and compares them with the baseline findings in the text.

The next section examines the relationship between adult male health and status in the

unreconstituted sample, before turning to the relation between (1) wife and (2) son health to male status in the unreconstituted sample.

B.1 Male Health and Status in the Unreconstituted Sample

Table A.2 provides information on the size and status composition of the unreconstituted sample.

Table A.2: Status in the Unreconstituted Sample

Status Group	Percentile (mid-point)	N	Fraction of Sample
Low	0.321	1,854	0.642
Intermediate	0.725	476	0.165
High	0.851	257	0.089
Very High	0.948	301	0.104
Total		2,888	1.000

Notes: Sample is all married men from seven Tongcheng clans in birth cohorts before 1775 with at least 16 years of lifespan whose vital statistics are not adjusted or estimated.

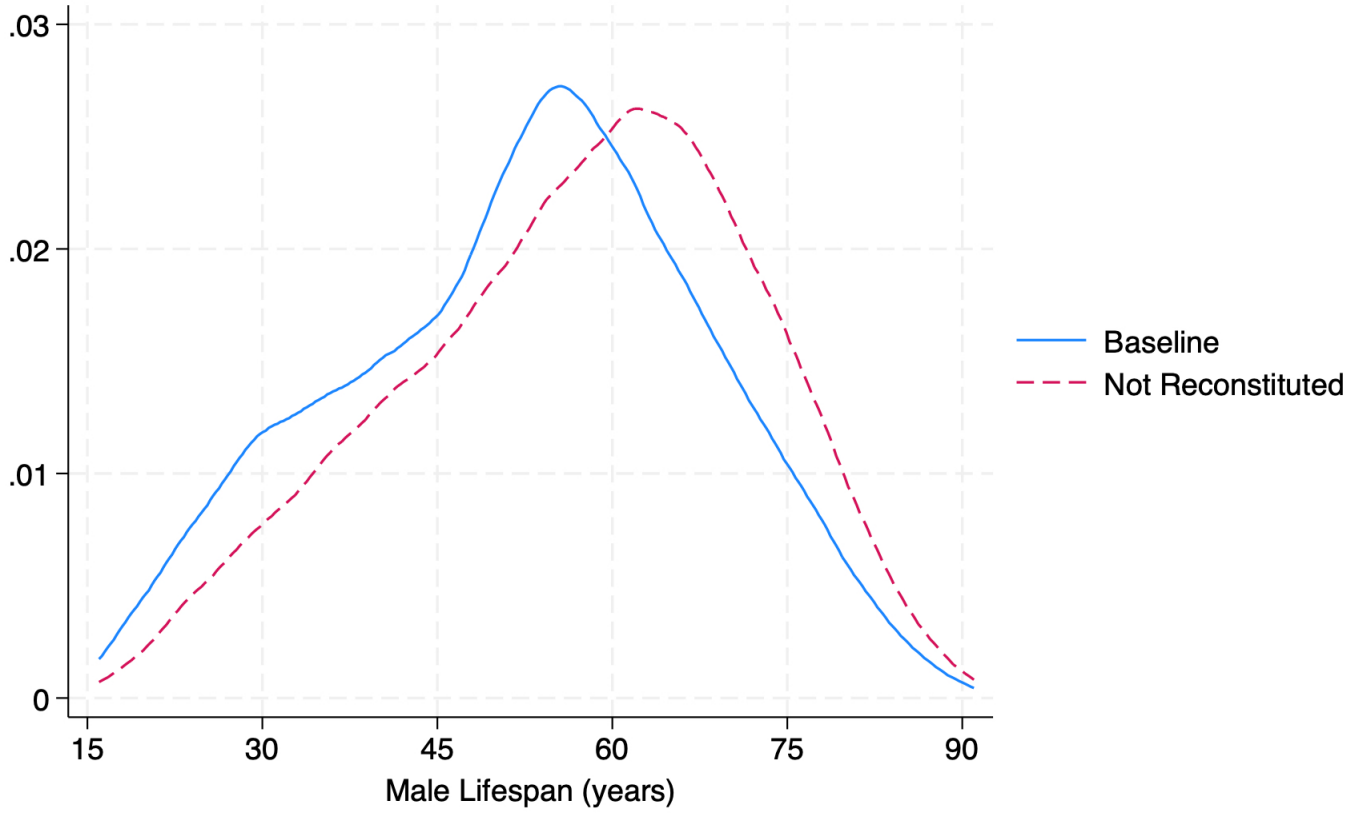
We see that the composition of status in this sample is shifted to higher-status men, compared to the baseline sample. In particular, the share of lowest-class men is 64 percent, in contrast to 71 percent in the baseline sample. Also, the share of men with very high status is now above 10 percent whereas in the baseline sample it is less than 7 percent.

Figure A.1 shows the lifespan densities for the baseline and the unreconstituted samples. We see that the latter is shifted to the right; average lifespan is 56.8 for the unreconstituted data, compared to 52.4 for the baseline. This difference in average lifespan in the two samples is largely due to 'missing' low status men. This is evident from Figure A.2, which compares the densities between baseline and unreconstituted samples separately for low status and very high status men (left versus right, respectively).

We see that the difference between the densities for low status men is much larger than for very high status men. As a result, the unreconstituted sample disproportionately includes high-status men that tend to have long completed lifespans. This difference in sample composition translates into different health-status gradients that one obtains from the two samples. We begin with the linear prediction of lifespan as a function of status, see Figure A.3.

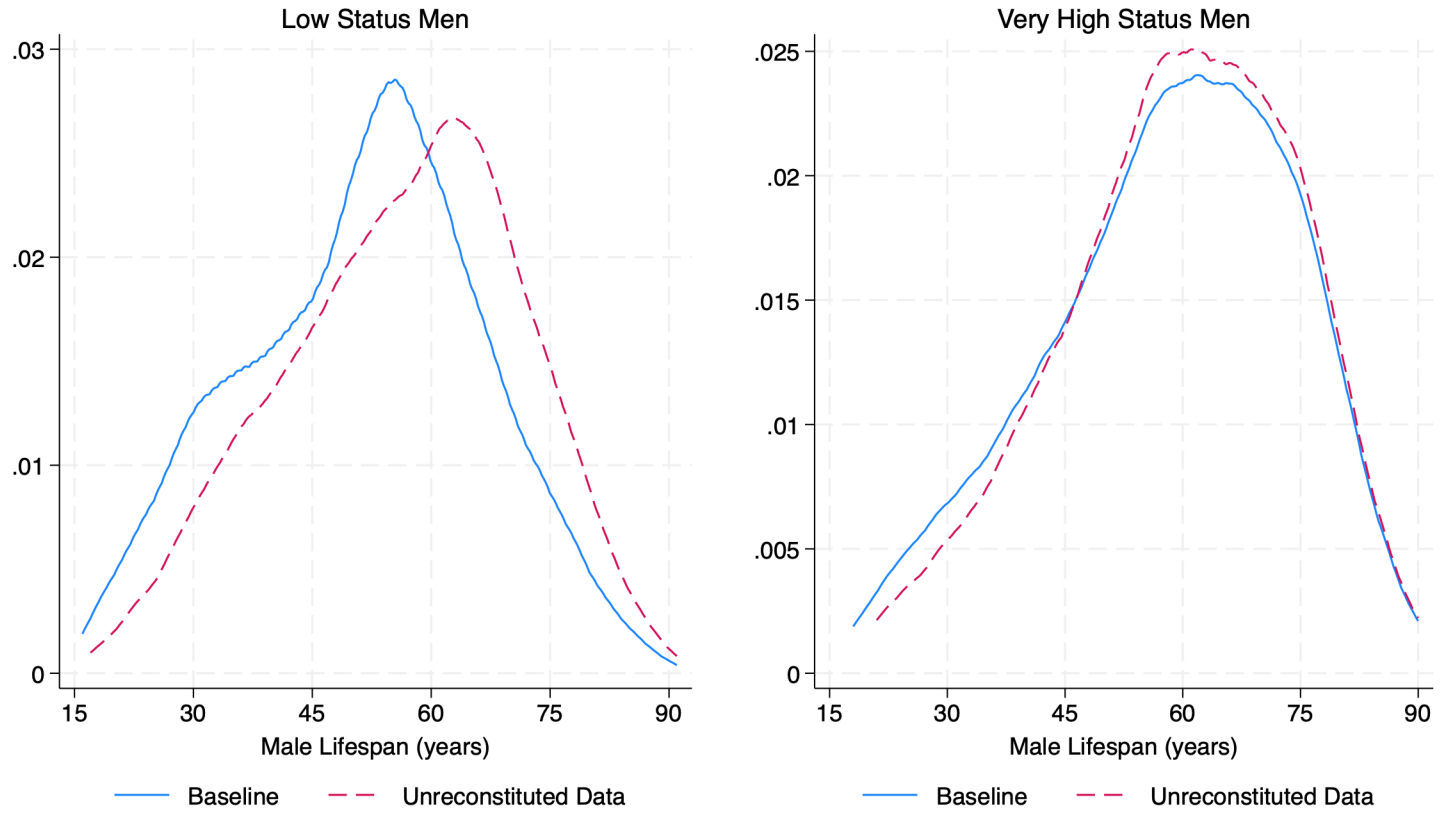
Baseline and unreconstituted results are different in two key respects. First, the gradient with unreconstituted data is located above that from baseline data. This follows because

Figure A.1: Lifespan Distribution: Baseline versus Not Reconstituted Data



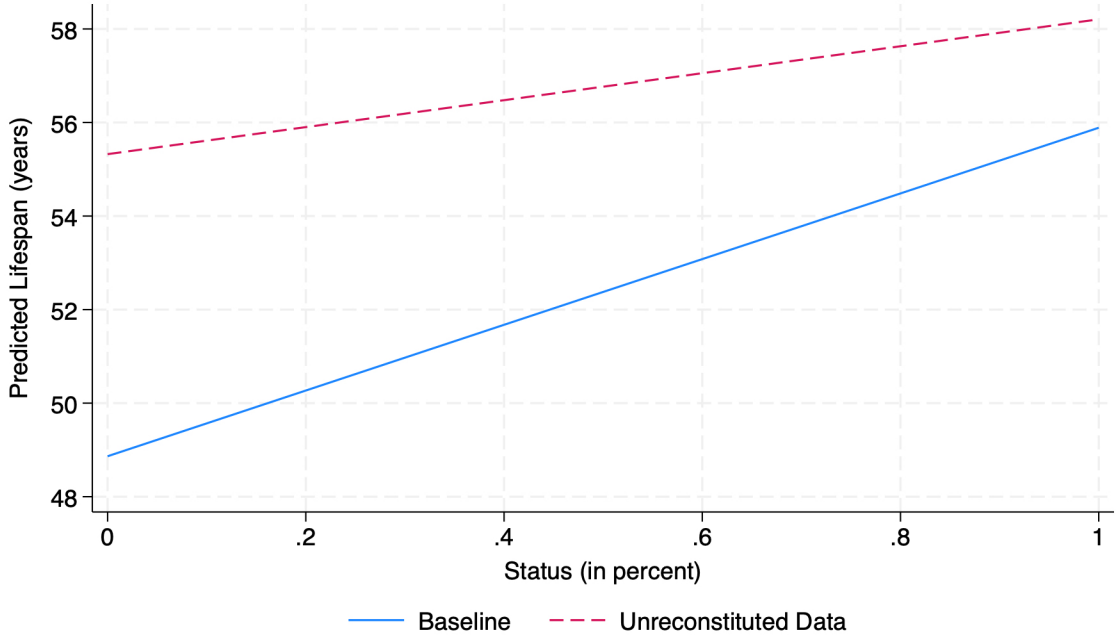
Notes: Completed lifespan of married men from seven Tongcheng lineages with birth by 1775 and at least 16 years of lifespan. baseline n = 4.656, limited sample n = 2,888.

Figure A.2: Lifespan Distributions by Status: Baseline versus Unreconstituted Data



Notes: Completed lifespan of married men from seven Tongcheng lineages with birth by 1775 and at least 16 years of lifespan. baseline n = 4.656, limited sample n = 2,888.

Figure A.3: Linear Prediction of Lifespan by Status: Baseline versus Unreconstituted Data

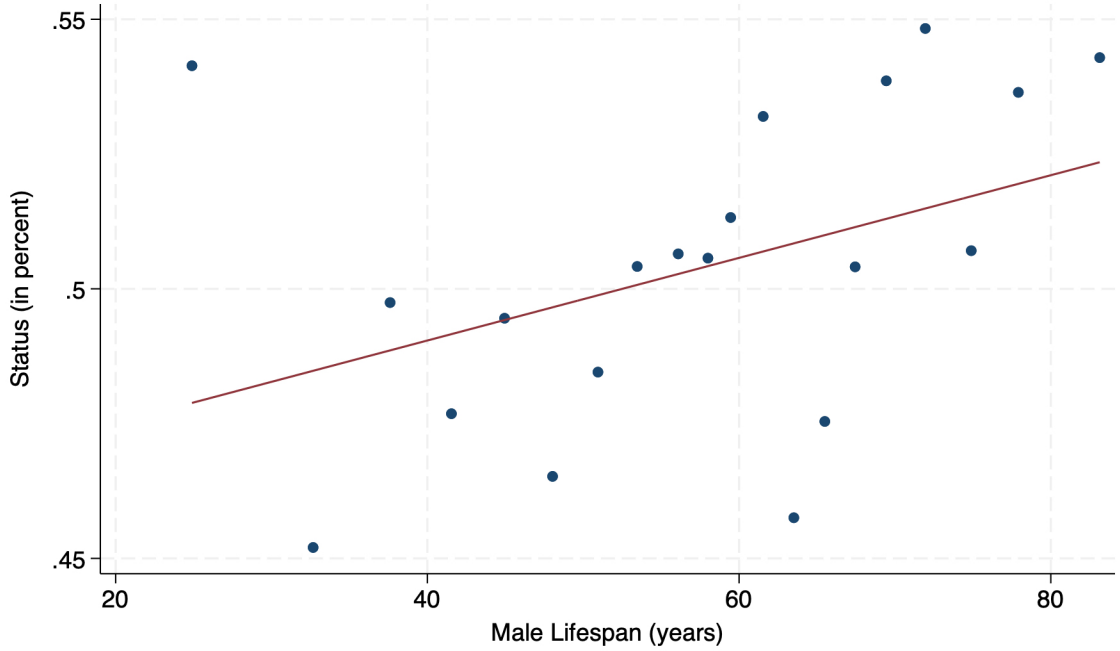


Notes: Completed lifespan of married men from seven Tongcheng lineages with birth by 1775 and at least 16 years of lifespan. baseline n = 4.656, limited sample n = 2,888.

missing data tends to be for men who died relatively early. Second, the gradient using unreconstituted data is less steep. This is the case because those men who died relatively early and whose data is disproportionately missing tend to be relatively poor. Put differently, the reason why the data is disproportionately missing is that these men have low status and poor health (see also Telford 1990).

However, while there are quantatative differences in the health-status gradient depending on whether one includes reconstituted data or not, qualitative findings are the same. This is also true when we consider the relationship between lifespan on the horizontal and status on the vertical axis, see Figure A.4.

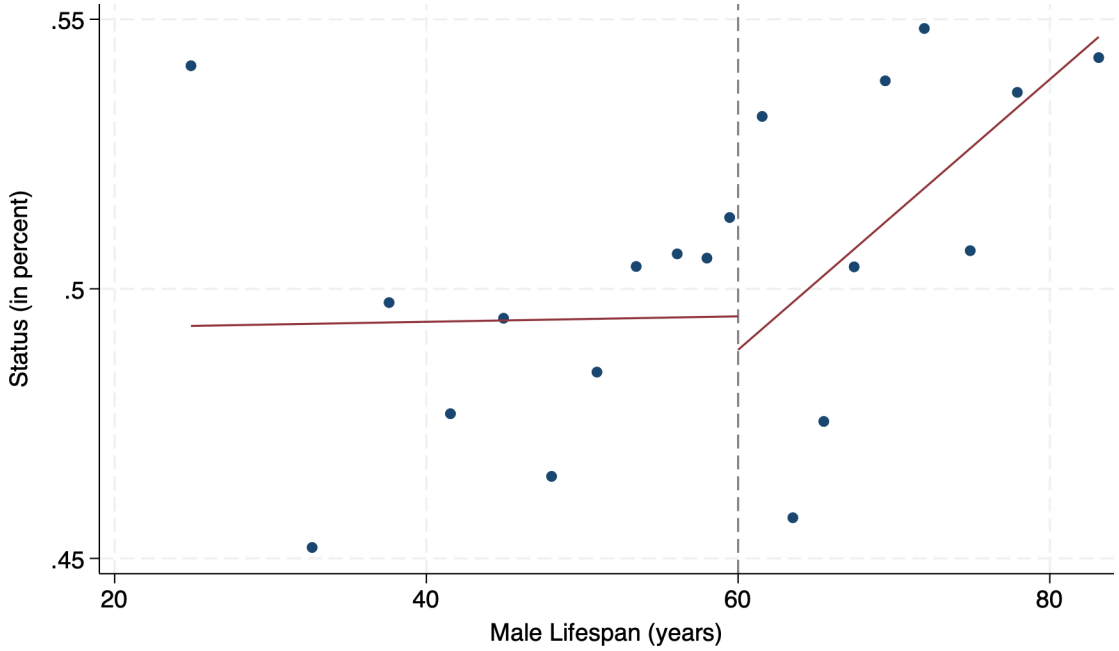
Figure A.4: Typical Status for Different Lifespan Bins with Unreconstituted Data



Notes: Completed lifespan of married men from seven Tongcheng lineages with birth by 1775 and at least 16 years of lifespan, unreconstituted data, N = 2,888.

As with the baseline data, there is a positive gradient between lifespan and status using unreconstituted data. Moreover, there is also evidence that the gradient is steeper at higher levels of lifespan and status, which parallels the finding in the text; compare Figures A.5 and 4, Panel B, respectively.

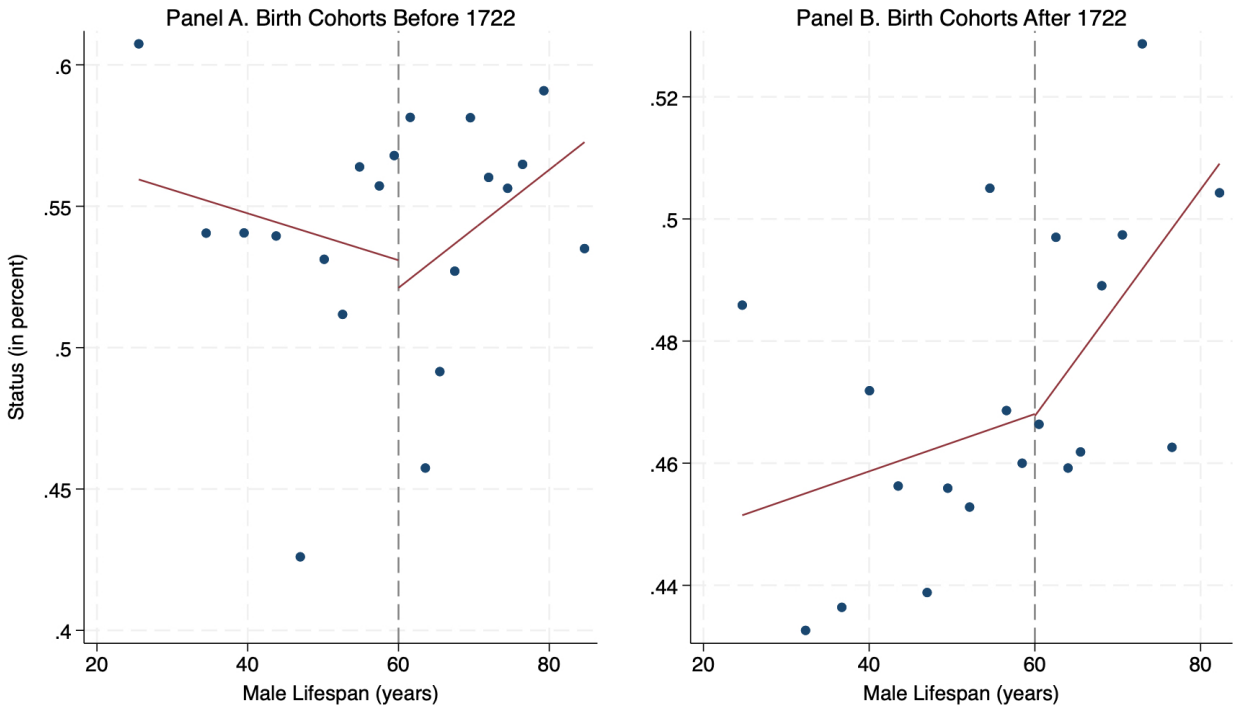
Figure A.5: Typical Status for Different Lifespan Bins with Unreconstituted Data: Nonlinear



Notes: Completed lifespan of married men from seven Tongcheng lineages with birth by 1775 and at least 16 years of lifespan, unreconstituted data, allowing for different slopes at lifespan of 60; N = 2,888.

Furthermore, also employing unreconstituted data we find evidence that over the sample period the range of the positive health-status gradient is increasingly extended to the entire population, see Figure A.6.

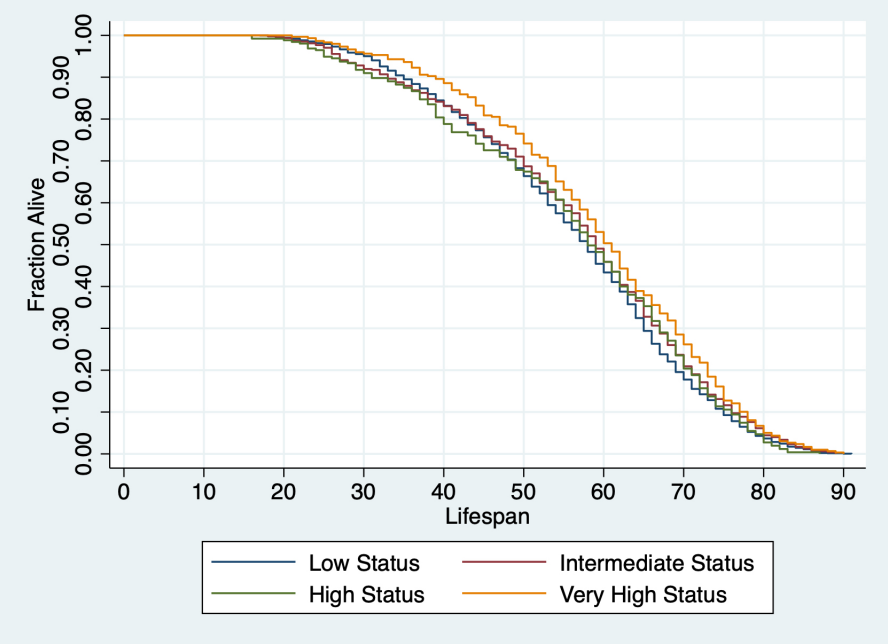
Figure A.6: Gradient Evolution: Unreconstituted Data



Notes: Completed lifespan of married men from seven Tongcheng lineages with birth by 1775 and at least 16 years of lifespan, unreconstituted data, allowing for different slopes at lifespan of 60; birth cohorts before 1722 on the left, birth cohorts after 1722 on the right.

We conclude this analysis of the health-status gradient with unreconstituted data for adult men by considering Kaplan-Meier survival estimates. They are shown in Figure A.7.

Figure A.7: Survival by Status: Unreconstituted Sample



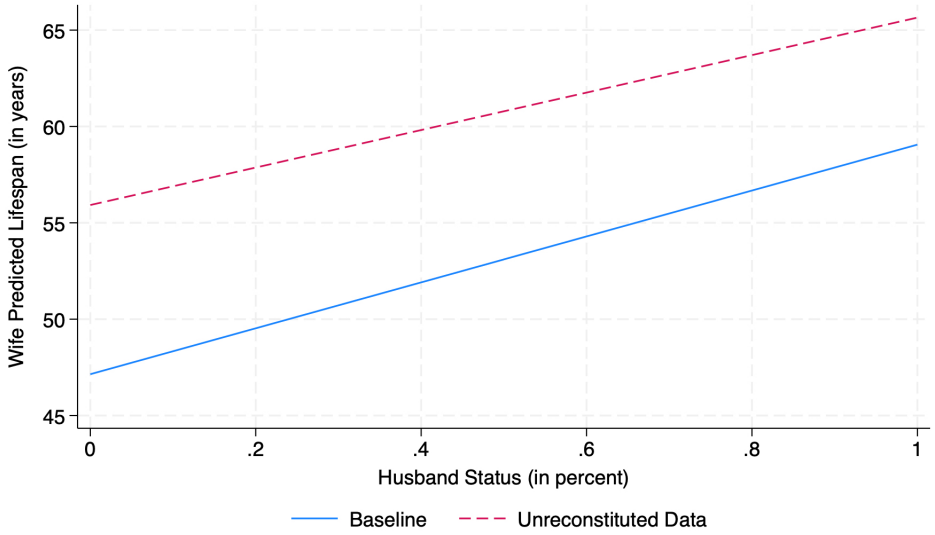
Notes: Kaplan-Meier survival estimates by status; unreconstituted data; n = 2,888.

Survival analysis for the smaller sample shown in Figure A.7 confirms the finding in the text that the chance of survival is higher for men of higher status. At the same time, the magnitude of survival differences for men of different status levels is smaller than in the full sample. For example, the difference in the probability of survival to 50 between highest and lowest status men is now 8 percentage points (75% versus 67%), in contrast to 14% for the baseline sample. This parallels our findings regarding the steepness of the health-status gradient above.

B.2 Wife Health and Husband Status in the Unreconstituted Sample

Figure A.8 shows the linear prediction of wife lifespan based on her husband's social status for the unreconstituted data.

Figure A.8: Wife Lifespan and Husband Status: Unreconstituted Sample



Notes: Shown is linear prediction of wife lifespan based on husband status; sample is all married men from seven Tongcheng clans in birth cohorts before 1775 who have at least 16 years of lifespan; baseline employs data for all women who are single-lifetime wives, while Unreconstituted Data sample employs the subset of wives of husbands whose vital data is not reconstituted and whose own vital data is also not reconstituted (N = 1,720).

Figure A.8 provides evidence for a health-status gradient irrespective of whether reconstituted data is included in the sample or not. Furthermore, as in the case of the male health-status gradient, the relationship becomes less steep if we drop reconstituted data (the OLS slope coefficient is 11.9 for the Baseline regression, compared to 9.8 for the Unreconstituted data). We note that the slope difference is smaller than in the the case of the male health-status relationship.[15]

B.3 Mortality of Son and Father Status: Unreconstituted Sample

In the following we examine the extent to which son mortality varies with the social status of their fathers in the unreconstituted sample. Table A.3 presents results for mortality rates at different ages. The very low infant mortality rates (death within first 12 months of birth) are expected, given changing norms over the study period about when a son and his premature death would be recorded in the vital statistics. Accordingly, account for omitted infant mortality in interpreting survival to age 5, 10, or 15, and in comparing these Chinese patterns of child survival to those available for historical European populations.

[15]Figure A.8 shows results for the sample that excludes observations in which either male or female vital statistics are reconstituted. At the same time, findings that include observations with reconstituted male *or* female vital statistics are similar.

Table A.3: Mortality Rates of Sons: Unreconstituted Sample

	N	One Year or less (12 months)	Five Years or less	Ten Years or less	Fifteen Years or less
Low Status	4,116	0.68	5.56	9.57	13.27
Intermediate Status	1,088	1.01	6.71	11.21	14.51
High Status	576	0.69	6.08	8.85	11.11
Very High Status	710	0.14	2.82	4.51	6.06
All	6,490	0.68	5.50	9.23	12.43

Notes: Mortality rates given in percent. Low Status is 32th percentile, Intermediate is 72nd, High is 85th, and Very High is 95th percentile; number of sons for which father status information in the limited sample is available is n = 6,490.

Generally, we see that son mortality rates in the unreconstituted sample are lower than the corresponding rates in our baseline sample. For example, mortality by age 10 is now 9.23%, compared to 11.62% in the baseline sample (see Figures A.3 and 3, respectively). This is to be expected given that we have seen above that going from the baseline to the unreconstituted sample, one drops disproportionately men of low status and poor health. At the same time, also in the unreconstituted sample it is the case that mortality rates tend to be lower the higher is the status of the father.

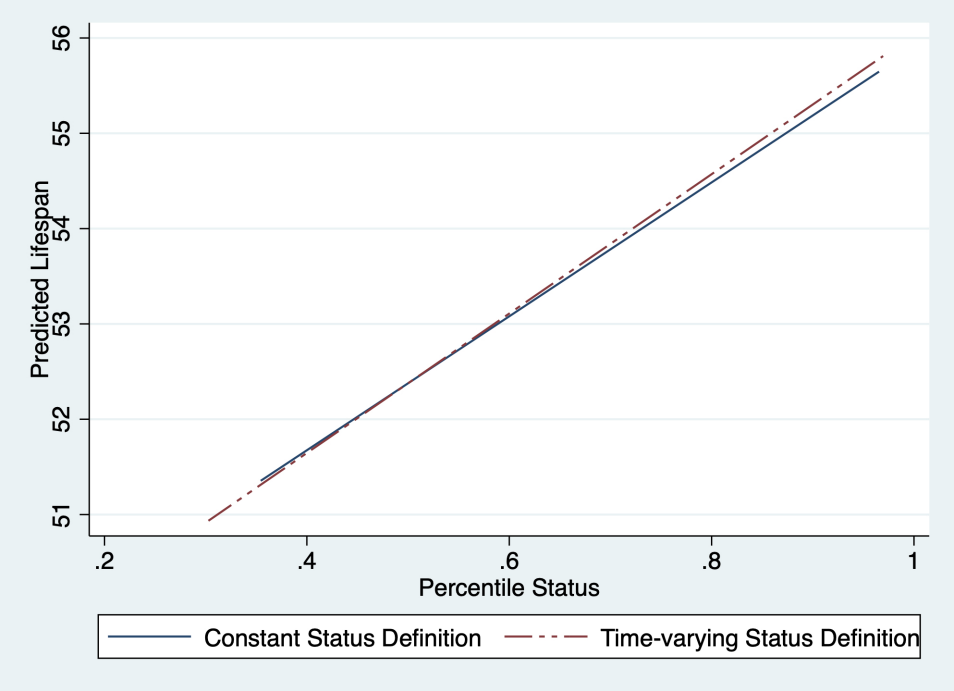
However, there are some differences in the patterns of how mortality rates between the baseline and unreconstituted samples varies with status. For example, the under-ten mortality rate for the highest status sons is virtually the same in the unreconstituted and the baseline sample (4.51% and 4.56%, respectively). In contrast, under-ten mortality rates for the lowest status sons are 12.4% in the baseline and 9.57% in the unreconstituted sample, which is a thirty percent difference. This pattern is in line with the idea that the unreconstituted data undersamples lower-status observations with relatively short lifespans.

C Robustness

C.1 Structural Factors Affecting the Status Distribution

Generally, social mobility during a particular time is affected by structural factors such as the number of high-ranking positions that are available (see Butaeva, Durlauf, and Shapoval 2026 for the case of Qing China). This section explores the role of changes in the ability of the men to achieve high status over the sample period, because we know that in the later part of

Figure A.9: Health-Status Gradient and Structural Factors

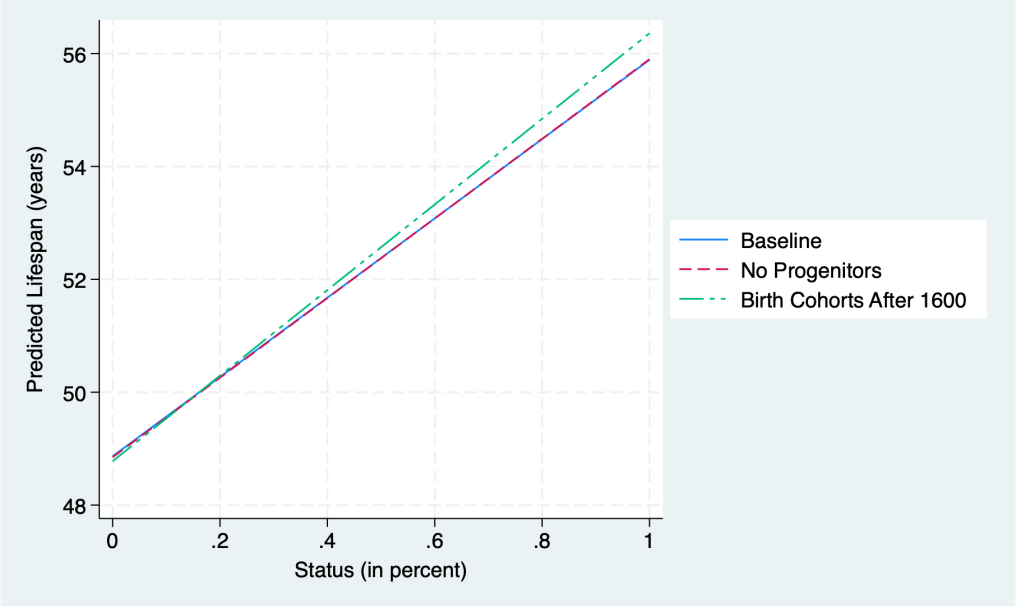


Notes: Linear fit for constant and time-varying status definition, the latter for birth cohorts before and birth cohorts after 1700; baseline sample of all men from seven Tongcheng clans in birth cohorts of 1775 or before with at least 16 years of completed lifespan; N = 4,656.

the Qing dynasty it became harder for men to achieve high status (Shiue 2025). One reason was that despite high rates of population growth, the number of official positions remained about the same. Thus, while the average percentile status for the entire sample period is equal to 0.5 by construction (see Table 1), the average status of cohorts born before 1700 is 0.55, while cohorts born after 1700 have only an average percentile status of 0.47.

To see what role this changing structural factor plays for the health-status gradient, we calculate the mens' percentile ranks separately for birth cohorts before and after 1700. This controls for the fact that status became harder to obtain after 1700 because by construction, the average percentile status rank is 0.5 for each subperiod (born before and after 1700). Figure A.9 compares the linear health-status gradient for the baseline definition of status with results for this time-varying definition of status. Figure A.9 indicates that results that account for status becoming harder to obtain are similar to findings based on an analysis that does not take account of this time-varying structural factor. We conclude that changes in the status distribution over the sample period have a limited influence on our health-status gradient findings.

Figure A.10: The Role of Progenitors and Early Cohorts



Notes: Linear fit between lifespan and status for the sample excluding progenitors, defined as generation 1 in the genealogies, compared with baseline results. All married males from seven Tongcheng clans in birth cohorts 1775 or earlier with at least 16 years of lifespan. Also shown is linear prediction of lifespan based on status for birth cohorts 1600 and later (N = 4,249).

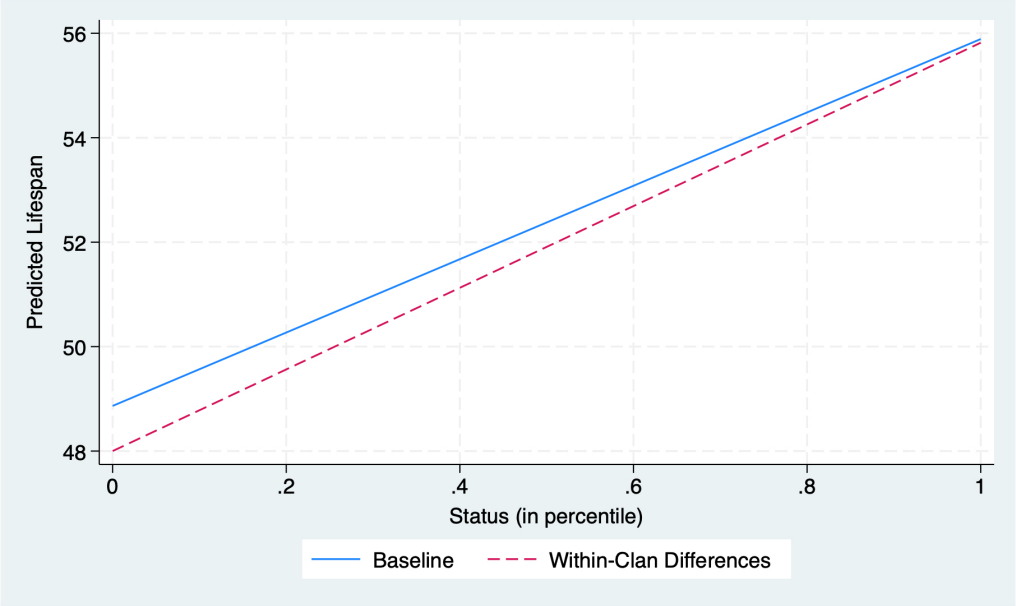
C.2 Progenitors

One might be concerned that the results are influenced by the fact that genealogical record keeping might often start when a lineage has a particularly noteworthy member, the person who in some sense starts the extended family (called the progenitor). Figure A.10 examines the linear health-status gradient that is obtained if such progenitors are dropped from the sample. We see that dropping the progenitors leads to virtually the same results. This is not surprising because there are only seven clans.

C.3 Early Birth Cohorts

A related concern is that individuals listed in the early parts of the genealogies are not be representative because their existence might have triggered the writing of the family's history in the first place. Moreover, given that the earliest birth in the genealogies is recorded for the year 1298, one might wonder what the role of data from the early centuries is for the analysis. Figure A.10 shows the linearly predicted health-status gradient for the sample that drops all birth cohorts before the year 1600. We see that results are similar to our baseline findings. A reason for this is that the genealogies do not have many records before the year 1600.

Figure A.11: Clan Differences and the Linear Health-Status Gradient



Notes: Linear fit between lifespan and status for the baseline analysis compared to the linear prediction when lifespan and status are residualized with clan fixed effects to focus on within-clan differences.All married males from seven Tongcheng clans in birth cohorts 1775 or earlier with at least 16 years of lifespan.

C.4 The Role of Clans for the Health-Status Gradient

Figure A.11 compares the linear prediction between status and lifespan in the baseline analysis with one that conditions on clan fixed effects. This is of interest for a number of reasons. One is that if results from using exclusively within-clan variation were very different from the baseline findings it would cast doubt on our approach of achieving a broadly representative sample by oversampling relatively poor clans. Figure A.11 shows evidence that the magnitude of the health-status gradient is similar whether one utilizes between-clan differences or not.