An Analysis with Five Linked Generations in China, 1300 – 1900[1]

Carol H. Shiue[2]
University of Colorado, NBER, and CEPR

November 2017

Abstract

This paper uses a multigenerational perspective to analyze social mobility over the long-run. The data covers information on about 10,000 men who lived in Anhui Province, their wives, and their children, for seven lineages of between fourteen to twenty generations. Among the more than 40,000 individuals in the sample, the earliest recorded birth is in the year 1298, and the last recorded death is in the year 1925. The results show that educational inequality is closely related to changes in mobility over time. Consistent with social and institutional trends during this period, times of greater inequality among men in the father's generation are times of less mobility in the son's generation. Heterogeneity in mobility over the distribution of status and asymmetry of mobility at the top and bottom contributes to this observed correlation. I next examine whether lineal or non-lineal effects are more important to son outcomes. The results show that although mobility is a multigenerational process, the lineal impact of grandfathers and older generations is overshadowed by non-lineal interactions coming from higher status men in the same generation as the father.

[1] Thanks to Murat Iyigun, Wolfgang Keller, as well as audiences at the CEPR Economic History Symposium, Cliometrics Society Conference, Conference on Deep Rooted Factors in Comparative Development at Brown, University of Michigan, University of Nebraska, Northwestern, the NBER Summer Institute Development of the American Economy, the Paris School of Economics Conference on China and India, and the Southwestern University of Finance and Economics for useful comments. T. Telford provided some of the data.
[2] shiue@colorado.edu

1. Introduction

Studies on intergenerational social mobility have shown that the extent to which outcomes between the parent and child are linked varies substantially. In societies with high intergenerational elasticity (IGE), a person's economic status depends to a greater extent on the status of his or her parents, while in low IGE societies, parental income or wealth matters much less. Estimated IGE's for different samples range from 0.2 to 0.6 or more (Black and Devereaux 2010, Solon 1999), and recent studies have focused on explaining these differences.[3]

One important correlate of mobility that has emerged in the recent literature is inequality. Figure 1 shows the positive relationship between inequality and intergenerational elasticity found in recent cross-country OECD data (Corak 2013), where mobility and inequality are measured at roughly the same point in time. The relationship, sometimes referred to as the "Great Gatsby Curve," shows that more equal societies, as measured by the Gini coefficient, tend to also experience greater intergenerational mobility. While this does not establish causality, the positive slope suggests that individuals in societies with more equality are more likely than their counterparts in less equal societies to experience movement in status relative to their parents. Because mobility is inherently a temporal phenomenon, however, in studying the mobility-equality relationship one would ideally like to observe how inequality in the previous generations is related to mobility in later generations. Moreover, one would want to observe that relationship repeatedly over many generations, rather than as a single cross-section.

This paper examines the determinants of intergenerational mobility over time, and focuses on the interaction between mobility and inequality. I examine multigenerational mobility patterns of households formed by around 10,800 men and 11,370 women. The sample provides consistent observations of individuals who resided in a single county in China from 1300-1900. This is a relatively homogeneous population—the extended paternal families of

[3] Regressive government policies, labor markets, heritability of endowments, regional inequality, and other factors could be important, see survey by Black and Devereaux (2010), also Becker and Tomes 1979, Björklund and Jäntti 1997, Krueger 2012, Solon 2014, Chetty, Hendren, Kline, and Saez 2014.

seven lineages who resided in the same county over this period—which should alleviate concerns that population diversity is primarily responsible for driving the results.

Furthermore, while the most often used approach centers on measuring the strength of the association in status between fathers and sons, with or without additional information on grandfathers, much less is known about the role of the extended family in mobility. If the family tends to enhance equality among its members non-lineally as well as lineally, then this will change measures of mobility. In this paper, I consider the sources of heterogeneity in the IGE over status and time, showing that mobility not only differs across the status distribution but the extended family plays a role both in preserving *and* enhancing mobility in specific ways.

The paper proceeds as follows: I first provide historical evidence documenting that from the Ming (1368-1643) to the Qing Dynasty (1644-1911), social and institutional change resulted in fewer social and economic barriers and was likely to have produced greater scope for mobility. I then construct an implied income distribution based on the biographical data in of individuals in the seven lineages based on their status percentile, and estimate the intergenerational elasticity to evaluate changes in absolute mobility and relative mobility over time.

Given the overall trends in mobility that can be observed, I ask what are the main correlates of mobility over time. I document that educational inequality in the father's generation is a central determinant of mobility, and the relationship is robust to the inclusion of a wide range of other correlates. The finding essentially documents a temporal "Great Gatsby Curve" in China over this period. The trend of decreasing inequality and increasing mobility is consistent with the history of gradual but slow moving changes leading up to the Qing Dynasty (1644-1911), in which occupational and social distinctions were weakened.

I next consider how families can affect the degree of mobility between generations. Intra-household allocation may be a way in which inequality is reduced or perpetuated among its members. In this context, I focus on a number of different aspects: 1) the intergenerational impact of grandfathers, great grandfathers, and great-great grandfathers; 2) men in the same generation within the lineage as the father, i.e. the uncles; 3) men in the same generation as the

father outside of the lineage, i.e., non-kin males; 4) marital unions, i.e., the role of the father-in-law.

A comparison of the role of these different parts of the family shows that the simplest, and most commonly assumed approach, which relate sons' status to that of their fathers, and possibly grandfathers, generally omits members who in fact add greater definition to son outcomes. Moreover, the degree of genetic and resource sharing between sons and grandparents or great grandparents is likely different from that between sons and uncles or other kin. Thus, certain family members may have more mobility preserving effects whereas other family relationships may be more mobility enhancing and tend to create greater equality within the family distribution of resources. Taking account of these members changes the estimated IGE. Senior men from other lineages, however, do not affect my results on mobility. This shows the IGE is sensitive only to the within-lineage dynamics.

Additionally, marriage has been studied as a mechanism affecting social mobility because it might preserve status-relevant group traits (Dunton and Featherman 1981, Clark 2014). The Tongcheng data provides an opportunity to examine this question because for each of the sons in the sample there is information on the lineage to which his mother belongs. I find that marriage can be either mobility enhancing or mobility preserving. High status families that marry tend to maintain their hold on status in the next generation relatively well, suggesting that marriage is an effective way of preserving status for these couples. Notably, in cases when men of lower status marry into a family where the woman is from a higher status family, there is greater mobility in the next generation. One explanation is that the marriage acted to increase equality of the two families' genetic or resource endowments, which in the following generation results in increased mobility as measured between the son and his father.

This paper contributes to the quantitative analysis of mobility from historical times to today.[4] We know very little about mobility historically as the data required for this type of

[4] Across administrative regions within the U.S. today, Chetty, Hendren, Kline, and Saez (2014) find that on average a 10-percentile increase in parent income is associated with a 3.4-percentile increase in child income, but the probability that a child in the bottom quintile will reach the top quintile could vary substantially from region to region.

analysis has so far been relatively difficult to access. An expanding literature has started to explore mobility trends in the 19th century and earlier: see Long and Ferrie (2012, 2013), Collins and Wanamaker (2015), Olivetti, Paserman, and Salisbury (2015). Clark and Cummins (2015) focus on English patterns of surname mobility, while Lindahl, Palme, Massih and Sjogren (2015) uses Swedish data. For China, Chen, Naidu, Yu, and Yuchtman (2014) study mobility in 20th century China, while Mare and Xi (2014) examine imperial and other specific populations starting in the 17th century.[5]

A key contribution of this paper is in characterizing long-run social mobility across a much broader gamut of the income distribution, in particular of those households at the middle to lower end of the distribution. With few exceptions, for periods prior to the 19th or 20th century, much more information about richer households, compared to poorer households, are available. Ho's (1962) classic study of social mobility in China, for example, focused exclusively on China's top (2-3% percentile) social status levels. But patterns of mobility at the very top of the income distribution are not representative of the rest of the households in the population. The results in this paper show not only that mobility varies over the status distribution: there is regression to the mean, but the speed of regression to the mean is different. Moreover, there is an asymmetry in which the mobility of households at the top of the distribution is relatively less sensitive to changes in inequality, compared to other households.

In addition, an important question that has been asked in this literature is whether more than two generations are needed to estimate the long-run trend in mobility. By using two-generation mobility estimates to predict higher order mobility, and comparing the predicted mobility with actual higher-generation mobility, this paper contributes to this empirical issue as well. The answer pivots on where in the income distribution we want to focus on: higher status families tend to regress much slower to the mean, so the one generation estimate would provide

[5] In particular, the findings of this paper complement work on three generations including Long and Ferrie (2012) and Olivetti, Paserman, and Salisbury (2015) on the United States, while four generations are examined in Lindahl, Palme, Massih, and Sjogren (2015), Mare and Xi (2014), and Braun and Stuhler (2015). These studies are for the more recent times, 19th century or later.

a skewed (and over) estimate of mobility. This drives home the fact again that there is heterogeneity in the determinants of social status that also changes over time.

This paper also contributes to the literature on the economic role of the family, in particular papers that have asked important questions about the contributions of the extended family (Altonji, Hayashi, and Kotlikoff 1992, Jæger 2012), and if the family is a force that tends to enhance equality, rather than contributing to the perpetuation of inequality (Griliches 1979, Bleakley and Ferrie 2016). My results demonstrate effective economic boundaries within and without the family, and suggest the larger economic role in perpetuating (or preserving) status for family members is different depending on whether one considers the horizontal cohort, or the vertical intergenerational relationships, or the relationships created through marriage.

The next section provides a discussion of the historical background, highlighting the temporal changes from that occurred in society during the sample period.

2. Social mobility in China during the sample period

By the 17th century, although there were large differences in wealth and status within Chinese society, there were few formal and institutionalized barriers that prevented commoners from rising across social classes. However, it had not always been this way. The process of change appeared to be more gradual and slow-moving, rather than a sudden shift. We observe hereditary aristocracies in China early on, but sometime over the Tang dynasty (670-906 AD), they had been eliminated. Beginning in the Song dynasty (960-1127 AD) officials of the state were selected on the basis of formal examinations. By about 1650, however, the only types of hereditary privileges and automatic status that remained belonged to the imperial lineage—where the throne was passed from the emperor to one of his sons—and the leading families of the Eight-Banner system. The latter was an exclusive hereditary institution that dominated certain military and command functions, and men born into banner families held a caste-like elite position.

The institution of using classical education to legitimize bureaucratic officials was a central aspect of governance until the last years of the Qing dynasty (1644-1911). Whereas in

Europe, nobility status could be passed down across generations, social status in China relied to a far greater degree on investments made by each generation on behalf of the next. The regular erosion of political power, at all levels below the throne, was in this way institutionalized from an early point on.

Thus, Chinese society was a meritocracy, where, at least in principle, even commoners with no family background in high status positions could gain entry into wealth, high status, and political power upon passing the state administered written examinations. At the same time, the question of the actual extent of mobility is debatable, since both genetics as well as resources are heritable, even if actual titles are not. Wealth alone would have meant better access to tutors for the sons of the rich, for example. In addition, it is possible that kinship networks and connections among the lineage in high status positions helped to extend personal advantages to other lineage members trying to stay in their high status positions or to climb the social ladder. It was not unknown for especially talented boys from families of low standing to be supported in study by a richer family member. Existing historical studies have attempted to examine the question of mobility by examining the family histories of officials.

For those who managed to become appointed to an official position upon passing the exams, the most important source of income for these officials was the state's compensation for their services (both in terms of official stipend and other remuneration). It was with this income that they were able to uphold their relatively high living standards, contribute to local community projects, and make investments in landed property.

Merchants who had accumulated fortunes could on occasion purchase minor titles and thus buy into some part of the governing elite. Degrees were offered for purchase in particular in times of revenue shortages. However, because by the Qing these degrees were distinguished by name and did not entitle the title-holder to any of the high ranked positions, participation in the state exams was the direct route and the only way to acquire the highest-level positions. Anecdotal evidence suggests that one use of merchant capital was to invest in the education of sons in order to climb the social ladder.

a. Institutional change from the 14th century to the 19th centuries

Of particular note is that despite the many traditions supporting social stratification from early on, the direction of change was towards greater social fluidity at both the upper levels of society, as well as for commoners. One indication was that merit was not just a quality inherited or a natural birthright, but it also required personal effort. Two types of inherited status categories deserve mention, however. One is the inherited titles of nobility. The second are the occupational status designations. In both, there were trends towards greater mobility and less automatic inheritance. From the Tang dynasty (670-906 AD), hereditary aristocracies in China had been eliminated. Beginning in the Song dynasty (960-1127 AD) officials of the state were selected on the basis of formal examinations. As the state examinations became a more central tool of upward mobility, the barriers to commoners from taking part in these examinations went down. This was a long process, and by the time that the sumptuary laws that prevented commoners from taking the examinations were eliminated in the latter half of the Ming, the practice had already turned into a common occurrence.

The second type of inherited status had more of a rigid character during the Mongol Yuan Dynasty, from the 13th century to 1368, when the state's requirement for labor services resulted in an institution in which certain occupational status groups, such as commoner, artisan, soldier, salt producer, miner, scholar, astrologer, and many other categories, were required to be registered and strictly segregated. The Ming Dynasty (1368-1644) carried over this practice of compulsory occupational status registration. Although there were still remnants of the older practice of having these occupations be hereditary, the family histories of some high ranking officials who came to prominence in the 15th century reveal backgrounds in the artisan status group, strongly suggesting that the hereditary nature of the categories had started to break down by the Ming (Ho, 1962). The occupational status categories appear to have lost the rigidity of the Yuan and social mobility became more fluid. From the lists that are available of the exam candidates, it is possible to see that the status categories of the family backgrounds of the highest level *jinshi* degree holders during the Ming Dynasty included also people from different special occupational status designations, such as the soldiers, army officers, horse

breeders, medical officials, official cook, and others (See Li Zhou Wang, 1746). In summary, despite the presence of different status groups in the registries of households, which can be observed during the Yuan and also the early Ming, the permanence of the groups gradually were eroded, first in practice, and then in legal terms.

The breaking down of these legal barriers to inherited titles, which in turn loosened occupational rigidities, together with the preservation of social stratification are the two essential characteristics of the social changes over the long run. From the early Ming—in the 14th century—to the start of the Qing—in the 17th century—there came to be, apparently, more mobility across groups over time. The Qing Dynasty discontinued the Ming practice of family status registration, by which time there were no effective legal barriers to social mobility due to occupational status of the family (Ho, 1962).

Anecdotal accounts point to a path of upward mobility earned through investment, sometimes over multiple generations. For example, according to a biography of Wu Chung-liang of Huizhou, he was a blacksmith who spent his nights studying, and in 1593 passed the degree that enabled him to become a magistrate. Other biographies show that a number of successful candidates came from households of prosperous Yangzi Delta silk weavers, as well the families of merchants who prospered financially and found themselves able to afford books and tutors for their sons. A few cases highlight very humble occupational beginnings: the grandfather of Shang Lo, a prime minister from the 15th century who ranked first in the national palace examinations in his cohort, earned his living through hunting and gathering firewood. Although these are only few examples, they each point to the lack of legal barriers to upward mobility all the way to the highest ranks.

The institution of using classical education to legitimize bureaucratic officials was a central aspect of governance until the last years of the Qing dynasty (1644-1911), and it was an important tool by which automatic hereditary status was eliminated. Social status in China relied on investments made by each generation on the next. The possibility for political power, at all levels below the throne and the imperial families, to erode over time was in this way institutionalized.

Given the fact that titles could not be inherited, downward mobility was a real possibility as well. The sons of a ranking official would have been part of the leisure class, and wealth and land could be inherited. But over time, the family's status would certainly have eroded if the descendants from later generations could not obtain the degrees and titles that would prevent sliding downward in the mobility ranks. Also here, anecdotal case studies have vividly illustrated the dramatic fall in household wealth among those descendants of famous officials who failed to earn any titles.

The question of the actual extent of mobility is still open to empirical analysis, however, since the lack of institutional barriers does not necessarily imply that commoners were more likely to move up in status; or that upper status sons were in danger of losing the status of the family in which they were born. Wealth alone would have meant better access to tutors for the sons of the rich, for example. In addition, networks and connections among the clan in high status positions could have helped as well. Richer lineage members could support talented boys from poor households, and the presence of lineage-supported schools would have made this type of investment feasible in practice.

The common descent group, or lineage, may have acted as an organization aiding the mobility of poorer members of the group through the sharing of resources among members.[6] Lineages that had more resources in particular could set aside common lands or lineage funds for the education of the children of poorer kin as well as the fees and expenses associated with taking the examinations. This was also a strategy that could potentially pay off for the lineage later if a member succeeded in the civil service examinations. The lineage as an organization form did not experience fundamental institutional change over the period, however, if the wealth had to be shared among a large group of people, the practice of resource sharing may have also created disincentives to accumulate wealth in the first place. In which case we would find wealthier and poorer lineages, but wealthier lineages as a group facing similar choices.

Other variables in the genealogical data can help determine the strategic factors that were used to influence mobility. Families could invest more in the education of their sons by

[6] Instead of lineage, the literature uses terms such as clan, extended family, and common descent group.

shifting spending and having fewer sons. Other demographic variables that could influence mobility could be life expectancy, health, and marriage. With respect to marriage, for example, the status of the wife's clan or the wife's father could be a factor in mobility of the sons.

3. Data Sources and Characteristics

a. The Tongcheng genealogies

The data of this paper comes from genealogies of individuals and households who lived in Tongcheng County of Anhui Province. Tongcheng County is approximately 30 miles by 60 miles, and is situated on the Yangzi River about 300 miles inland from the coast of the East China Sea. The county is about 150 miles from Nanjing, the early Ming Dynasty capital, and 650 miles from Beijing, the later Ming and Qing capital. Anhui Province was representative of the more developed and densely settled regions of China, with Tongcheng considered a centrally important economic region in the relatively developed agricultural economies of the lower Yangzi. The region was mainly a rice-producing area where the wealthiest families were typically landowning gentry (Beattie, 1979, pp. 130-131). Over the Ming and Qing Dynasties, the region gained some fame for having produced a number of the highest officials of the empire.

The dataset is created from genealogies of seven lineages of Tongcheng County (Shiue, 2017). Typically, genealogies start with the progenitor of the lineage from which all following lineage members descend (Shiue, 2016). In the Tongcheng genealogies, the lineages' progenitor is recorded usually in the 14th century, with the earliest date being the year 1298. The Tongcheng genealogies cover typically 18 consecutive generations, with a maximum of 21. The latest death recorded in the data set is 1925.[7] Generally, the coverage of genealogies at the turn to and into the 20th century becomes patchy (for example, see Harrell's 1987 analysis of lineage data from Taiwan). While my sample covers part of the Yuan (1271 to 1368) and the Ming

[7] The Tongcheng genealogies are not unique in the length of the period covered; Fei and Liu (1982), for example, examine ten lineages over the period of 1400 to 1900.

(1368 to 1644) dynasties it should be noted that the large majority of observations are for the Qing (1644 to 1911), as shown in Figure 2.[8]

One important purpose of the genealogy was to document its members. Genealogies were compiled and updated by the literate members of the lineage to aid in the ritual of ancestral worship. They were valued and kept privately in the hometown of the family in ancestral halls, providing future generations with a record of the location of graves, texts relating to grave worship, family rules of conduct, biographies of prominent members and their achievements or contributions, a record of lineage lands, and an overall history of the family.[9]

Given their purpose and method of collection, genealogies do not completely match up to census data, official population registers, and other administrative data. Census data typically record the observed population at a certain date, either at the time of registration or in retrospect. One would need repeated observations throughout the lifetime of the same individual in order to determine the highest lifetime achievement of that person. By contrast, genealogical data presents one entry per person in biographical format. When the birth and death dates are given, the achievement listed in that individual's record can be considered the highest position achievement over the lifetime of the individual. In addition, the voluntary nature of data assembly of genealogies may induce selection, the retroactive updating of the genealogy might lead to recall bias, and there may be survivor bias. Given these issues are studied in Shiue (2017, 2016), less detail is given here.

Genealogies give a window to examine questions that are hard to address otherwise, in China or elsewhere. This is particularly true when the link between generations is of central interest, as is the case here. Census data in the U.S., for example, becomes available around the year 1800. The influential study of Long and Ferrie (2013) studies occupational mobility based on two generations, in 1850 and 1880 (see also Collins and Wannamaker 2015, Olivetti, Paserman, and Salisbury 2015). Chinese genealogies approach the same issues, but the linking of across generations is more straightforward. The following section introduces the data and

[8] This is typically the case in Chinese genealogies (see Fei and Liu, 1982).
[9] Surveys of the content and scope of Chinese genealogies include Liu (1978), Telford (1986), and Shiue (2016).

provides summary statistics for the estimation sample. I will also compare this Tongcheng data with other available information on China during this time.

b. Sources of information in the data

Generally, genealogies provide information on male lineage members, their wives, and their children. The data can be inter-generationally linked because male children who survive to adulthood and marry re-appear in the genealogy as grown men.[10] The core information in the genealogies is on married men, their wives, and their children. One may also employ elements of information other than the status of a man, which might potentially matter for social mobility, such as the status of the man's wife's father, or how many brothers he had. I begin therefore by summarizing the available information on the men, women, and children separately before turning to the five-generation linked sample on which most of the results are based.

The data gives information on a total of 9,787 men, all of whom were over the age of 17. Around the year 1790, Tongcheng county had a population of approximately 1.3 million, suggesting that my sample covers about 1.5% of the Tongcheng population.[11] Table 1a provides summary statistics for the 9,787 married men in my sample. There are vital statistics such as the year and month of birth and death, with somewhat more complete data on birth than on death. The status coding, detailed in Table 1b, follows Telford (1986), which in turn draws on Chang (1956) and Ho (1962). The status measure is the highest status obtained during each man's lifetime, which ranges from 0 (no status) to 22 (highest status). Table 1a shows that the mean status level of these men is 1.65. Because this is a lifetime measure, it is not subject to the frequently encountered problem that status (or income) is measured early in the career, leading to measurement problems. The status of the father-in-law is also a lifetime measure.

[10] The potential selection arising from survival to adulthood and marriage are discussed below.

[11] I observe about 3,600 men that would be alive in the year 1790 in my sample. These men had more than 4,200 wives, and the data records more than 7,500 sons and 4,100 daughters, for a total of just under 20,000 persons. Gazetteers were local histories about a certain place. Three county-level gazetteers about Tongchong cover the period under analysis: *Tongcheng xian zhi* (1490), *Tongcheng xian zhi* (1696), *Tongcheng xuxiu xian zhi* (1827).

Furthermore, in this data set all status measures of the linked generations are lifetime measures, so that life-cycle issues do not arise.

About 70% of the sample consists of men without status. During the sample period, covering part of the Yuan and all of the Ming and Qing dynasties, high status came in no small part by holding an official government position. Generally, obtaining such a position required to successfully participate in tournament-style state examinations, given in several rounds. Holders of the highest degree, which was the gateway to the highest government positions and the highest status levels, were called *jinshi*. They account for about 0.2 percent of the sample.

Preparation for the state exams, not only passing, entailed substantial human capital investments, and Table 1b shows as well a simple indicator variable which is equal to one for individuals that are educated in the sense of making these human capital investments, and zero otherwise.[12] Some of the lower government positions were also open to men who purchased degrees, as noted in Table 1b. Intermediate levels of social status were held by wealthy landowners, farmers, and merchants (status level 8), or by students of the Imperial Academy (status level 12).

Because these 23 status levels are, for some purposes, too detailed for analyzing social mobility, I aggregate them to a classification with nine status classes. None of my main results is sensitive to plausible changes in the status coding or aggregation. Thus, unless noted otherwise, all results presented in this paper are based on the 9-level status classification. The mapping from 23 to 9 status classes is shown in Table 1b, columns 1 and 2. The table shows that in the son's generation, about seventy percent of the men has the lowest status level, and 0.2%, or 18 individuals, are *Jinshi*.

c. Summary statistics

The following are summary statistics for the five-generation linked sample. The number of men in this sample is n = 7,328; the number is lower than 9,787 because linking additional

[12] Shiue (2017) analyzes the relationship between human capital investments and fertility in China using this data.

generations implies losing observations, analogously to lagging in time series analysis. The link is based on the mens' birth vitals (month and year) and death vitals (month and year), together with their father's status. This leads to a successful match rate of 97%. I have experimented with more identifiers (e.g. lineage, generation) to establish the intergenerational link, finding that this does not affect the main results. I refer to the five generations from the perspective of the youngest generation as "Son", "Father", "Grandfather", "Great-Grandfather", and "Great-Great-Grandfather". In the father's generation of this sample, about two thirds of the men have the lowest status level, and 0.3% are *Jinshi* (see Table 2 Distribution of status). Another 1% in my sample can be considered in the group of *Juren* and *Jinshi* (see Table 2, bottom).

Because more than 70% of men have no status (level 0), the average status of men is 0.69 at the son's level (note that the status included in the top of Table 2 is based on the 9-status sample). At the father's level, status is on average equal to 0.93. The slightly higher average status at the father's level is related to the secular decline in the probability that men could obtain an official position—this is in turn explained by the fact that the Qing government did not expand official positions in proportion to the population increase. Below I will examine the role of this for my estimates in a number of ways to ensure it is not driving the key results.

The oldest five-generations-linked observation in my linked sample has 1443 as the birth year of the son. The average birth year in the sample is 1770, indicating that the bulk of the men lived during the Qing, not Ming dynasty. Live expectancy of the parents was about 57 years, and the men were typically born when their parents were about 30 years.[13] These figures suggest that the typical five-generation linked observation covers more than one hundred fifty years. Table 2 indicates that there are typically five siblings in the sample, with a higher recorded share of men. The fact that genealogies provide more information on men than on women is related to their focus on patrilineal relationships.

[13] The life expectancy is conditional on (1) surviving to childhood, (2) marrying, and (3) the year of death is recorded. Life expectancy at birth was below 40 in China for much of this period.

4. Absolute Mobility

a. Sons and fathers

A reading of the historical background during this period suggested that there may have been institutional changes that made it more likely barriers to occupational mobility were reduced. In this section, I exploit the relatively long time series of the sample to study whether there were changes in mobility over time that can be observed, based on individuals' biographical information from the genealogies. Because the appropriate mobility measure depends on normative objectives, it is useful to examine multiple measures (Fields and Ok 1999). I begin by employing transition matrices between different status levels, followed by the canonical regression of son status on father status.

The simplest approach in assessing absolute mobility is to divide the sample into poor and rich groupings and ask two questions. First, "what are the outcomes of children from poor families?", and second, "what are the outcomes for children from rich families?" Here, a man is "poor" if status equals zero, and "rich" if status is greater than zero. In the sample, the fraction of fathers without any of the special notations that provide evidence of above-normal status is 66%. Table 3 indicates that the great majority of status transitions across generations are from no-status fathers to no-status sons (top left cell).

The off-diagonal cells in this transition matrix are measures of absolute mobility. The chance for a son from a poor family to become rich is about 12% (top right cell). In contrast, the chance that a child from rich parents will be equally rich is about 58%. Notice that there are about half as many cases of upward mobility compared to downward mobility (1,064/582 is about 2).

I then assess whether the chance of moving out of the lowest status has changed over time (upward mobility). To do so, I break the sample period into 100 subsamples with roughly the same number of observations, and for each subsample I compute the probability that the son has some status ($s > 0$) while the father has none ($s = 0$). Second, I ask whether the chance that a son has the lowest status ($s = 0$) even though his father has some ($s > 0$) has changed over time (downward mobility).

The analysis shows that both upward mobility and downward mobility have changed over time. Figure 5 shows increases in absolute mobility (as defined earlier). The chance of moving up is about 17% around the year 1500 and about 19% in the year 1850. The chance of moving down has changed more in comparison, from about 28% in the year 1500 to approximately 50% in the year 1850. In sum, both upward and downward mobility have increased over time, with downward mobility exhibiting a greater change.

This distinction between rich and poor is relatively coarse, so in the next assessment I implement finer status distinctions. Table 4 gives a transition matrix where rows indicate the nine status levels of the father and columns give the same nine status levels of the son generation. The transition matrix provides information on the outcomes of children from families with different status levels, addressing specifically the question of "What are the outcomes of children from poor families?" as well as the question, "What are the outcomes for children from rich families?" Close to 88% of all sons of fathers without status (status level 0) end up at the same status level 0 (top left corner). The probability that the son of a father without status reaches status level 7 (a *juren*) is 4 in 10,000, while the chance that the son reaches *jinshi* (status level 8) status is zero in this sample (top right of Table 4). This is consistent with mobility being a gradual process of investment over multiple generations.

We also see that among fathers with status level 1, about 80% of their sons will be without status (status level 0), and of the sons of *jinshi* (status level 8), 92% will achieve a lower status level than their fathers. Furthermore, 13% of the sons of *juren* fathers (status level 7) stay at the same level, 1.3% move up to become *jinshi*, and more than 85% of the sons have lower status than their father. For the families with the intermediate status level 4, mostly students at the Imperial Academy, the chance that their son rises in status, versus the chance that he falls is about 1 in 6. These figures shed new light on status mobility, especially for the non-elites on whom we know very little, in China or anywhere else during this period.

b. Mobility Five-generations

To compare the two-generation transition matrix with that over more generations, Table 4b presents evidence on absolute mobility over five generations, that is, between the generations of the Great-Great-Grandfather (row) to Son (column). In particular, if the Great-Great-Grandfather is without status, the probability that the Son is also without status is 82% (compared to 88% for the father-son probability). The table also reveals that the chance of a *jinshi* descendant to be without status over five generations is 2.4% (lower left corner), whereas the descendent of an Imperial Academy member (Great-Great-Grandfather status level 4) to without status over five generations is about 60%. These figures provide direct evidence on long-run social mobility. In the absence of a long-run linked sample those transition probabilities have to be estimated, typically by iterating on two-generation transition probabilities (as those in Table 4). A key assumption for this to be accurate is that the transition matrix does not change over time—in the next section I will address how good the assumption is.

Additional evidence on long-run mobility comes from analyzing the evolution of status over five consecutive generations. I follow Chetty et al. (2014) and convert the status levels into ranks in the percentile status distribution. In particular, note that 72% of men in the son generation have status level equal to zero (Table 2, bottom). Each of them is assigned the percentile rank of 0.36, which is the midpoint for the status class (=[0+0.72]/2). Next come the men with status level 1, which account for about 8.5% of all men; these men are assigned the percentile rank of 0.763 (= 0.72 + [0.805-0.72]/2), and so forth. Analogous mappings into percentile ranks are performed also for the Father, Grandfather, Great-Grandfather, and Great-Great-Grandfather generations.[14] Furthermore, for present purposes I aggregate the nine status levels to three, called No Status, Moderate Status, and High Status in order to avoid a too

[14] In a later section below I will drop this mid-point procedure in favor of modeling the within-status level distribution.

thinly parsed analysis. No Status is status level 0, Moderate Status corresponds to status levels 1 to 4, and High Status corresponds to status levels 5 to 8.[15]

To begin with, consider the top of Figure 4. High-Status Great-Great-Grandfathers are roughly at the 97th percentile in the status distribution (0.968 in the figure). Given their high status, regression to the mean would imply that consecutive generations have on average lower status, and that is what Figure 4 shows: conditional on the Great-Great-Grandfather being at the 97th percentile, their sons—the Great-Grandfathers—will be on average at the 86th percentile, followed by the Grandfathers at the 83rd percentile, Fathers at the 78th percentile, and Sons at the 75th percentile. This evolution is based on n = 468 observations for which the Great-Great-Grandfather status is High. Naturally, the average is based on rather diverse status transitions across families. For example, while in some families after two generations the Grandfather is a top-level *jinshi*, in other families the Grandfather is in the group of men with No Status.

Figure 4 also shows the group of Great-Great-Grandfathers with Moderate Status (n = 3,192). While lower in status they are still in the upper half of the status distribution, and over five generations they experience on a fall in their status. With a Great-Great-Grandfather at the 72nd percentile of income, the Son four generations later will typically be at the 53rd percentile. Finally, the lower series in the figure shows the evolution of status for those families in which the Great-Great-Grandfather had No Status (n = 3,652). Being at the bottom up is the only one way in the status distribution, and Figure 4 shows that after four generations the son in the typical family has moved to the 44th percentile in the status distribution, up from the 25th percentile. It is striking that low status families did not simply stay in the bottom rungs of society forever.

The lower line in Figure 4 gives a visual depiction of how upwardly mobile non-elites were in China during the sample period, as aspect of mobility that has so far eluded existing studies. The figure also quantifies *differences* in social mobility across status levels. Perhaps most striking is the degree to which High Status families can maintain their status. Note that

[15] The transition matrix for this 3-status level analysis is shown in Table 5.

the percentile rank of the High and Moderate Status groups in the Great-Great-Grandfather generation is 0.968 and 0.718, respectively. Given both of these are above the mean, regression to the mean implies that status will fall for both on average in the following generations. This is confirmed by the figure; for example, after two generations the rank of the High group is 0.831 (down from 0.968), while the rank of the Moderate group after two generations is 0.576 (down from 0.718).

At the same time, notably, the difference between High and Moderate group does not fall by much across generations. Specifically, the advantage of the High Status group in the Great-Great-Grandfather generation is 0.25 = (0.968-0.718), whereas two generations later, in the Grandfather generation, it is 0.255 = (0.831-0.576). If there would be a common rate of mobility in society, the descendants of High Status families would, given their initial position, be expected to lose (more) percentile ranks than the descendants of Moderate Status families.

The fact that High Status Great-Great-Grandfathers can preserve their status so well—even in the fifth generation the difference is 0.23 to Moderate Status, only down 2 percentage points—indicates that (downward) mobility at the top is much lower than in other parts of society. Note that high rates of persistence at the top are not necessarily only a feature of pre-industrial Asian countries. They have been found as well by Bjorklund, Roine, and Waldenstrom (2012) for 20th century Sweden. The results in Figure 4 differ in that they apply to a larger share of the population, the top 3-4 percent as opposed to the top 0.1 percent in Bjorklund, Roine, and Waldenstrom (2012).

5. Estimation results

a. The Intergenerational Elasticity, 1300 - 1900

The simplest and perhaps most widely applied method in mobility studies is a regression of the lifetime income of the son outcome on the lifetime income of the father to estimate what is known as the intergenerational elasticity (IGE).[16]

$$\log Y_1 = \alpha + \beta \log Y_0 + \varepsilon \quad (1)$$

To allow a comparison of the mobility in this sample to existing estimates, I estimate the IGE by regressing son status on father status, employing the percentile rank as in Chetty et al. (2014). I begin by estimating the mean IGE using OLS before allowing for different degrees of social mobility below.

The OLS results are shown in Table 6. The IGE is estimated at 0.58, as shown in Column 1. On average, a 10 percentage point higher father status, e.g. from the 70th to the 80th percentile, is associated with a 5.8 percentile higher son status. Including either a trend (Column 2), or, more flexibly, one hundred time fixed effects (Column 3) does not change the IGE estimate by much. Thus, the secular decline in status during the sample period has a negligible impact on the estimated average IGE. I have also calculated the rank in the percentile status distribution separately for each quarter of the sample separately to account for a changing distribution of status over these six hundred years (Column 4), finding that this does not have a major effect on the IGE estimate. The last specification shows the results when I employ the nine status classes (0 to 8) directly, instead of the respective percentile ranks; the IGE is estimated to be quite similar (Column 5).

How do the IGE estimates of Table 6 fit into the existing literature? Most existing estimates are in the range of 0.2 to 0.6, see Solon (1999), Jantti and Jenkins (2013), and Clark (2014). In particular, Chetty et al. (2014) report an estimate based on income data of 0.34 for

[16] This is a measure of relative (not absolute) mobility because implicitly the regression framework compares the mobility of individuals with a certain status with that of other individuals.

the U.S. around the year 2000. For historical samples, Lindahl et al. (2014) and Braun and Stuhler (2016) present evidence on IGEs from education and occupation regressions starting in late 19th century Sweden and Germany, respectively, finding IGE estimates around 0.33 and 0.56, respectively. In comparison, my IGE estimate of 0.58 appears to be on the high side (relatively low levels of social mobility). However, compared to these studies, my sample period is several hundred years earlier, essentially ending when the sample period in these studies begins.[17] An exception to this is Clark (2014) whose analysis over several centuries tends to yield high IGE estimates, perhaps around 0.75, although it is not clear whether Clark's (2014) approach, based on elite surnames, can be directly comparable. Another difference is this IGE is based on a sample of 5-generations', which restricts the sample somewhat. One might be concerned that over time there were changing numbers of single men—because single and unmarried men by definition do not have long intergenerational linkages—the IGE might be different. However, adjusting for the changing numbers of single men over time (as sons) does not change the average IGE estimate.[18]

6. Mobility across the status distribution, and over time

The previous section estimated the average IGE for the entire sample period. Given the long time period covered, I obtain the IGE estimate using OLS for one hundred subsamples, the same subsamples used to estimate absolute mobility trends.

$$R_{ics} = \alpha_s + \beta_s P_{ics} + \nu_{ics} \quad (2)$$

where s=1,2,3,...100 subsamples for c $\in$ s, and c = birth cohort determined by son's birth year, c=1,2,3,...C. The rank-rank slope coefficient, β_s, is subsample specific.

[17] In addition, the extent to which educational and status mobility are comparable is unknown.

[18] Trend lines in IGE over time are very similar across samples that include or do not include single unmarried men of different ages, showing that accounting for unmarried men does not explain temporal trends.

Figure 5 shows the results of relative mobility plotted by year. It is apparent that the IGE has fallen over time, or equivalently, that relative mobility has increased. Up until the 16th century, the IGE point estimate is about 0.7 it is around 0.5 by the middle of the 19th century. The result is obtained for a given sample region so differences in measurement or definitions should not play a major role.

a. Distribution of IGE across status levels

The previous sections have shown that the evidence from regressions estimating the IGE strongly suggests the IGE differs depending on household income of status. Here, I directly estimate the distribution of IGEs across status levels, and in the next step examine mobility changes over the distribution for the first half of the sample relative to the later half.

Consider the IGE for a particular percentile p of the status distribution, denoted by IGE(p), $p \in (0,1)$. For example, IGE(0.8) would be the intergenerational elasticity at the 80th percentile of the status distribution. In principle, this can be estimated using a standard quantile regression. However, in the present case I have only nine distinct status levels, from level 0 to 8 (see Table 2, bottom). All sons with status level 0 have percentile rank 0.36, those with status 1 have percentile rank 0.763, and so forth, which limits the quantile analysis.

Because of this limitation, in the following, I model the unobserved within-status distribution to address this issue. Specifically, let *son(i,s,m)* be the percentile rank of son i with status s, where $i = 1,..,7{,}317$, and $s = 0,1,..,8$, and sample m. So far, for status level $s = 0$, so far I have employed *son(i,0,m)* = 0.36, all i, (see Table 2, bottom). Now, I now model the unknown within-status distribution by assuming the percentile rank of son i with $s = 0$ is randomly drawn on the interval 0 to 0.72 in a given sample m; this is repeated for all i with status level $s = 0$, and analogously for the other status levels. For example, the sons with status level $s = 1$ are randomly assigned percentile ranks between 0.72 and 0.805, and so forth. This random sample of percentile ranks *son(i,s,m)* of size n = 7,317 has the same status means but the percentile rank is continuous between 0 and 1. Finally, to ensure that specific assignments do not matter I repeat the procedure M=100 times and report the average across

the hundred quantile regressions to obtain the intergenerational elasticity at percentile p. An analogous approach is adopted for the percentile rank of the father.

Across the M = 100 random samples, the average OLS IGE estimate is virtually equal to that in Table 6, column 1. This is as expected because my approach does not change the means. Figure 6 shows results for IGE(p) by decile, as well as the 1^{st}, 5^{th}, 95^{th}, and 99^{th} percentile of the status distribution. The figure shows the distribution of IGE estimates across the status distribution together with the confidence interval (dashed lines). The IGE estimates vary strongly from around 0.9 at the 30^{th} decile to less than 0.05 at the 1^{st} percentile. In contrast, recall that the OLS estimate is equal to 0.58. Evidence for high levels of mobility (low IGE) is found especially at the extremes. To some extent this is not surprising. A son who is at the 99^{th} percentile of the distribution will typically have risen substantially over the status of his father, and similarly a son who ends up at the 1^{st} percentile will have typically dropped considerably relative to his father's status; both reflects relatively high levels of mobility.

Notice, however, that mobility at the top and at the bottom is not symmetric. This becomes clear from Figure 7, which shows the IGE(p) for the top and the bottom 10 percent of the distribution. I compare the IGE estimate for the 1^{st} (Poor) and the 99^{th} (Rich) percentile. It turns out that the IGE estimate for the 99^{th} percentile is almost 4 times as large as the IGE for the 1^{st} percentile. There is less mobility for the top 1% than for the bottom 1%. Relatively rich men can maintain their high status relatively well. Conversely, on the lower end of the distribution mobility is higher, both upward and downward (this analysis does not distinguish the two). Figure 7 indicates that this result is not limited to the top and bottom 1%, as it is also at the $5^{th}/95^{th}$ and $10^{th}/90^{th}$ percentiles. Overall, the analysis confirms the descriptive finding from Figure 4 above and confirms mobility differs quite strongly across social status.

b. Distribution of IGE for two sample periods

To see how mobility changed over time from higher to lower status households in the sample, I employ the same quantile approach as in the previous section except that I estimate quantile-specific IGEs separately for two subsamples, Early versus Late (before and after the

year 1786). The two periods, Early and Late, are divided so that the number of observations is roughly the same. Figure 8 shows the results. Overall, there is a higher level of mobility over time in the sense that the coefficients for the Late-period are typically below those of the Early-period.[19] The IGE coefficients estimated by OLS are 0.65 (s.e. 0.016) and 0.50 (s.e. 0.021) for the Late and Early period, respectively. The exception to this finding is mobility for the top 5% (95th percentile and higher), which is virtually unchanged from the Early to the Late sub period. Is this because downward mobility did not change for these rich men, or because the probability of moving up into these ranks did not deteriorate over time? One interpretation is that high-status families were relatively well able to maintain their status. This conjecture is based on the idea that in China as a (largely) pre-industrial society during this time, high status is more associated with maintaining high status than moving into it.

A more detailed examination of these changes by status is shown in Figures 9 and 10, which indicates that increased mobility over time varies according to status. Figure 9 shows the upper half of the status distribution. I show the average status of men from families in which the Grandfather had Moderate, High, or *jinshi* status, for two cohorts before (Great-Great-Grandfather and Great-Grandfather) and two cohorts after (Father and Son) the Grandfather achieved his status. This gives the figure the character of a mobility event study. With the age of death being above forty years and father's age at the birth of the son being about 30 in much of the sample, typically the status evolutions shown in Figure 9 cover around one hundred fifty years.

For each Grandfather status level (Moderate/High/*jinshi*), the figure shows two series, one for the earlier and one for the later half of the sample.[20] Consider first the set of Moderate Status Grandfathers at the bottom of the figure. The general pattern is inverse-V shaped. Given that I condition on the Grandfather having Moderate Status, and given that Moderate Status is higher than the status mean, *rising* status in the cohorts before and *declining* status in

[19] The figure does not report confidence intervals for readability; they are available upon request.

[20] The earlier subsample is for 1443-1785, the later for 1786-1863, by birth year of the son.

the cohorts after Grandfather is a reasonable outcome in light of a tendency to regress towards the mean.

Turning to the changes over time, the series for the earlier years is dashed while that for the later years is solid. There is virtually no difference between the earlier and the later years in the climb up to Moderate status—it is about 8 percentile ranks between Great-Great-Grandfather and Great-Grandfather, and about 7 percentile ranks more in the following generation. The figure shows that in cohorts after the Grandfather, however, the descent in status is faster, on average, in the later years. While in the earlier years the step down occurs from the 78th percentile to the 66th, in the later period the step is larger—it is to the 64th percentile rank in the status distribution (a 17% larger descent). To be sure, this difference in the average speed of downward mobility masks substantial heterogeneity, though it is exactly these effects that explain the overall increase in mobility over time that were documented above.[21] For families of High Status (i.e., those with a High-Status Grandfather), both upward and downward mobility is higher in the later than in the earlier years: the inverse-V has become sharper. Notice that the increase in downward mobility is larger than the increase in upward mobility (8%, compared to 4%, measured at two cohorts before and after).

Finally, for the subset of families where the Grandfather was a *jinshi* (n = 34), there is also a pattern of increased mobility over time.[22] However, in contrast to both Moderate and High Status levels, now the increase in mobility over time is entirely the result of greater upward mobility, not downward mobility. In the later period, men who become *jinshi* climb on average 21 percentile ranks in the status distribution (0.79 in the Great-Great-Grandfather cohort to 0.997 as *jinshi* in the Grandfather cohort), whereas in the earlier years *jinshi* men climbed only 4 percentile ranks over the previous two generations.

Overall, these transitions paint a rich and varied picture of how the overall increase in mobility in China during this period resulted from heterogeneous changes in the mobility at different status levels. For the highest status men, the *jinshi*, upward mobility substantially

[21] As a measure of dispersion, the standard deviation of Father status here is about 0.25.

[22] There is roughly the same number of *jinshi* in the earlier and later years (16 and 18, respectively).

increased while downward mobility did not: individuals were able to shield themselves from substantial downward mobility. Men with Moderate status tended to fall faster than they had before, with no corresponding increase in upward mobility. Only for men with High status is there both a substantial increase in up- and downward mobility. The overall picture is much more complex than existing conclusions in the literature, which tend to emphasize population growth and declining in opportunities for upward mobility. I find that this is not in fact true in general, and is not true in particular for the highest status men that have thus far received the bulk of study.

Figure 10 shows the patterns at the lower end of the status distribution. The patterns now are V-shaped because these families have the lowest status level in the grandfather cohort by construction, and any achievement of status in cohorts before and after them will yield the V-shaped pattern. We see that downward mobility—in the two cohorts before the No Status Grandfathers--decreased somewhat over time, while upward mobility—the Father and Son cohorts—increased. The descent into poverty is not as rapid in the later compared to the earlier years. This finding enriches the mobility picture further, because the decline in downward mobility goes against the general trend of increased social mobility over time. The reason why this decline does not determine overall mobility is that the increase in upward mobility exceeds the increase in downward mobility: Great-Great-Grandfather status is 3.9% lower in the later than in the earlier years, while Son status is 5.4% higher in the later than in the earlier years.

In summary, the previous analysis indicates that the earlier finding of increased social mobility is present at most status levels, although not to the same extent. In particular, there is no evidence that a more rapid descent into lower income or status is important for my overall finding of increased mobility over time in this period. Rather, greater possibilities for upward mobility across the spectrum while more rapid downward mobility at moderately high status levels all play a part.

7. Mobility and its correlates over time

The previous sections presented evidence that mobility changed over time, in particular, it increased from the first to the second half of the sample over most of the distribution, although less so at the top. Notably, the increases in mobility were not confined to any one class of society, but appear to have had broader impacts across moderate upper and throughout the lower half of the status distribution. The changes are consistent with anecdotal accounts suggesting a lowering of barriers of entry out of lower classes, and into moderate or upper classes.

In this section, I examine temporal factors that may be correlated with the increased social mobility observed. I regress the IGE estimate from 100 periods on measures of demographic change and parental resources, based on equal sized divisions of the data sample:

$$\beta_s = \delta_o + \delta_1 \bar{X}_s + \delta_2 SD(X_s) + \epsilon_s \quad (3)$$

where the mean and standard deviation of father's education and father's status rank are introduced, together with additional demographic variables that describe health and resources of the parents. The sample of about 7,300 observations is thus divided into one hundred equal-sized subsamples. This implies that the following results exploit variation across 100 time periods, where each variable is based on about 73 observations.

Table 7 shows the results of the IGE average regressed on a time trend (column 1), and shows it is significantly negative, which means the IGE is falls over time, consistent with increasing mobility. In Columns 2-5, the IGE is regressed on the mean of father's status, father's age at death (which captures aspects of longevity and health), the number of marriages that the father had (a measure of household resources and health), and the father's education (a 0/1 variable of whether the father had any education). The results shows that in sub periods when the mean status, resources, and education tend to be high, the IGE is high. This is consistent with the idea that the rich experience low mobility because they are relatively more capable of insulating themselves from a loss of status, or downward mobility. Notably the strongest predictor of IGE is the father's education, and it is the only variable that renders the time trend insignificant.

In Columns 7-10 in Table 7, I include the time trend and variation in the measures of status, resources, and education of the father. This shows whether, in addition to the mean, dispersion of education matters for social mobility. The results on standard deviation of father's status and father's education suggest that both matter, above and beyond the means of these variables. The demographic variables, marriage and longevity, do not have a significant role in explaining patterns IGE over time once variation in education and status is accounted for. Economically, the effect of variation in education is strongest, with a beta coefficient of 0.65, compared to the also important coefficient 0.36 for variation in status.

Figure 11 plots the Educational Inequality of the father on the Intergenerational Mobility of the son-father estimate, across n = 100 subsamples from 1300-1900. The two figures show the strong temporal relationship between educational inequality in the father's generation, and subsequent intergenerational mobility in the son's lifetime outcome. During times of greater educational inequality, there is also less mobility, and this is the most important explanation for trends in mobility over time. This relationship holds even while controlling for the secular trends in education over time. Figure 12 reveals the timing of the periods by labeling each point chronologically from 0 to 99. The negative relationship is also robust to dropping certain periods or certain points from the graph.

The result that inequality is associated with lower mobility in the time series is consistent with negative correlation of inequality and mobility across countries (Kruger 2012, Corak 2013) and across U.S. regions (Chetty et al. 2014), both for the late 20th century. These results demonstrate the overriding importance of education inequality underlying this relationship.

8. The impact of the extended family on IGE

Having estimated the average mobility for the father-son relationship, the previous sections show that over different periods, times of greater educational inequality are times of lower overall mobility. Furthermore, there is heterogeneity in mobility not only over the entire distribution of status, but also across lineages that cannot be fully explained by average status

of the lineage. In this section, I examine how the extended family affects mobility. Below, R and P are the rank status of the individual son and father, respectively:

$$R_{ic} = \alpha + \beta P_{ic} + \boldsymbol{BX_{ic}} + u_{ic} \quad (4)$$

where individual i is in cohort c. $\mathbf{X}$ is a vector that characterizes the rank of relatives identified as previous generation (e.g., grandfather, great grandfather, great great grandfather); or the rank of relatives in the father's generation (e.g. uncles) as a proxy for networks within the family; or the rank of non-kin men in the same generation as the father, and u is the error term. If $\mathbf{B}$ is insignificant, this would imply no coordination or relationship to the son, whereas $\mathbf{B} > 0$ and significant suggests cooperative action or compensation for existing differences, and $\mathbf{B} < 0$ suggests competition or reinforcement of existing differences.

a. Mobility as a multigenerational process

Table 8a shows the IGE estimates by regressing son status on father status and comparing that to the coefficient resulting from regressing son status on previous generations. That is: son on grandfather status, son on great-grandfather status, and son on great-great-grandfather status. This 5-generation linked sample gives actual estimates on mobility over longer time periods, and usefully sheds light on whether one can validly iterate the son-father coefficient for several generations in order to get accurate estimates of mobility over longer periods.

These estimates can shed light on the empirical question of whether the first-order autoregression between two generations allows us to understand the multigenerational association across higher-order associations. Using the son-father coefficient twice would yield 0.335 (0.579 x 0.579) as the IGE mobility coefficient between son and grandfather, as given in the last row—whereas the actual, directly estimated coefficient is 0.398. Notably, the percentage difference between actual and iterated mobility gets larger with every generations, so that one would make an increasingly larger error by iterating the son-father coefficient.[23]

[23] Note that: (0.398-0.335)/0.398 < (0.326-0.194)/0.326 < (0.273-0.112)/0.273.

The reason why the predicted coefficient deviates from the actual estimate over higher generational links is likely related to how intergenerational mobility varies with status in the population. In particular, the IGE varies in a hard-to-predict way over the income distribution so that generalizations about the applicability of the first-order association to higher-orders cannot reasonably be made for all populations. There are a number of important sources of heterogeneity. First, intergenerational mobility varies with status and thus mobility is conditional on where the household is in the income distribution. This can be seen comparing the average IGE coefficient, which implies that a 10-percentile difference in the status of fathers translates into a 5.8 percentile difference for sons, with the difference between a High and Moderate status Great-great-grandfather (which is produced in Figure 4). The latter shows a 25-percentile difference in status, which implies that the predicted difference in the next generation should be about 14.5. However, the actual percentile difference in the next generation is 23. The implication is that high status households tend to regress to the mean at a slower rate than moderate status households.

Second, along with status, mobility also varies sometimes quite substantially across surname lineages, which may in turn be a proxy for genetic or environmental factors, but also capture differences in education and income across families. Table 8b shows the IGE estimates by each of the seven clans that are in the sample. The IGE for clans that have a higher status is on average higher than for low average-status clans, as we would expect (around 0.8 and 0.69 for the *Ma* and *Ye* clans, respectively).[24] Also worth noting is that the average status for the clan with the lowest IGE estimate is not too different from the average status of some other clans—evidence of heterogeneity in IGE even across lineages with similar average status, thus suggesting that there is more to IGE than a mechanical relationship to income.

[24] In separate results (not shown in the table), the IGE is 0.54 for the 5-generation relationship in the *Ma* clan (a relatively educated clan), much higher than the essentially zero IGE estimate found in four other (less educated) clans, and points to higher status persistence.

b. Non-linear (Horizontal) sources of mobility

Beyond the vertical relationship of the son to father to grandfather, it is possible that horizontal relationships of people sharing the same generation as the father are important. The baseline relationship between father and son is shown in Column 1 of Table 8c. Adding grandfather status to this OLS regression yields a small positive coefficient, significant at the 10% level (Column 2). Going further back in the sequence of ancestors, Great-Grandfather as well as Great-Great-Grandfather status enter positively as well (Column 3 and 4, respectively). Notice that the point estimate of the Grandfather coefficient is not larger than that of Great-Grandfather or Great-Great-Grandfather, as one might have expected based on the likelihood of personal interaction.

I also examine the role of non-lineal relationships by adding another variable, the status of the son's uncles, "Average Rank Uncle". It is computed as the average status of all lineage members of the father generation, except the son's father. The Uncle variable enters with a highly significant positive coefficient (Column 5). Next, in Column 6, I include the maximum rank among males in the lineage in the father's generation as the variable 'maximum uncle'; it enters positively as well. Finally, it is reasonable to believe that a high-status males in the lineage might matter more so when the son's father is of relatively low status. This is captured by the variable 'max uncle to father', which is defined as the ratio of the maximum status of the males in the lineage in the father's generation relative to the father's status. We see that having an uncle whose status is much higher than the father's is associated with a higher status of the son (Column 7).

Given the high degree of autocorrelation in status across generations, it is of interest to ask which is the most important. When Grandfather is included together with either earlier lineal relationship variables (Great-Grandfather, Great-Great-Grandfather), the significance of the Grandfather coefficient weakens (Column 9). Beta coefficients are reported in brackets, and show that the standardized size of the great-grandfather coefficient is only around 10% of the father's coefficient. Columns 10 shows the analysis for the uncle effects, and shows that having a high status uncle when the son's own father is relatively unaccomplished is much more

important than the average status of the uncles. This is plausible if the son does not seek the support of all uncles equally, but rather the uncle with the highest rank relative to his own father.

Finally, I have combined both the lineal and the non-lineal status men within the lineage (Column 12). The results confirm that in addition to the father, who is the most important, it is uncles of higher status relative to the son, who is the most important in the lifetime status achievements of the son. This provides, in this comparison, evidence against a strong influence of the grandfather relative to the Uncles.

c. Intra-lineage and Inter-lineage effects

The results from the previous section suggest that within the lineage there are network effects that suggest cooperation between fathers and other males in the same lineage because sons benefit from high status of men in the lineage, especially so if their own father is relatively lower than average. This suggests cooperation within the family and that uncles promote greater mobility by enhancing equality within the lineage.

Are the same effects observed across members living in the same county who are not related by kinship? The presence of other non-lineage networks may be positive, neutral, or negative. In addition to the role of uncles, similar effects might also exist within the region, through resource pooling not only within the lineage but with other lineages. By contrast, negative, suggestive of inter-lineage competition for available resources.

To investigate these questions, I define the average status in the father's generation for 100 subsample periods so that there are similar numbers of observations in each sample, 'Average Status All'. This variable captures the overall prevalence of high status men in the father's generation across all lineages. Further, I define an analogous average across the 100 subsamples that is lineage-specific, denoted as 'Lineage Specific Status'. This captures high status achievements in the father's status for a given lineage.

I also define the variable 'Average Status Other Lineages' as the average status among men in the father's generation outside of the son's lineage, also by subsample. High-status achievements in other lineages might lower the status achievements of a particular son if the

senior members in other lineages, through either genetic or resource endowments, improve the status of their own members at the expense of other lineages. An alternative possibility is that high status achievements outside of one's own lineage might generate positive spillovers that benefit those outside the lineage.

Column 1 of Table 8d shows that the subperiod mean status does not significantly affect son status. This is in part because the specification includes a trend variable, over which the mean status does not add much. In contrast, the average lineage-specific status in the father's generation does enter with a positive sign (Column 2). Although somewhat weaker, it is roughly comparable to the effect of average uncle status in the previous Table 8d.

The key new result in Table 8e is that high status outside one's own lineage is negatively associated with status achievement of the sons in one's own lineage (Column 3). This indicates that the inter-lineage competition effect is stronger than any positive externalities that high status individuals outside of one's own lineage might generate. Finally, the positive within-lineage, and the negative inter-lineage effects, are both significant at standard levels, see Column 4. The negative competition effect is about 40% of the size of the positive within-lineage effect in terms of economic magnitudes, based on standardized regression coefficients. This suggests that competition between lineages plays a sizeable role, even if lineage pooling within lineages matter more for the status mobility of sons.

d. Marriage effects: the role of the father-in-law

While the above subsection demonstrates that there appear to be strong differences in behavior depending on whether members identify as being in the same lineage or in another lineage, marriage is of interest because two lineages can decide to join together. In this section, I examine the potential role of the in-law family on social mobility. Sons cannot pick their father or their lineage, but they might be able to influence their upward mobility by marrying a woman with a high-status father. Traditionally, women who marry into another family would also be considered as having broken off from their natal family, but in practice, cooperation could have been a better strategy.

It is not also unusual for marriage decisions to be influenced by the size of the dowry or the status of the in-law family. While the status of the wife is not observed in my data, many traits that affect status, such as IQ and resources are likely positively correlated with the woman's father status, and a positive correlation between the status of a son and the status of his father-in-law would be evidence for positive assortative matching in the marriage market.

I have information on the status of the father-in-law for about 60% of the sample. There is no major difference in the relationship between father and son status in this subsample compared to the full sample, see columns 1 and 2 in Table 8e. Next, when I include the status of the father-in-law in the regression it enters with a positive sign (column 3). The status of a son is increasing in the status of his father-in-law. At the same time, the coefficient on the father status variable falls. This suggests that there is no independent additional status provided by the father-in-law. Indeed, the predicted status of the son is similar in columns 2 and 3, i.e., whether or not father-in-law status is included in the regression. For example, if the father's status is the 88^{th} percentile, the predicted son rank status is the 74^{th} percentile according to both columns 2 and 3. Across all status levels, the correlation of predicted son status is 0.96.

Adding the interaction between *Rank Father* and *Rank FatherInLaw* lowers the size of the linear coefficients substantially (column 4). Dividing the sample into father-in-laws with no versus with some status yields the results shown in columns 5 and 6. With an IGE of 0.46 for father-in-laws without status and 0.86 for those with at least some status, there is evidence that marrying into a family with status is associated with higher status persistence.

Finally, the last column I return to the question whether there is a complementarity between father and father-in-law rank for the status of the son. We see that conditional on a son marrying a wife whose father has status, mobility is not affected by the particular level of the father-in-law's rank in the distribution. This is evidence that high-status fathers tend to have son's who marry wives from high-status families. The evidence for direct in-law effects beyond this sorting is limited.

As noted above, all men of the Tongcheng sample are members of one of seven lineages (see Table C). The women that these men have married over the centuries are from a far larger

set of lineages, namely 180 different lineages. If a son's mother is member of the same lineage as his father, both parents share information from the progenitor of the lineage, and in this sense their son is the product of intermarriage. Among these 180 lineages are also the seven to which all of the men belong to, giving me the opportunity to examine the role of intermarriage for social mobility.

To do so, I distinguish the two lineages that tend to have relatively high status levels among the men, the Ma and the Ye. Therefore, I create a new variable to indicate if a man's mother is a member of the Ma or the Ye. If she is, she is considered a High Status Lineage Mother, which is denoted by *Mother High Stat* equals to one, and zero otherwise. Observations for which *Mother High Stat* takes on the value zero are the five other lineages of the men in the sample, as well as the other 173 lineages from which the mothers come.

Results are shown in Table 8f. I begin by adding the interaction *Rank Father x Mother High Stat* together with the *Mother High Stat* indicator variable. As seen from column 1, the interaction is not significant. I find the same result when the sample is restricted to the seven male lineages (n = 705; Column 2). It may not be surprising that having a high-status mother has no clear association with mobility because the extent to which intermarriage affects mobility should depend on father status. I examine these issues first by defining an indicator variable equal to one if both father and mother are members of a high-status lineage, namely the Ma and the Ye, and zero otherwise; denote this variable by *Both Parents High*. I also define *Both Parents Low* as an indicator variable for the case when father and mother are both from lineages other than the Ma and the Ye.

As shown in Columns 3 and 4, including interactions of these indicators with father status yield coefficients with different signs. In particular, when both parents come from high-status lineages, the IGE estimate is relatively high while when they both come from relatively low-status lineages it is relatively low (Columns 3 and 4, respectively). These results hold also when both indicators are included at the same time (Column 5).

Next I separate the fathers from low-status lineages and those from high-status lineages. The hypothesis is that when the father has high status, a high-status mother locks in high

status, thereby resulting in lower mobility. In contrast, when the father is not of high status, marrying a high-status mother yields higher mobility because the son might move up even though that would not have been expected based on his father's status.

The results show support for this hypothesis. For sons from a marriage of a father from a lower-status lineage with a high-status mother, the marginal effect of *Rank Father* is estimated at 0.2 (= 0.49 – 0.29). This contrasts sharply with the marginal effect of *Rank Father* for sons from a high-status father with a high-status mother, which it is 0.92 (= 0.61 + 0.31). Clearly, status mobility for intermarriages between high-status parents is much lower than for other marriages. Overall the results indicate that intermarriage between similar status groups might be a powerful mechanism that reduces social mobility.

9. Conclusions

Employing genealogical data on the individual status of thousands of Chinese men over many successive generations, this paper examines multi-generational mobility during the years 1300 to 1900. I document a pattern of higher mobility with time, which is consistent with social and institutional change that occurred during this period. The results also show a great deal of heterogeneity in mobility at all levels of status, where educational inequality is the most important factor related to social mobility. Periods of less inequality in the father's generation correspond to periods of greater mobility in the son generation. This provides temporal evidence that echoes the cross-sectional evidence in other studies showing a positive relationship between the IGE and inequality across regions.

The reason why periods of higher inequality are significantly and positively correlated with the IGE may be in part explained by the pattern of mean regression. Regression to the mean implies that mobility is fastest at the bottom (where the only direction is upwards), and the top strata. However, empirically, when inequality increases, there is a lack of symmetry: mobility at the top is largely unaffected (IGE unchanged), while the bottom rungs become relatively less mobile because the poor are less able to climb out of poverty–together these dynamics mean that the IGE on average increases.

In addition, analysis of a 5-generation linked sample shows that iterating a two-generation estimate would overestimate multi-generational mobility relative to what is actually found in the data. Because gains from the past are built up slowly, and show more persistence than the two-generation IGE estimate implies, inequalities from the past have more pronounced impacts on mobility than we would otherwise expect from the two-generation results.

Even though the actual persistence of past generations is more pronounced, however, I find that the impact of the grandfather and older generations on the status of the current generation is in fact relatively small. Other aspects of the family show families function in cooperative, or mobility enhancing ways. In particular, men in the lineage in the same generation as the father, but who have higher status than the father, influence the status of the son positively. This suggests that the son is being pulled up in his rank through wider family contributions. Notably, this positive relationship is confined to family members. Similar effects are not observed with members in the same generation as the father but not in the same lineage, consistent with competition and the reinforcement of differences. This suggests senior men might promote greater equality within the extended family, but that does not extend to those outside the family. Also, the estimated magnitudes are greater than the impact from lineal members: in this sense, even though mobility is a multigenerational process, the lineal impact is overshadowed by network effects of those in the same generation as the father. The dynamics of non-lineal interactions of the family is an area where more research is warranted.

10. References

Altonji Joseph, Fumio Hayashi, and Laurence J. Kotlikoff (1992), "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data." The American Economic Review, 82:5 1177-1198.

Beattie, Hillary (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.

Becker Gary S., Tomes Nigel (1986), "Human Capital and the Rise and Fall of Families." Journal of Labor Economics 4(3):S1–39

Becker Gary S., Tomes Nigel (1979), "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." Journal of Political Economy 87(6):1153–89.

Bjorklund, A., and M. Jantti (1997), "Intergenerational income mobility in Sweden compared to the United States", American Economic Review 87: 1009-1018.

Bjorklund, A., J. Roine, and D. Waldenstrom (2012), "Intergenerational Top Income Mobility in Sweden: Capitalist Dynasties in the Land of Opportunity?", Journal of Public Economics 96: 474-84.

Black Sandra, Devereux Paul (2010), "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics, Volume 4B*, ed. Ashenfelter O., Card D., 1487–1541. Amsterdam: Elsevier.

Bleakley, Hoyt and Joseph Ferrie (2016), "Shocking Behavior: Random Wealth in Antebellum Georgia and Human Capital Across Generations." The Quarterly Journal of Economics, 131(3): 1455-1495.

Braun, Sebastian, and Jan Stuhler (2015), "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany", mimeo.

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

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014), "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", Quarterly Journal of Economics 129, pp. 2633-2679.

Clark, Greg (2014), *The Son Also Rises. Surnames and the History of Social Mobility.* Princeton University Press.

Clark, Greg, and Neil Cummins (2015), Intergenerational Wealth Mobility in England, 1858-2012: Surnames and Social Mobility", Economic Journal.

Collins, William, and Marianne Wannamaker (2015), "Up from Slavery", paper presented at University of Colorado, Boulder.

Corak, M. (2013), "Income Inequality, Equality of Opportunity, and Intergenerational Mobility", IZA DP 7520.

Daqing luli 大清律例 (The Great Qing Code), 1740 edition.

Daqing huidian shili 大清会典事例 (1899 edition) Ch. 370.

Dunton, Nancy & Featherman, David L. (1983). "Social Mobility through Marriage and Careers," in Janet T. Spence (ed.), *Achievement and Achievement Motives*, San Francisco: Freeman.

Durand, John D. (1977), "Historical Estimates of World Population: An Evaluation." *Population and Development Review* 3(3): 253-296.

Elman, Benjamin (1991), "Political, Social, and Cultural Reproduction via Civil Service Examinations in Late Imperial China", *Journal of Asian Studies* 50(1): 7-28.

Elman, Benjamin (1994), "Changes in Civil Service Examinations from the Ming to the Ch'ing Dynasties," in Benjamin Elman and Alexander Woodside, eds., *Education and Society in Late Imperial China, 1600-1900.* Berkeley: University of California Press.

Elman, Benjamin (2002), "The Social Roles of Literati in Early to Mid-Ch'ing." Part one: The Ch'ing Empire to 1800. Ed. Willard J. Peterson. Cambridge University Press.

Fei, John C. H. , and Ts'ui-jung Liu (1982), "The Growth and Decline of China Family Clans", *Journal of Interdisciplinary History* XII:3 (Winter): 375-408.

Fields, G. S. & Ok, E. A. (1999), "The measurement of income mobility: An introduction to the literature" [Electronic version]. In J. Silber (Ed.) *Handbook on income inequality measurement* (pp. 557-596). Norwell, MA: Kluwer Academic Publishers.

Griliches, Zvi (1979), "Sibling Models and data in Economics: Beginnings of a Survey." Journal of Political Economy, 87(5) Part 2: Education and Income Distribution, S37-S64.

Harrell, Stevan (1985), "The Rich Get Children: Segmentation, Stratification, and Population in Three Chekiang Lineages, 1550-1850." In S. Hanley and A.P. Wolf, eds., *Family and Population in East Asian History,* Stanford: Stanford University Press.

Hertz, Tom (2008), "A Group-Specific Measure of Intergenerational Persistence", Economics Letters 100, September, pp. 415-417.

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

Ho, Ping-ti (1959), *Studies on the Population of China, 1368-1953.* Cambridge: Harvard University Press.

Jæger, Mads Meier (2012), "The Extended Family and Children's Educational Success." American Sociological Review, 77(6).

Jantti, M., and S. P. Jenkins (2013), "Income Mobility", IZA DP No. 7730, November.

Krueger, A. B. (2012), "The Rise and Consequences of Inequality in the United States", speech at Center for American Progress, January 12.

Lee, C.I., and G. Solon (2009), "Trends in Intergenerational Income Mobility", Review of Economics and Statistics.

Leung, Angela Ki Che (1994), "Elementary Education in the Lower Yangtze Region in the Seventeenth and Eighteenth Centuries," in Alexander Woodside and Benjamin A. Elman eds., *Education and Society in Late Imperial China.* Berkeley: University of California.

Li, Zhou Wang 李周望, *Guochao like timing beilu chuji* 国朝历科题名碑录初集. 1746 edition.

Lindahl, Mikael, Marten Palme, Sofia Sandgren Massih, and Anna Sjogren (2015), "The Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations", Journal of Human Resources 50: 11-33.

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.

Liu, Ts'ui-jung (1980), "The Demographic Dynamics of Some Clans in the Lower Yangtze Area, c. 1400-1900", paper presented at the International Conference on Sinology, Taipei.

Long, Jason, and Joseph Ferrie (2013), "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *American Economic Review*, 103(4): 1109-37.

Long, Jason, and Joseph Ferrie (2012), "Grandfathers Matter(ed): Occupational Mobility across Three Generations in the U.S. and Britain, 1850-1910, mimeo, Northwestern University.

Mare, Robert D. (2014), "Multigenerational aspects of social stratification: Issues for further research", Research in Social Stratification and Mobility 35, March, pp. 121-128.

Mare, Robert D. , and Xi Song (2014), "Social Mobility in Multiple Generations", UCLA Working Paper, September.

Olivetti, C., M. D. Paserman, and L. Salisbury (2014), "Intergenerational Mobility Across Three Generations in the 19th Century: Evidence from the US Census", mimeo, February.

Shiue, Carol H. (2017), "Human Capital and Fertility in Chinese Clans Before Modern Growth", forthcoming, *Journal of Economic Growth*.

Shiue, Carol H. (2016), "A Culture of Kinship: Chinese Genealogies as a Source for Research in Demographic Economics", *Journal of Demographic Economics*, December 2016, 82(4), pp. 459-482.

Solon, Gary (2015), "What do we know so far about multigenerational mobility?", NBER # 21053.

Solon, Gary (2014), "Theoretical Models of Inequality Transmission across Multiple Generations", *Research in Social Stratification and Mobility* 35, March, pp. 13-18.

Solon, Gary (2002), "Cross-Country Differences in Intergenerational Earnings Mobility", *Journal of Economic Perspectives* 16(3): Summer, pp. 59-66.

Solon, Gary (1999), "Intergenerational Mobility in the Labor Market", Chapter 29 in Orley C. Ashenfelter and David Card (eds.), Handbook of Labor Economics, vol. 3A, Amsterdam, North-Holland, 1999.

Solon, Gary (1992), "Intergenerational Income Mobility in the United States", *American Economic Review* 82, pp. 393-408.

Stuhler, Jan (2014), "Mobility Across Multiple Generations: The Iterated Regression Fallacy," mimeo, UC3M.

Telford, Ted A. (1995), "Fertility and Population Growth in the Lineages of Tongcheng County, 1520-1661." *Chinese Historical Microdemography,* ed. S. Harrell. Berkeley: University of California Press.

Telford, Ted A. (1990), "Patching the holes in Chinese genealogies." *Late Imperial China.* 11(2): 116- 35.

Telford, Ted A. (1986), "A survey of social demographic data in Chinese genealogies." *Late Imperial China* 7: 118-48.

Twitchett, Denis and Frederick W. Mote (1998), *The Cambridge History of China* (eds.), Vol. 8. New York: Cambridge University Press.

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

Figure 1. Inequality and intergenerational elasticity (household income for 1985 and child's adult-outcomes in mid to late 1990s)

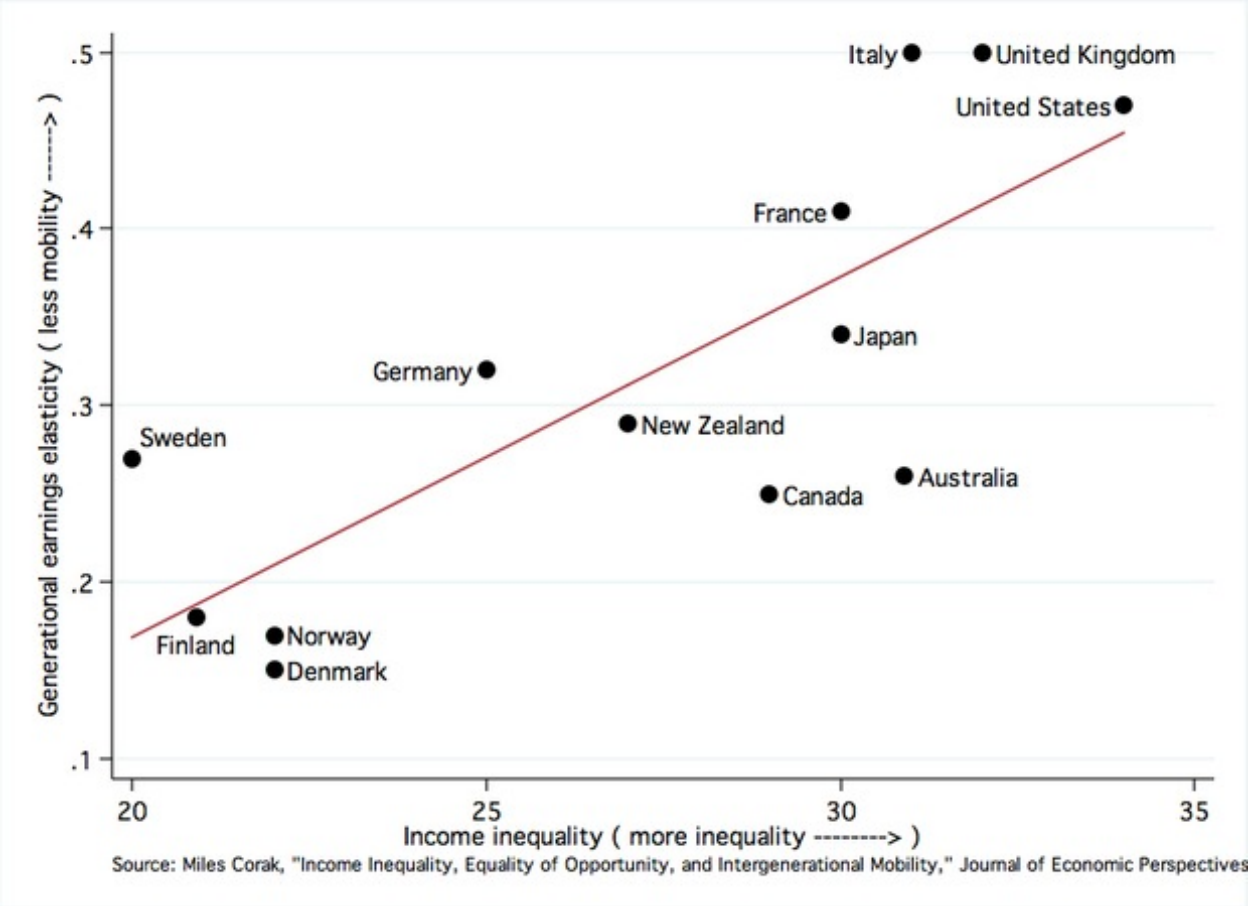


Figure 2. Frequency of birth year

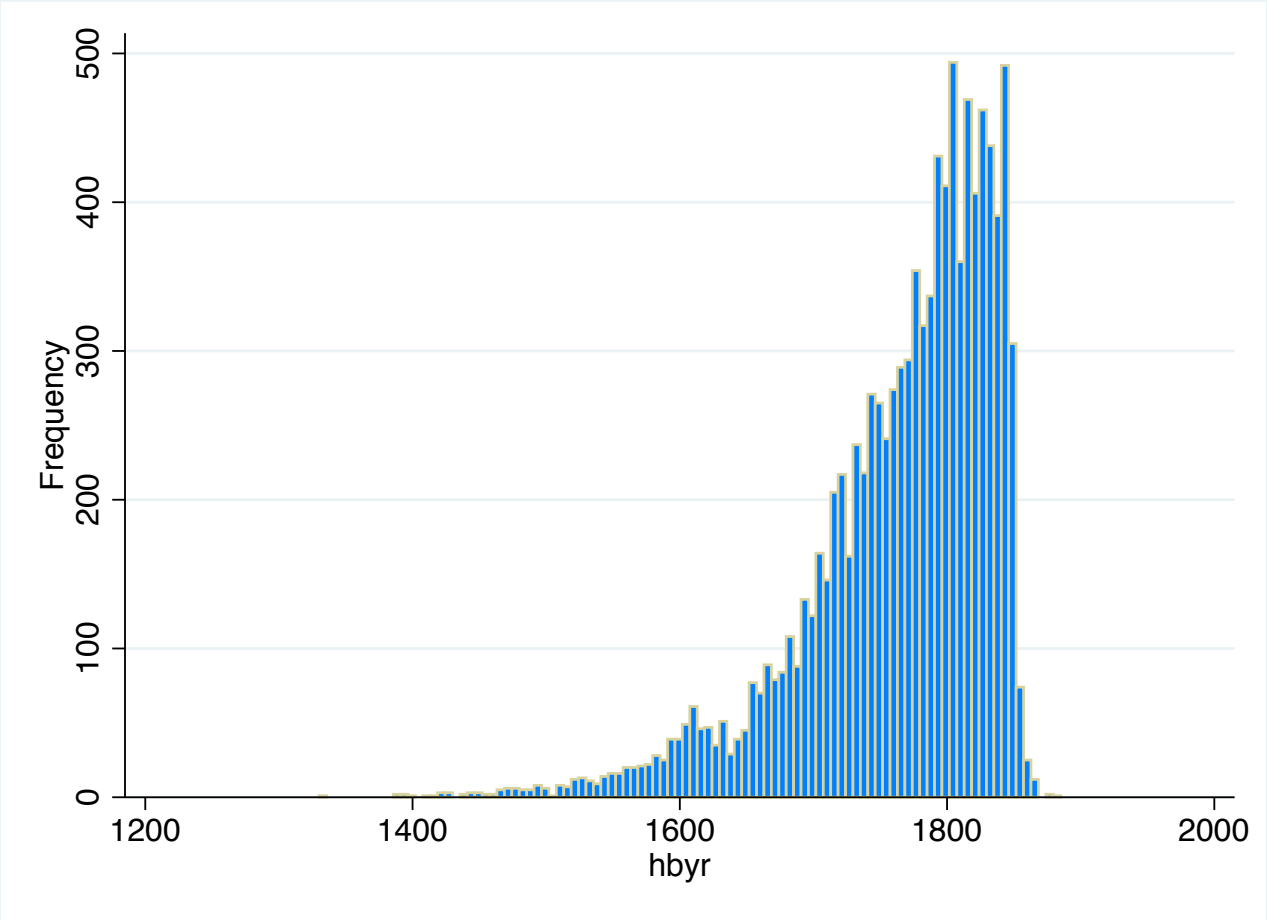
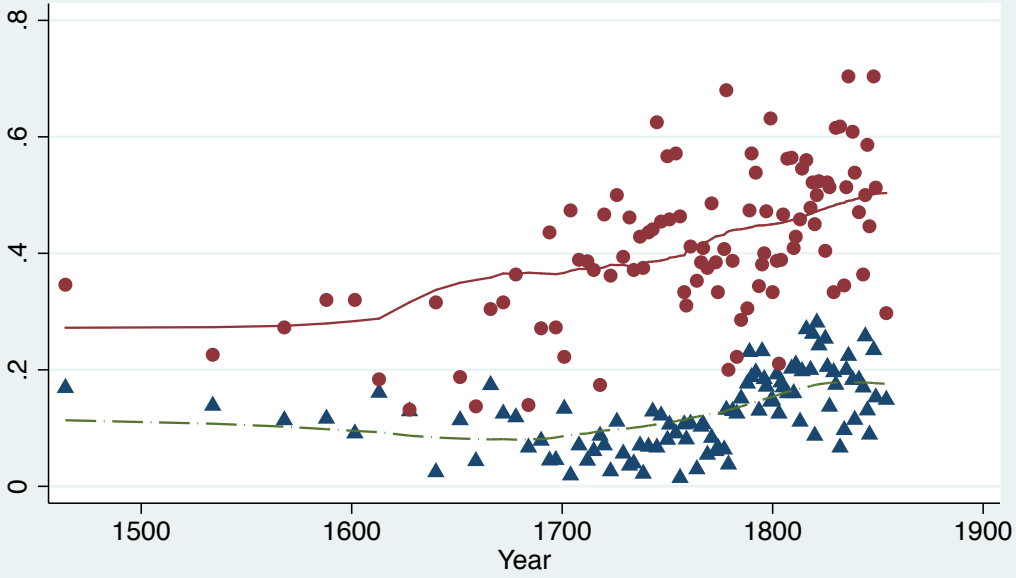


Figure 3. Upward and Downward Mobility over Time



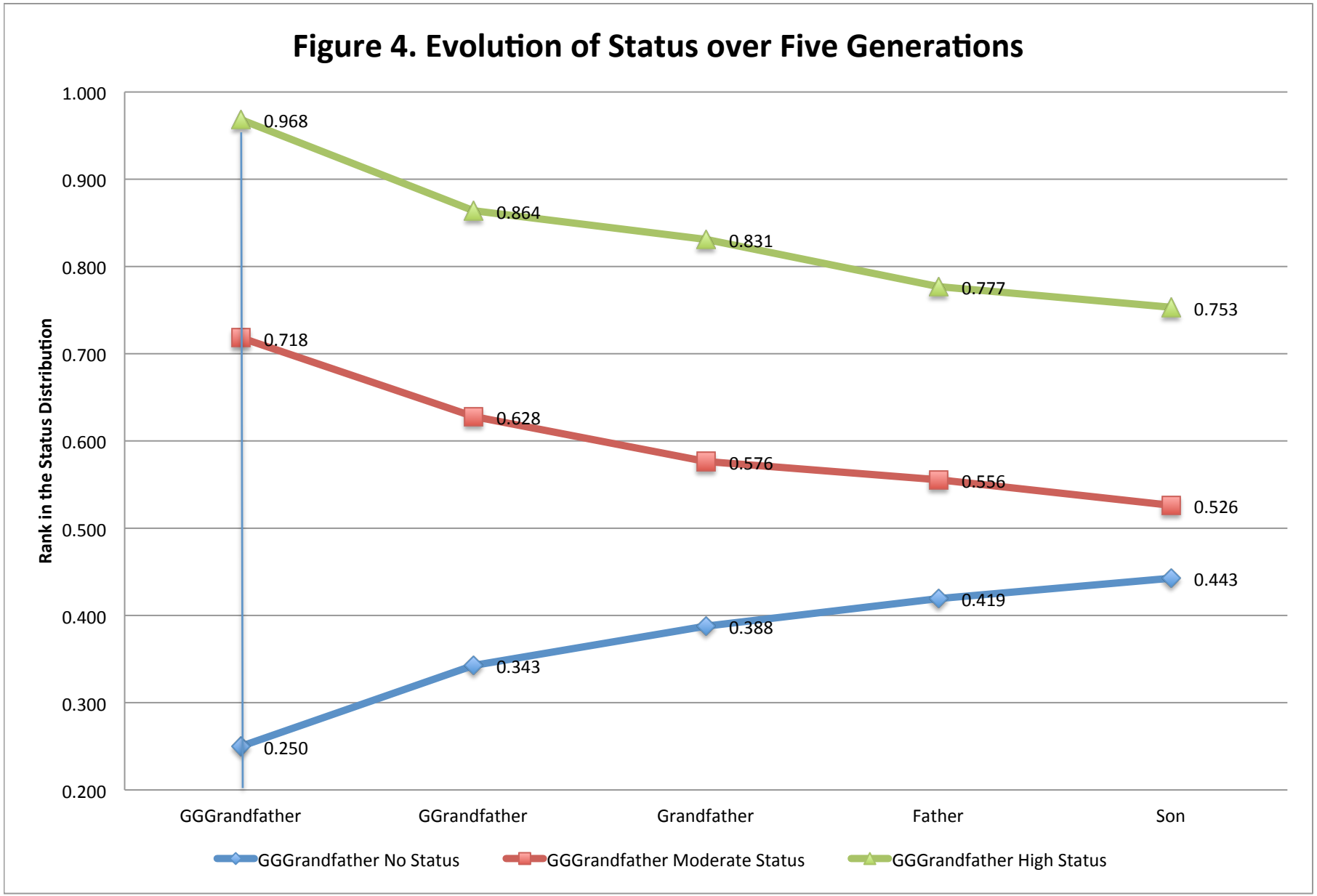
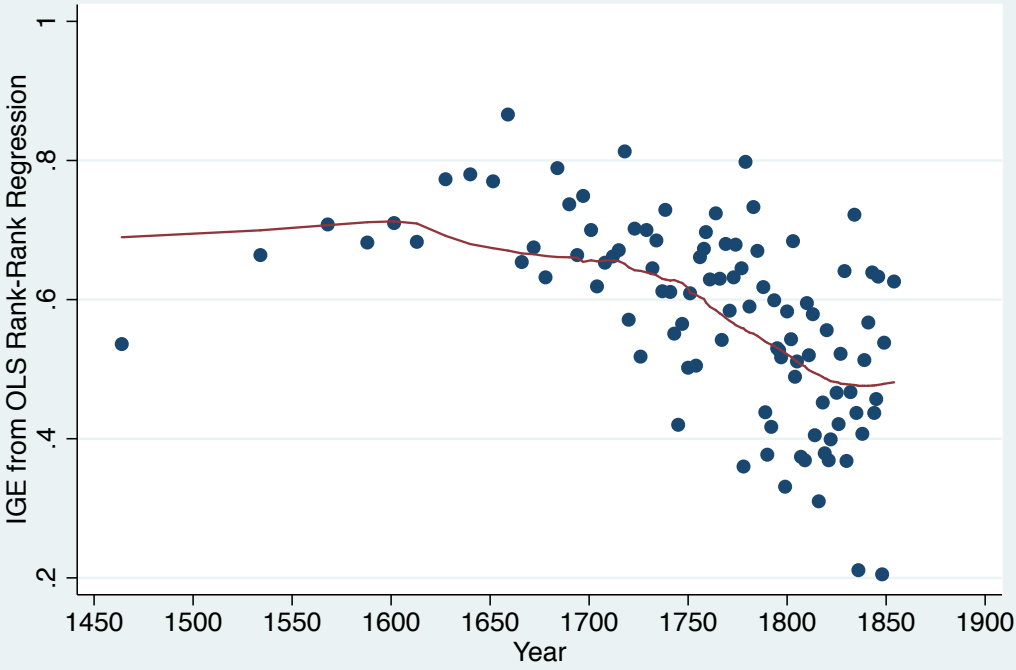


Figure 4. Evolution of Status over Five Generations
Rank in the Status Distribution
1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.968
0.864
0.831
0.777
0.753
0.718
0.628
0.576
0.556
0.526
0.443
0.419
0.388
0.343
0.250
GGGrandfather
GGrandfather
Grandfather
Father
Son
GGGrandfather No Status
GGGrandfather Moderate Status
GGGrandfather High Status

Figure 5. Relative Mobility over Time



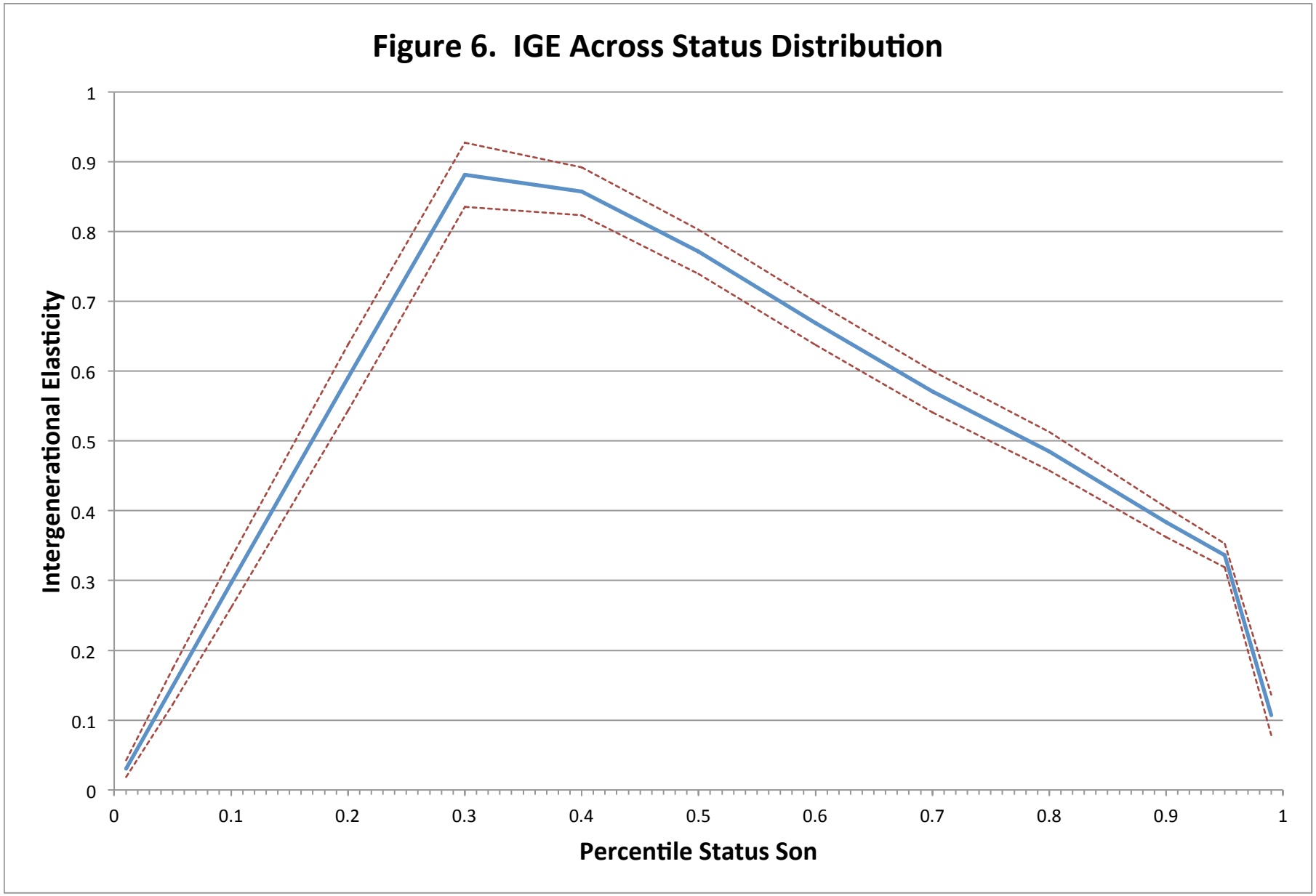


Figure 6. IGE Across Status Distribution
Intergenerational Elasticity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Percentile Status Son

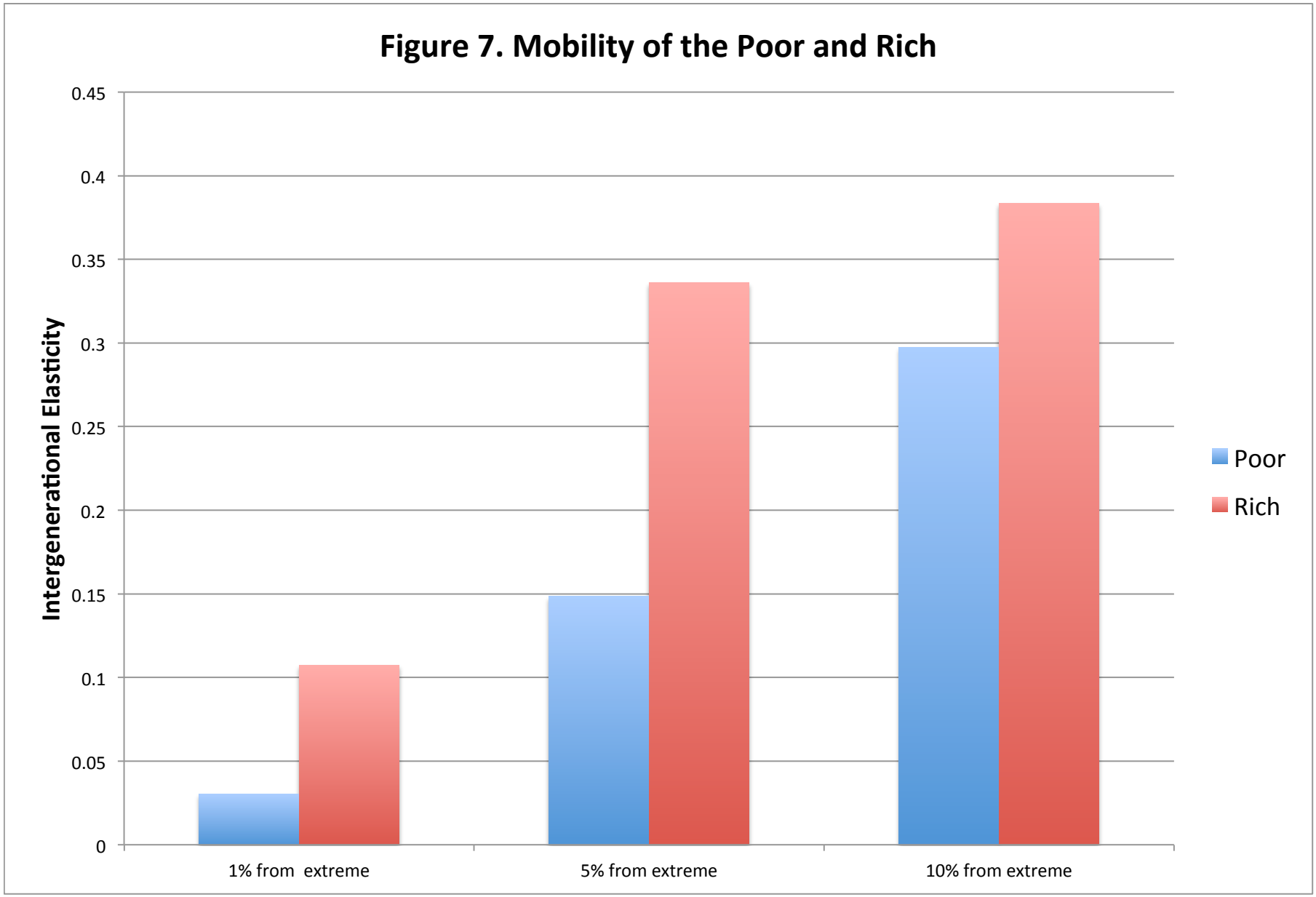


Figure 7. Mobility of the Poor and Rich
Intergenerational Elasticity
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1% from extreme
5% from extreme
10% from extreme
Poor
Rich

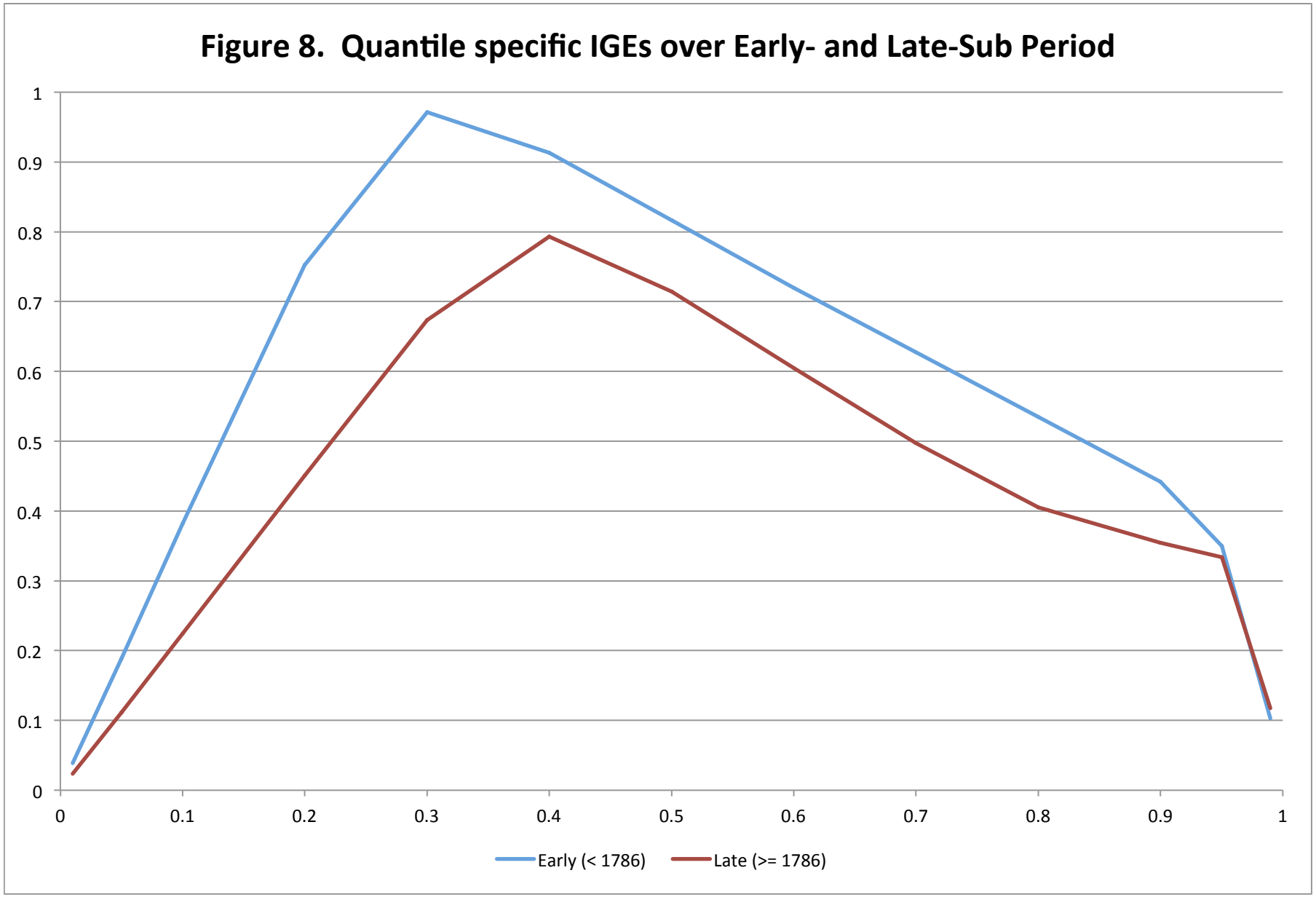


Figure 8. Quantile specific IGEs over Early- and Late-Sub Period
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Early (< 1786)
Late (>= 1786)

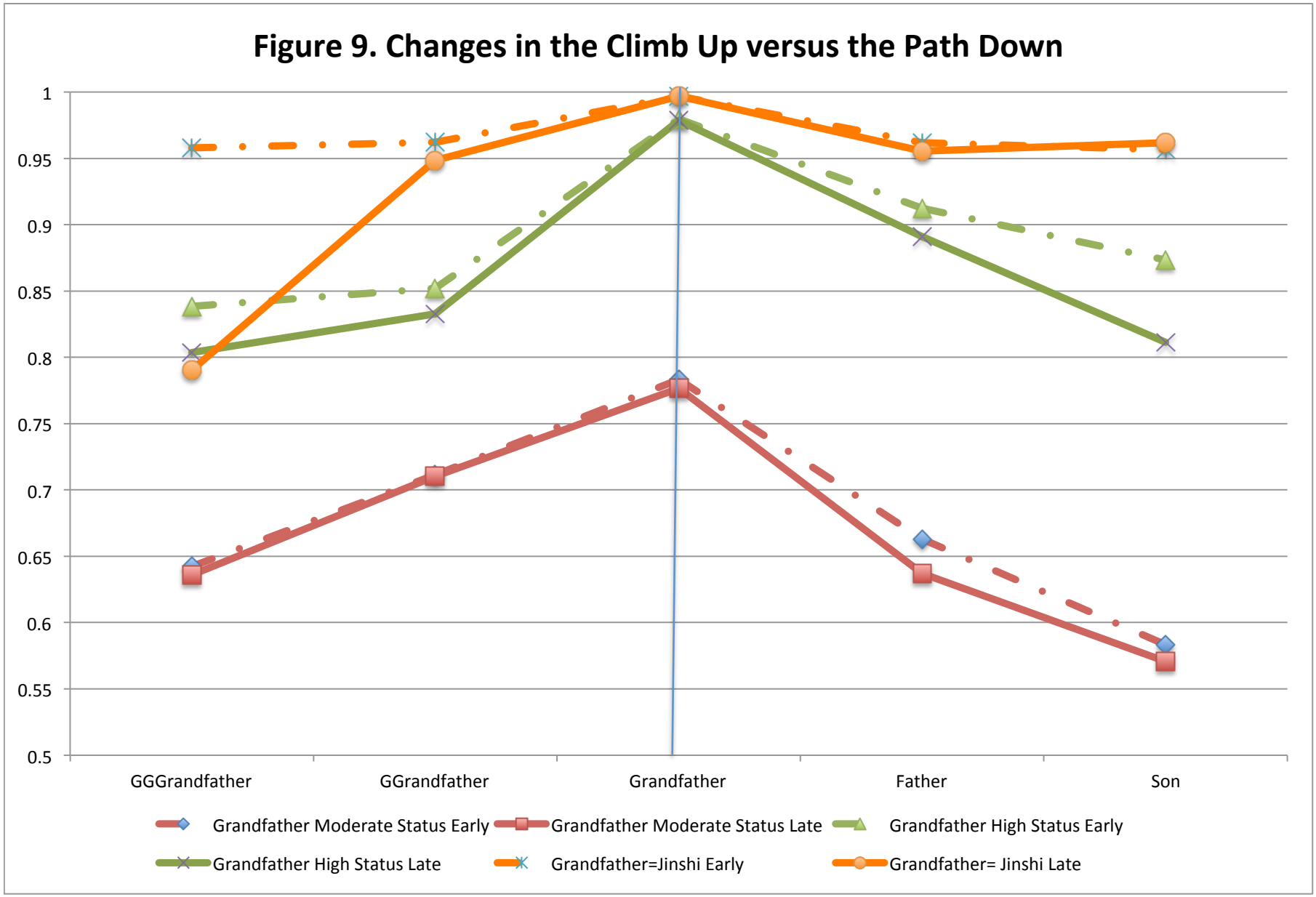


Figure 9. Changes in the Climb Up versus the Path Down
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
GGGrandfather
GGrandfather
Grandfather
Father
Son
Grandfather Moderate Status Early
Grandfather Moderate Status Late
Grandfather High Status Early
Grandfather High Status Late
Grandfather=Jinshi Early
Grandfather= Jinshi Late

Figure 10. Changes in Mobility over Time for Low-Status Families

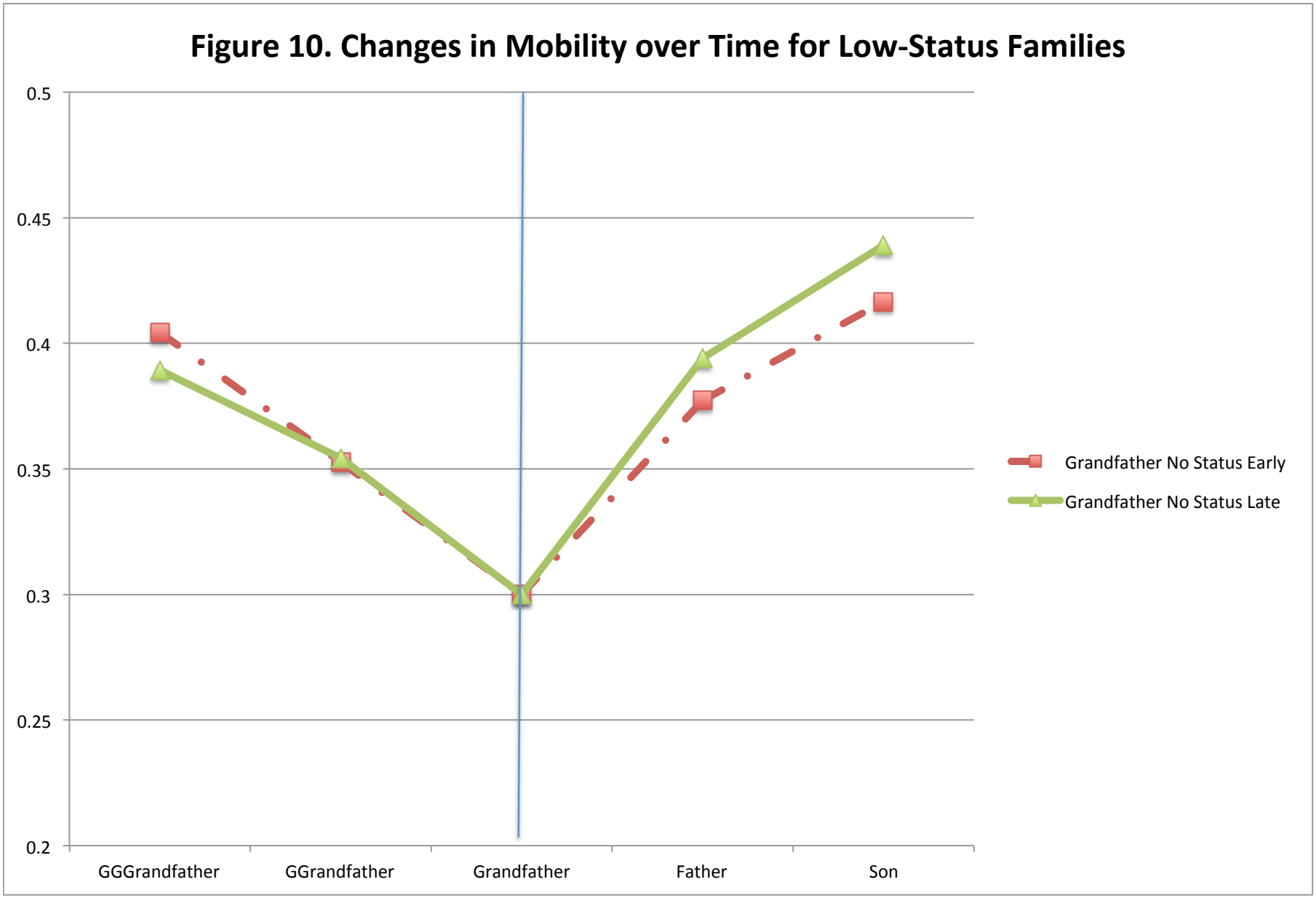


Figure 11.Educational Inequality and Intergenerational Mobility

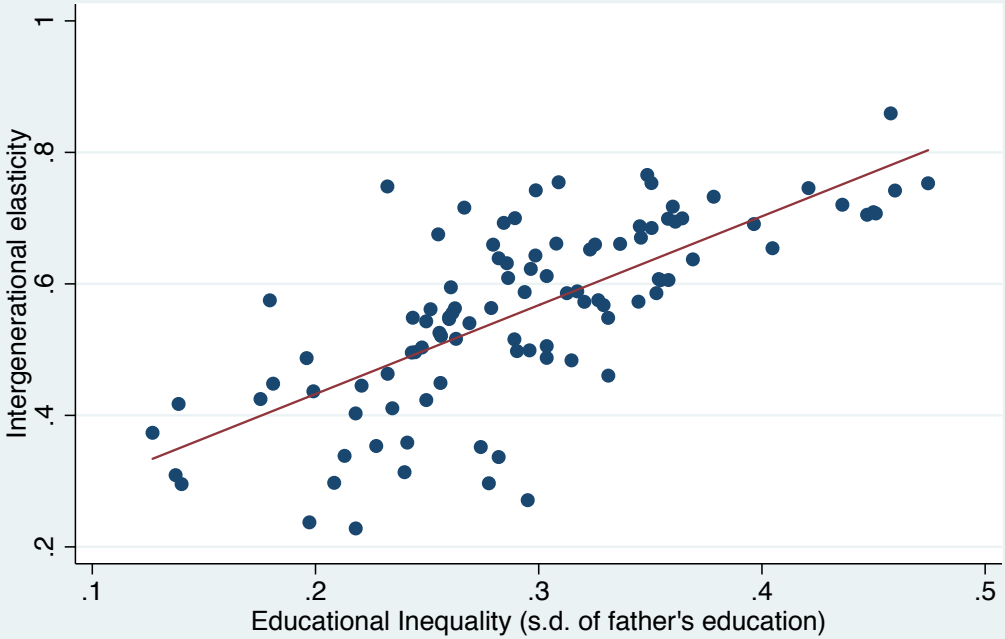


Figure 12. Educational Inequality and Intergenerational Mobility

Across n = 100 subperiods from 1300 to 1900

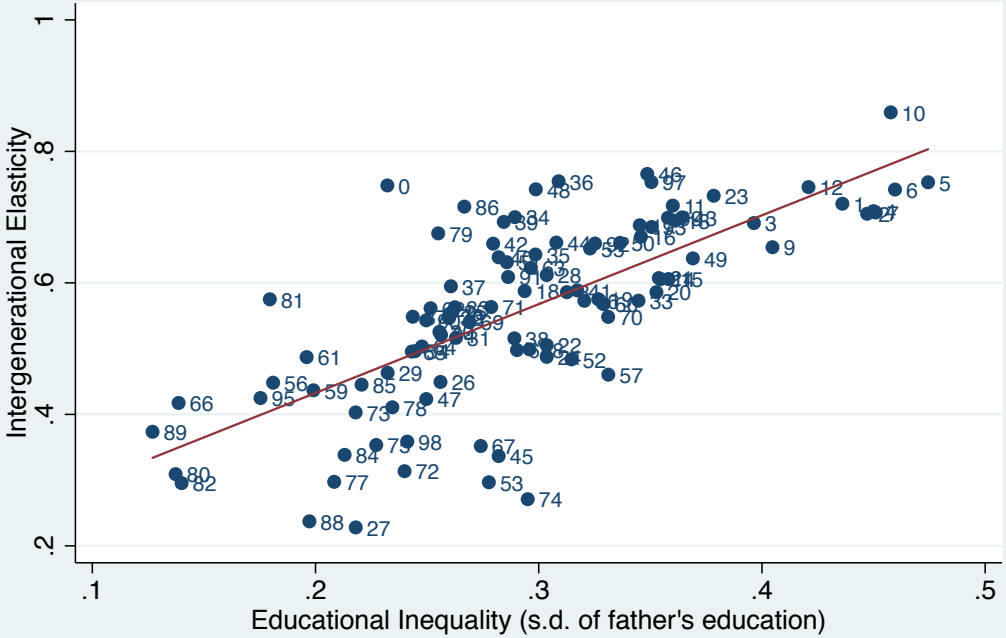


Table 1a. Characteristics of men and their wives

	Variable	Obs	Mean	Std.	Min	Max
Men						
	Demographics					
	Birth year	9,787	1760.55	74.69	1298	1885
	Birth month	9,787	6.90	3.51	1	12
	Death year	8,142	1796.64	72.51	1348	1929
	Death month	8,142	6.53	3.41	1	12
	Wealth					
	Status	9,786	1.65	3.67	0	22
	No. of wives	9,783	1.17	0.44	1	5
	Education					
	Education	9,787	0.07	0.25	0	1

	Variable	Obs	Mean	Std.	Min	Max
Women						
	Demographics					
	Birth year	11,378	1766.24	75.31	1300	1887
	Birth month	11,378	6.75	3.47	1	12
	Death year	8,708	1797.10	73.66	1355	1930
	Death month	8,708	6.54	3.39	1	12
	Status					
	Father's status	6,179	1.86	4.74	0	22

Table 1b. Status: Description and Coding

Status 9	Status 23	# of obs	Percent	Description	Education
0	0	6,956	71.08	No titles, degrees, office, other evidence of wealth	0
1	1	36	0.37	Honorary or posthumous titles; main guest at the county banquet; village head	0
1	2	786	8.03	Multiple wives in consecutive marriage, two or more not living at the same time	0
2	3	912	9.32	Father a *sheng-yuan*, minor official, or official student; evidence of wealth, *jian-sheng*, expectant official	0
2	4	23	0.24	Grandfather a *juren*, *gongsheng*, *jinshi*, or official	0
2	5	38	0.39	Father a *juren*, *gongsheng*, *jinshi*, or official	0
3	6	163	1.67	Educated, scholar, no degrees or office; editor of genealogy, refused office, or prepared for but did not pass exam	1
3	7	89	0.91	Concubinage (i.e. polygyny, two or more wives or concubines at the same time	0
3	8	11	0.11	Substantial evidence of wealth and property; set up ancestral estates, large donations, philanthropy; wealthy farmer, landowner, or merchant	0
3	9	193	1.97	Official students	1
4	10	1	0.01	Military *sheng-yuan*, minor military office	0
4	11	146	1.49	Purchased *jian-sheng* and/or purchased office	0
4	12	99	1.01	Students of the Imperial Academy (non-purchased)	1
5	13	53	0.54	Civil *sheng-yuan*; minor civil office	1
5	14	102	1.04	Expectant official, no degrees	0
5	15	4	0.04	Expectant official with one of the lower degrees	1
6	16	28	0.29	Military *juren, jinshi*; major military officer	1
6	17	45	0.46	Civil official with no degree, minor degree, or purchased degree	0
7	18	27	0.28	*Juren, gongsheng*, with no office	1
7	19	56	0.57	*Juren, gongsheng*, with expectant office	1
8	20	0	0.00	*Jinshi*, no office	1
8	21	11	0.11	*Jinshi* with official provincial post or expecant official	1
8	22	7	0.07	*Jinshi* with top-level position in Imperial bureaucracy (Hanlin Academy, Grand Secretariat, Five Boards, Prime Minister, etc.)	1

Notes: Table gives education and status for 9,787 of the adult men in the sample; coding developed using Chang (1956), Ho (1962), and Telford (1986)
See text for descriptions of the degrees titles.

Table 2. Summary Statistics

Variable	Obs.	Mean	Std. Dev.
Son Status	7,326	0.69	1.38
Father Status	7,327	0.93	1.56
Grandfather Status	7,327	1.16	1.70
Great-Grandfather Status	7,327	1.36	1.80
Great-Great-Grandfather Status	7,327	1.52	1.85
Father-in-law Status	4,488	0.58	1.58
Rank Son Status	7,326	0.50	0.23
Rank Father Status	7,327	0.50	0.24
Av. Rank Uncle Status	7,322	0.50	0.11
Son Birth Year	7,327	1770.22	66.14
Father Birth Year	7,327	1738.33	65.74
Father Longevity	7,135	57.48	12.92
Mother Longevity	6,793	58.49	14.70
No. of Son's Siblings	7,327	4.87	2.22
No of Son's Brothers	7,327	3.28	1.62
Share Male Siblings	7,326	0.71	0.24

Distribution of status			
	Frequency	Percent	Rank
Father Status Level		of sample	Father Status
0	4,812	65.67	0.330
1	520	7.10	0.690
2	877	11.97	0.790
3	486	6.63	0.880
4	354	4.83	0.940
5	144	1.97	0.970
6	30	0.41	0.984
7	79	1.08	0.991
8	25	0.34	0.998
Total	7,327	100.00	

Table 3. Transition Matrix between No Status to Positive Status

	Son No Status Freq [Percent]	Son >0 Status Freq [Percent]	Total Freq [Percent]
Father No Status	4,218 [87.88]	582 [12.12]	4,800 [100.00]
Father >0 Status	1,064 [42.29]	1,452 [57.71]	2,516 [100.00]

Table 4a . Transition matrix between nine status levels

	Son Status									
Father Status	0	1	2	3	4	5	6	7	8	Total
0	4,211	435	56	54	20	16	3	2	0	4,797
	87.78	9.07	1.17	1.13	0.42	0.33	0.06	0.04	0	100
1	415	78	10	5	7	6	1	0	0	522
	79.5	14.94	1.92	0.96	1.34	1.15	0.19	0	0	100
2	650	88	28	46	32	14	9	5	3	875
	74.29	10.06	3.2	5.26	3.66	1.6	1.03	0.57	0.34	100
3	0	0	312	81	43	20	14	12	4	486
	0	0	64.2	16.67	8.85	4.12	2.88	2.47	0.82	100
4	0	1	207	59	45	21	7	12	2	354
	0	0.28	58.47	16.67	12.71	5.93	1.98	3.39	0.56	100
5	0	0	73	30	9	14	5	11	2	144
	0	0	50.69	20.83	6.25	9.72	3.47	7.64	1.39	100
6	0	0	7	9	2	10	1	0	1	30
	0	0	23.33	30	6.67	33.33	3.33	0	3.33	100
7	0	0	14	19	10	16	9	10	1	79
	0	0	17.72	24.05	12.66	20.25	11.39	12.66	1.27	100
8	0	0	3	2	3	6	7	2	2	25
	0	0	12	8	12	24	28	8	8	100
Total	5,276	602	710	305	171	123	56	54	15	7,312
	72.16	8.23	9.71	4.17	2.34	1.68	0.77	0.74	0.21	100

Notes: For each cell of the transition matrix, table shows frequency on top and fraction of the row sum on bottom.

Table 4b. Transition Matrix over Five Generations

Son Status										
Great great grandfather status	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Total**
0	3,004	330	210	69	22	21	4	3	0	3,663
	82.01	9.01	5.73	1.88	0.60	0.57	0.11	0.08	0.00	100.00
1	324	46	17	6	5	3	1	1	0	403
	80.40	11.41	4.22	1.49	1.24	0.74	0.25	0.25	0.00	100.00
2	795	81	92	36	19	13	2	2	2	1,042
	76.30	7.77	8.83	3.45	1.82	1.25	0.19	0.19	0.19	100.00
3	686	84	222	74	61	33	12	15	6	1,193
	57.50	7.04	18.61	6.20	5.11	2.77	1.01	1.26	0.50	100.00
4	337	55	80	27	9	18	21	8	2	557
	60.50	9.87	14.36	4.85	1.62	3.23	3.77	1.44	0.36	100.00
5	80	8	44	30	21	14	2	1	2	202
	39.60	3.96	21.78	14.85	10.40	6.93	0.99	0.50	0.99	100.00
6	19	2	11	23	8	5	5	4	0	77
	24.68	2.60	14.29	29.87	10.39	6.49	6.49	5.19	0.00	100.00
7	34	4	32	19	19	13	7	14	3	41
	23.45	2.76	22.07	13.10	13.10	8.97	4.83	9.66	2.07	100.00
8	1	0	2	21	7	2	2	6	0	41
	2.44	0.00	4.88	51.22	17.07	4.88	4.88	14.63	0.00	100.00
Total	5,280	610	710	305	171	122	56	54	15	7,323
	72.10	8.33	9.70	4.16	2.34	1.67	0.76	0.74	0.20	100.00

Notes: For each cell of the transition matrix, table shows frequency on top and fraction of the row sum on bottom.

Table 5. Transition matrix between Low, Moderate, and High Status

	Son Status			
Father status	Low	Moderate	High	Total
Low	4,211	565	21	4,797
	87.78	11.78	0.44	100.0
Moderate	1,065	1,042	130	2,237
	47.61	46.58	5.81	100.0
High	0	181	97	278
	0	65.11	34.89	100.0
Total	5,276	1,788	248	7,312
	72.16	24.45	3.39	100.0

Notes: For each cell, the table gives the frequency on top and the share of the row sum below.

Table 6. Relative mobility results

	(1)	(2)	(3)	(4)	(5)	(6)
	Rank Son Status	**Rank Son Status**	**Rank Son Status**	**Rank Son Status (timevar)**	**Son Status (23)**	**Son Status (9)**
Rank Father	0.579**	0.577**	0.575**	0.574**		
	(0.013)	(0.013)	(0.013)	(0.013)		
2nd quartile				0.039**		
				(0.007)		
3rd quartile				0.047**		
				(0.007)		
4th quartile				0.028**		
				(0.008)		
Trend		-0.010*			-0.089	-0.040
		(0.004)			(0.075)	(0.028)
Father stat (23)					0.567**	
					(0.017)	
Father stat (9)						0.620**
						(0.014)
Observations	7,319	7,319	7,319	7,319	7,318	7,319
R-squared	0.384	0.385	0.398	0.375	0.434	0.492

Notes: Regression by OLS; robust standard errors clustered by father in parentheses; **/*/+ is significance at the 1%/5%/10% level. Ranks of Father and Son are based on the quartile status distribution in column 4; first quartile is son's birth years 1443 to 1739, second quartile: 1740 to 1785, third quartile: 1786 to 1818, and fourth quartile 1819 to 1885.

Table 7. Correlates of Mobility

	VARIABLES	(1) IGE	(2) IGE	(3) IGE	(4) IGE	(5) IGE	(6) IGE	(7) IGE	(8) IGE	(9) IGE	(10) IGE	(11) IGE	(12) IGE	BETA COEFF
	Trend	-0.002**	-0.002**	-0.002**	-0.003**	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	-0.009
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	
Mean	Father status		1.078**				-0.080							
			(0.257)				(0.499)							
	Father's longevity			-0.004			0.001							
				(0.005)			(0.005)							
	# Wives of father				0.370**		-0.060							
					(0.120)		(0.196)							
	Father's education					1.335**	1.454**	0.666*	0.672*	0.749**	-0.705*	-0.496	-0.496	-0.229
						(0.222)	(0.319)	(0.269)	(0.276)	(0.282)	(0.351)	(0.485)	(0.588)	
Std deviation	SD_Father's status							2.711**	2.705**	3.290**	2.063*	2.475**	2.475**	0.359
								(0.807)	(0.818)	(0.825)	(0.794)	(0.937)	(0.909)	
	SD_Father's longevity								0.001			-0.003	-0.003	-0.032
									(0.009)			(0.009)	(0.008)	
	SD_# Wives									-0.198+		-0.103	-0.103	-0.097
										(0.106)		(0.113)	(0.102)	
	SD_Father's education										1.424**	1.233**	1.233*	0.654
											(0.395)	(0.456)	(0.517)	
	Constant	0.668**	0.096	0.881**	0.253+	0.438**	0.446	-0.143	-0.154	-0.232	-0.267	-0.256	-0.256	
		(0.021)	(0.146)	(0.298)	(0.145)	(0.047)	(0.368)	(0.182)	(0.201)	(0.186)	(0.162)	(0.197)	(0.203)	
	Observations	100	100	100	100	100	100	100	100	100	100	100	100	
	R-squared	0.207	0.322	0.211	0.262	0.449	0.451	0.505	0.505	0.527	0.549	0.555	0.555	

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

Table 8a. 5-Generation Linked Sample

VARIABLES	(1) son	(2) son	(3) son	(4) son
Rank Father	0.579** (0.013)			
Rank Grandfather		0.398** (0.016)		
Rank Great-Grandfather			0.326** (0.015)	
Rank Great-Great-Grandfather				0.273** (0.015)
Constant	0.208** (0.006)	0.300** (0.007)	0.336** (0.007)	0.363** (0.007)
Observations	7,315	7,315	7,315	7,315
R-squared	0.385	0.200	0.141	0.104
Robust standard errors in parentheses ** p<0.01, * p<0.05, + p<0.1				
COMPARISON-->		0.335 .579^2	0.194 .579^3	0.112 .579^4

Table 8b. Heterogeneity in status across lineages

VARIABLES	ALL sons	1 CHEN	2 MA	3 WANG	4 YE	5 YIN	6 ZHAO	7 ZHOU
father	0.579**	0.245**	0.807**	0.436**	0.692**	0.486**	0.481**	0.435**
Observations	7,315	256	500	4,016	1,282	474	538	249
average status of lineage		0.408	0.720	0.467	0.578	0.461	0.467	0.409

Table 8c. Beyond Intergenerational Relationships

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(10)	(12)
VARIABLES	Rank Son	Rank Son	Rank Son	Rank Son	Rank Son	Rank Son	Rank Son	Rank Son	Rank Son	Rank Son
Rank Father	0.577**	0.558**	0.538**	0.547**	0.525**	0.570**	0.891**	0.554**	0.799**	0.804**
	(0.013)	(0.017)	(0.014)	(0.014)	(0.014)	(0.013)	(0.028)	(0.017)	(0.036)	(0.037)
								[0.593]	[0.855]	[0.860]
Rank Grandfather		0.025						-0.041+		-0.052**
		(0.016)						(0.021)		(0.020)
								[-0.045]		[-0.058]
Rank Great grandfather			0.071**					0.066**		0.060**
			(0.012)					(0.020)		(0.019)
								[0.075]		[0.069]
Rank Great great grandfather				0.064**				0.036*		0.012
				(0.011)				(0.015)		(0.014)
								[0.042]		[0.014]
Average of Uncles					0.294**				0.163**	0.136**
					(0.030)				(0.036)	(0.037)
									[0.075]	[0.063]
Maximum Rank Uncle						0.073**				
						(0.019)				
Max Uncle Relative to Father							0.101**		0.080**	0.081**
							(0.008)		(0.010)	(0.010)
									[0.294]	[0.295]
Trend	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
								[-0.029]	[-0.044]	[-0.047]
Constant	0.389**	0.379**	0.363**	0.370**	0.130+	0.386**	0.215**	0.371**	0.103	0.124
	(0.074)	(0.074)	(0.074)	(0.074)	(0.073)	(0.073)	(0.075)	(0.073)	(0.078)	(0.076)
Observations	7,312	7,312	7,312	7,312	7,306	7,312	7,312	7,312	7,306	7,306
R-squared	0.385	0.385	0.390	0.389	0.400	0.386	0.405	0.391	0.409	0.412

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

Table 8d. Competition between Lineages

	(1)	(2)	(3)	(4)
VARIABLES	Rank Son	Rank Son	Rank Son	Rank Son
Rank Father	0.576**	0.542**	0.574**	0.541**
	(0.013)	(0.014)	(0.013)	(0.014)
Average Status All	0.030			
	(0.057)			
Lineage Specific Status		0.216**		0.209**
		(0.023)		(0.023)
Average Status Other Lineages			-0.177**	-0.132**
			(0.039)	(0.039)
Constant	0.358**	0.198**	0.612**	0.371**
	(0.091)	(0.071)	(0.089)	(0.085)
Trend	Y	Y	Y	Y
Observations	7,313	7,252	7,313	7,252
R-squared	0.384	0.394	0.386	0.395

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

Table 8e. Status Mobility and the Role of the In-Laws

	(1) Full Sample	(2) InLaw Sample	(3)	(4)	(5) Father In Law = 0	(6) Father In Law > 0	(7) Father In Law > 0
Rank father	0.577**	0.594**	0.492**	0.116*	0.459**	0.859**	0.871**
	(0.013)	(0.014)	(0.017)	(0.053)	(0.019)	(0.044)	(0.157)
Rank father in law			0.314**	-0.342**			
			(0.023)	(0.086)			
Rank Father x Rank Father-in Law				0.793**			-0.012
				(0.097)			(0.157)
Observations	7,320	4,486	4,486	4,486	3,896	590	590
R-squared	0.381	0.390	0.426	0.436	0.243	0.493	0.493

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

Table 8f. Status Mobility and Intermarriage

	(1)	(2)	(3)	(4)	(5)	(6) Son low status lineage	(7) Son high status lineage
Rank father	0.577**	0.551**	0.513**	0.549**	0.511**	0.491**	0.611**
	(0.013)	(0.043)	(0.043)	(0.039)	(0.043)	(0.069)	(0.059)
Mother Hi Status	0.030	0.019				0.129+	-0.150**
	(0.046)	(0.050)				(0.065)	(0.045)
Rank Father x Mother High Status	-0.037	-0.009				-0.290+	0.313**
	(0.091)	(0.101)				(0.150)	(0.065)
Rank father x Both parents High			0.421**		0.423**		
			(0.051)		(0.051)		
Both Parents High			-0.196**		-0.198**		
			(0.032)		(0.032)		
Rank father x Both parents Low				-0.135**	-0.141**		
				(0.025)	(0.025)		
Restrict to 7 husband lineages	N	Y	Y	Y	Y	Y	Y
Observations	7,330	705	705	705	705	473	232
R-squared	0.382	0.359	0.377	0.359	0.378	0.244	0.528

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1