Mobility through Marriage over Multiple Generations: Evidence from Imperial China*

Wolfgang Keller and Carol H. Shiue†

March 21, 2024

Abstract

This paper studies the role of marriage for mobility with intergenerationally linked genealogical data ranging from 1500 to 1900 in Central China. Augmenting the father-son relationship with a status measure of the in-law family indicates that marriage is central for understanding mobility. We show that a 10 percentage point status premium of the in-law family typically translates into a 2.5 percentage point higher status of the son, on par with the influence of his own father. Marriage quality matters even for brothers, accounting for up to half of their status differences. The degree of assortative mating between biological and in-law families is considerable–a correlation of 0.5–, and ignoring it means overestimating mobility. The analysis of three linked generations confirms the role of son's own marriage for his mobility and also shows that the quality of the son's parents' marriage influences his status. Quantitatively, the role of marriage for mobility is at least as large as the role of higher-order members of the biological family.

Keywords: Intergenerational Transmission, Assortative Mating, Chinese Clans

*Thanks to Ran Abramitzky, David de la Croix, Hao Dong, and Rui Esteves for comments, as well as to seminar audiences at the University of Hong Kong and CEPR Paris Symposium. Ted Telford provided data employed in this paper. Financial support from NIH under grant NICHD Award HD042731-01 (Shiue) is gratefully acknowledged.

†shiue@colorado.edu, kellerw@colorado.edu; University of Colorado, CESifo, CEPR, and NBER.

1 Introduction

Upward mobility through marriage is truly an old idea. It took only one clever engagement between Philip of Austria and Juana of Castile and Aragon for the House of Habsburg to end up with the entire Spanish Empire, for example.[1] Yet, while marriage in many countries is considered to be key to status and wealth, outside of the elites for the broader population relatively little is known on these matters. This paper examines the role of marriage for intergenerational mobility, and how marriage compares with other investments that parents make into their children.

This analysis studies a sample of approximately 5,000 observations of married men over three generations centered on the period from 1650 to 1870 in Tongcheng, a county of China's Anhui province. To examine the role of marriage for mobility, we consider the relationship between son status and a measure of the status of the wife the son marries–her father's status. Father-in-law status turns out to be highly predictive of son status, and the coefficient of 0.4 is similar to that in the typical intergenerational specification of son status on biological father status. We show that key to this is positive assortative mating, with a correlation of biological and in-law family status of about 0.5.

Next, we evaluate the roles of the biological and the in-law family side by side by including own father status and father-in-law status simultaneously. Both enter with a positive coefficient, and they are similar in size. This indicates that marriage plays a role for son status independent of that of his biological father. For example, our results show that while the best prediction of a son with a low-status father is that he ends up at percentile rank 44 in the status distribution, marriage into a top in-law family would mean the son moves up to the 58th percentile rank, an improvement in status of more than 30%. For most sons, the typical range of status for most successful marriage (top in-law status) to least successful marriage (lowest in-law status) is about fifteen percentile ranks.

While there is historical evidence that both groom and bride families search for a match, generally little is known on why some families are successful in historical marriage markets while others are not. In particular, there is little evidence on the role of within- versus between-family variation for marriage success. Do we find assortative mating in the aggregate because rich families' sons tend to marry rich families' daughters and that goes more or less equally for all their sons and daughters, or does marriage play a role for son status even within-family? To address this question we focus on the role of marriage for the status of brothers.

Compared to the population overall, brothers have a relatively large part of their genetic and resource endowment in common because they are children of the same families. The amount of status variation among brothers in our sample is only about one fifth of that among all sons. And yet, we find that quality differences in the marriages of brothers translate in a big way into differences in status of the brothers. Specifically, for two brothers with marriages in which the status of the in-law

[1] A well-known saying describing Habsburg's strategy is *Bella gerant alii, tu felix Austria nube* (let others wage war: you, happy Austria, marry). The House of Habsburg benefited in this case from a series of early deaths of throne successors higher up the ladder, and, at any rate, Habsburg did not shy away from waging wars.

fathers are ten percentage points apart, the brother with the better marriage can expect to have a status almost five percentage points higher than his brother. The result highlights that marriage matters for intergenerational mobility even conditional on the resources of the family.

Our second set of results compares the role of marriage for mobility with that of higher-order members of the son's own biological family. The analysis begins by confirming that grandfathers play a role for son mobility independent of the status of the father, and that accounting for multi-generational effects lowers the estimated rate of mobility, here by 15 percent.

We proceed by including status of the son's father-in-law and the son's mother's father–his maternal grandfather– to capture the role of current and past marriages (i.e., the son's own and his parents' marriage, respectively). Results indicate that both biological family and in-law family characteristics play a significant role for mobility. Quantitatively, coefficients for father and father-in-law status are estimated with similar magnitude, in line with our earlier findings. Paternal and maternal grandfather estimates are also similar (about 0.075). Accounting for higher-order biological family and marriage (in-law family) decreases the estimated rate of mobility by 44 percent compared to the simple father-son estimate. Overall, marriage patterns play a similar if not larger role for social mobility compared to accounting for higher-order biological family.

It is important to consider that our analysis is based on a comparatively small sample of privately held family records, in contrast to recent studies based on large administrative datasets. Early work on the role of women for social mobility in modern societies, however, employs quite small samples as well.[2] Even recent studies using US administrative data for 19th or 20th century may not employ samples that are orders of magnitudes larger than our sample.[3] The key concern with genealogical data is its non-representativeness because it required literacy and resources to compile a genealogy. Employing all extant genealogies of Tongcheng county, for example, would yield a sample in which wealth is higher than typical in the population, something which is true for genealogical samples based on crowd-sourcing as well (Stelter and Alburez-Gutierrez 2022). Given that we do not have census-like data for benchmarking, we adopt a targeted approach in which the sample is constructed with the explicit goal to match key statistical moments–including the shares of different status levels in the population–on which there is broad consensus (see sections 3, A.3). Information from Chinese family genealogies is generally considered to be reliable because it served important private economic functions for the clan (see Harrell 1987, Shiue 2016, 2024).

In addition, the present paper broadens the scope of analysis by considering not only the role of marriage for intergenerational mobility but also for intragenerational mobility. We relate son status to a measure of the status of his wife, which is an intragenerational relationship. While the analysis of intragenerational mobility is common in some social sciences, economic analysis tends to focus on the role of marriage for intergenerational relationships.[4] One reason may be that

[2]E.g., Chadwick and Solon (2002) is based on just over 500 observations from the Panel Study of Income Dynamics.

[3]One reason is the need to establish intergenerational links ex-post. For example, Espin-Sanchez, Ferrie, and Vickers's (2023) multigenerational study is based on N = 23,736 observations.

[4]For example, social stratification has been defined as "the inter- and intragenerational reproduction of social

intergenerational relationships are more plausibly exogenous compared to intragenerational links. At the same time, intergenerational relations in the vast majority of studies including the present one are correlations, not causal effects, so including intragenerational marriage relationships hardly changes the fundamental nature of the analysis. Furthermore, our measure of the status of the wife's family is the status of her father, which is predetermined at the time of her marriage. This reduces the concern that what we measure as wife status is simply the status bestowed upon her by her husband.

While most research on the intergenerational transmission of economic status assumes a single-parent household model to study mobility between father and son, by now there is a small but growing literature on the role of women and marriage for social mobility in historical periods (including Olivetti and Paserman 2015, Eriksson, Lake, and Niemesh 2019, Espin-Sanchez, Gil-Guirado, and Vickers 2022, Bailey and Zin 2022, Curtis 2022, Olivetti, Paserman, Salisbury, and Weber 2022, Espin-Sanchez, Ferrie, and Vickers 2023, and Althoff, Gray, and Reichardt 2023).[5] A key insight is that the correlation in status between fathers and sons is a function of both the effect of fathers and mothers on son, and thus it depends on the degree of assortative mating (e.g., Espin-Sanchez, Gil-Guirado, and Vickers 2022). Our estimate of 0.48 is higher than most, in particular for 19th and 20th century US where comparable assortative mating correlations in different studies range between 0.04 and 0.52 (Buckles, Price, Ward, and Wilbert 2023).[6] Given that the downward bias from using the father-son correlation is increasing in the degree of assortative mating (shown below), in the present analysis the implied role of women for mobility is greater than elsewhere. Furthermore, our analysis considers intragenerational marriage mobility in addition to intergenerational marriage effects by relating son status to the status of his in-law family, which turns out to be quantitatively important (see also Van Bavel, Peeters, and Matthijs 1998).

It has been increasingly recognized in studies of intergenerational status transmission that the focus on two generations, parent-child, may lead to biases because previous generations affect the status of the child up and beyond the influence of the parent generation (Mare 2011, Solon 2014). In particular, it has been shown that occupational income or wealth is positively correlated in the US and the UK in the late 19th and early 20th century over multiple generations (Long and Ferrie 2018). A common result is that the focus on two generations overstates mobility because grandparents exert a positive influence that is not entirely captured by the father coefficient. The present paper presents two findings in this respect. First, we confirm that abstracting from higher order generations overestimates mobility in a relatively early sample from Imperial China. Second, we extend the study of multigenerational mobility to include marriages in previous generations, showing that the role of marriage for mobility is larger than the influence that grandparents have on the status of their children.

Much progress has been made in connecting census data to obtain large intergenerational samples

inequality through homogamous marriage patterns and status inheritance over families and generations" (Haller 1981).

[5] Chadwick and Solon (2002) and Ermisch, Francesconi, and Siedler (2006) study modern economies.

[6] Buckles, Price, Ward, and Wilbert's (2023) preferred instrumental variables estimate of assortative mating is 0.68.

through automation and genealogical crowd-sourcing, and it has led to several novel findings (see Abramitzky, Boustan, Eriksson, Feigenbaum, and Perez 2021, Buckles, Price, Ward, and Wilbert's 2023). At the same time, broadly representative official data is a fairly recent development, going back for the United States to the mid-19th century, for example, while China had its first census of the modern era in 1953.[7] There is typically no census data to study intergenerational mobility in pre-industrial societies, and the analysis of mobility in the very long requires alternative data sources. Data sources for analysis over the very long run are important more generally because the historical roots of economic development often lie far back in the past (Nunn 2020).

One approach to long-run mobility analysis is to employ pseudo-intergenerational links based on first or last names instead of actual intergenerational links (Clark 2014, Olivetti and Paserman 2015, and Guell, Rodriguez, and Telmer 2015). While this approach has produced interesting insights, meta analysis shows that limited informational content in names and other issues can also lead to large biases (Santavirta and Stuhler 2024). Family genealogies are an attractive alternative data source not only because actual intergenerational links are well documented because they are central to genealogies but also because they are common in many parts of China going back many centuries. Among other issues, they have been employed to study long-run behavior in terms of fertility, migration, and human capital acquisition (Shiue 2017, Shiue and Keller 2024, Hess 2024, and Hu 2023).

In the remainder of this paper, the following section 2 provides a synopsis of marriage and mobility in Ming-Qing China. Section 3 describes the data employed in this study, with additional analysis of sample characteristics and representativeness shown in the Appendix. Our empirical results are presented in section 4. We begin by analyzing the role of marriage for mobility in two generations, before turning our attention to the influence of marriage for the status of brothers. The second part of our empirical results concerns the role of marriage in the context of multigenerational mobility. Section 5 provides a summary and some conclusions, and further details on the data as well as robustness checks are shown in the Appendix.

2 Marriage and Mobility in the Ming-Qing

Social scientists have made much progress on the subjects of marriage, inequality, and the specific role of women during the Ming and Qing dynasties (Watson and Ebrey 1991 is an introduction). For example, the conventional view in the literature holds that marriage was the ladder of success for women in late imperial China, that women stayed typically in the inner part of the Chinese house, and that marriage was primarily arranged by the parents.[8] At the same time, much of what we know is drawn from general sources such as the teachings of Confucius, the Qing code, or etiquette booklets. As far as actual behavior is concerned, historical sources impose tight limits on what is

[7] Hao and Xie (2023), e.g., analyze assortative mating on education in 20th century China.
[8] See Mann (1991, p.204), Ebrey (2023), and Shan (2013), respectively.

known, especially on lower status individuals and women.[9] While this paper seeks to contribute some new evidence from a sample population of Tongcheng, the following section sketches some generally agreed-upon features for Ming-Qing China.

Typically in a Chinese marriage the husband would gain the productive value of the wife, and in exchange the wife would receive a claim to the privileges of the husband's ancestral rites, as well as a stake in the economic fortunes of her husband, even if he died (Watson and Ebrey 1991). Marriage was the union of two extended families (clans), not of individuals, and as such the parents, grandparents, and even the father's brothers and sisters determined who a younger relative should marry (Shan 2013). The families of women and men could gain status through marriage, which was an occasion to adjust their economic position (Shan 2013), and would require often substantial expenditures.

Children were typically betrothed at the age of 8 to 10 years, and even younger for high-status families. In general, marriage was an occasion in which the wife changed her residence. Not only would the wife move from her father's family to her husband's family (patrilineal), but she would also move from her father's home to her husband's home (patrilocal). A man had usually a single female partner, his married wife, during his lifetime. The couple might fulfill their marital goals, in particular a certain number and gender of children, or not. Producing a male heir to ensure that the family line lives on was one of the most important roles of the wife. If the wife does not have (male) children, or the wife dies, as a rule the family line was broken, unless an adoption of a male heir could be arranged.

The exception to this is that the husband remarries and has a male child with a second or later wife. The ability to remarry depended largely on the status of the husband–given the considerable marriage expenses, most men could not afford to marry a second time. From the point of view of the second wife, remarriage came with certain risks of alienating offspring (from the first wife) and losing one's authority over the coming generation (Mann 1991). A stepmother had to deal with such things as the resentment of her stepchildren who would consider her an outsider, and the insubordination of her daughter-in-law who they were supposed to manage (Mann 1991).

On the other hand, if the husband died early, his widow ranked in the household under the male heir if there was one, and she could be a chaste widow. If there was no male heir, the woman's position in the household was precarious, leading often to her remarriage even though this was generally frowned upon in this society (Mann 1991). Women who would remarry due to their husband's death tended to do so with a man of lower status (Telford 1992).

As a rule families had a choice of whether to invest in the marriage of their daughters or in the education of their sons. Well connected in-laws or an educated wife could provide benefits for the

[9] Anecdotally, Ebrey (2023) notes that already by the Song dynasty (960-1279) there were Chinese widows who ran inns, midwives delivering babies, pious women who spent their days chanting sutras, nuns who called on such women to explain Buddhist doctrines, girls who learned to read with their brothers, famers' daughters who made money by weaving mats, childless widows who accused their nephews of seizing their property, and women who drew from their dowries to help their husband's sisters marry well.

next generation because of access to well-connected networks of knowledge, and an educated mother would have been the first teacher to the young males of the next generation. In that sense, marriage was a ladder to success for women. Because status could not be passed on directly from parents to their children, the economic success of the next generation depended to a greater extent on the investments parents made, and with that, it depended on the match made at the time of marriage. Marriages, and the associated investments through searching for a good marriage, were one of the channels through which wealth was passed on to the next generation.

The question of whether marriage perpetuated intergenerational inequality would depend on the specific patterns of marital sorting. If marriage is mostly between likes it raises cross-household inequality, while if differences attract each other marriage is a force towards equality. While one might expect that families want their daughters to always to "marry up", there were reasons that men might want to "marry down"– because such brides were thought to be harder workers and more easily satisfied with their situation (Ahern 1974, Freedman 1979).[10]

The patterns of marital sorting depended also on the availability of spouses, the age of bride and groom, and other time-varying factors (Mann 1991, Telford 1992). For example, women of the 17th century Jiangnan regions increased their cachet as wives by becoming increasingly educated, triggering a rising incidence of (non-arranged) "companionate marriages", and this can be traced to commercial growth and urbanization (Ko 1994). There are also examples of girls during the 18th century who despite being in an arranged marriage with a scholar husband would stay in their parents household and learn alongside their older brother from their mother until puberty mandated sex segregation (Mann 1997).

Meanwhile the costly expenses of investments in education for sons competed with the costs of dowry in marriage. As the returns to education declined, however, fewer families would have found this a productivity channel of investment. Since positive assortative matching typically requires considerable search costs, declining likelihood of success in officialdom would likely have decreased the returns to high marriage search costs or dowry. Thus, the flip side of the decline in the return to education is a decline in the return to marital sorting (Watson and Ebrey 1991). One would expect a decline in marriage matching as the returns to male education fell.

3 Data

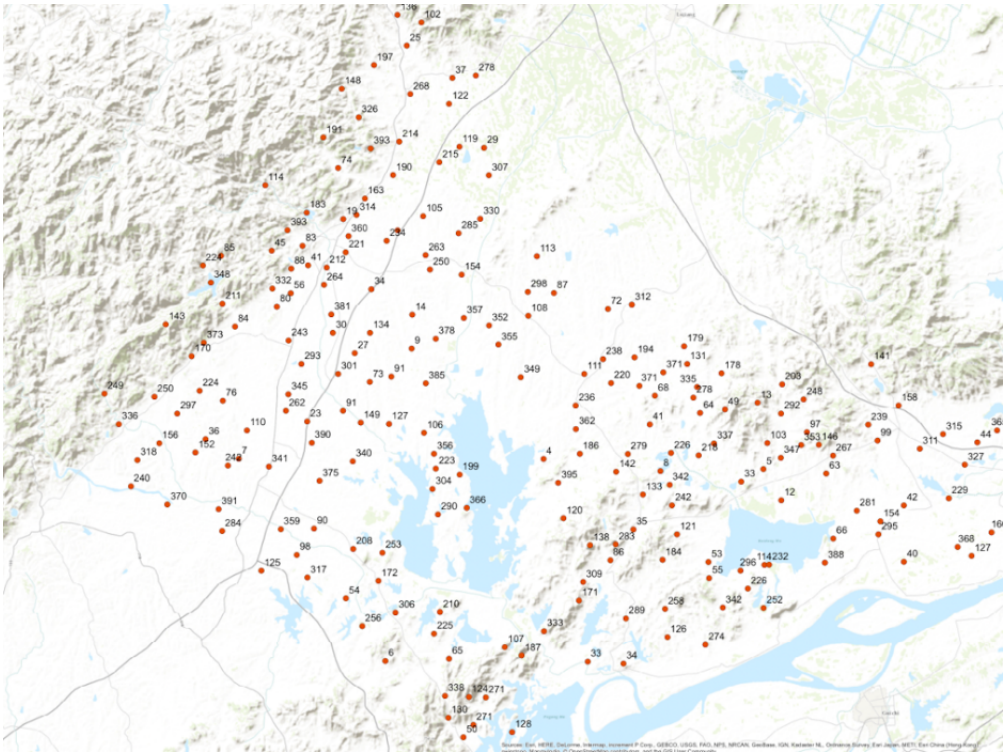
3.1 Tongcheng County: Location

This study examines marriage and multigenerational mobility using genealogical data for six clans in Tongcheng county, part of Anhui province. The lifetimes of our three generations span roughly

[10] Often, parents would seek a so-called "matching doors" marriage in which families sought matches between couples of comparable social backgrounds. While gifts were presented both from bride's family to groom's (akin to a dowry) and from groom's family to bride's (akin to a bride price), in no period was China a dowry society comparable for example to India (Watson and Ebrey 1991). It was important that gifts were not too one-sided lest one gives the impression that a spouse is "purchased" (Mann 1997).

the years 1500 to 1900, centered on the period 1650 to 1870. Thus, the sample period includes parts of the Ming (1368-1644) and Qing dynasty (1644-1911).[11] Genealogies are a classic source of socio-economic data for China, and estimates put the number of existing genealogies into the tens of thousands (Wang 2008). Earlier research such as Beattie (1979) has employed dozens of genealogies from Tongcheng county.

Figure 1: Tongcheng County and Villages in the Sample



Notes: Location of towns and villages in Tongcheng county, in Anhui province, with basic topography and roads.

Figure 1 shows some of the historical towns and villages of Tongcheng county, which is located north of Anqing, the capital of Anhui province. The county spans about 100 kilometers East-West and 60 kilometers North-South. The Yangzi river is shown in the south-eastern part of Figure 1.

3.2 Status Information

Because documenting intergenerational links is a key purpose, genealogies are particularly useful for studying intergenerational family responses. The genealogies employed in this paper are essentially richly annotated pedigree charts, though for practical reasons the information is organized by married couple, which underlines the central importance of marriage to these families. We also know which of the men married, and which did not. Among the men who married, we know how many wives they had, and, through examining their death dates, we can determine if the wives were alive at the same time or whether it is likely that the marriages happened sequentially. The importance

[11] Chinese clans are also referred to as lineages or common descent groups in the literature.

of descent is underlined by the that the genealogies list the children that were born to each wife separately. An example of the genealogies employed in this study is shown in Figure A.1.

Status of men in this study is based on up to 30 descriptors in the genealogy that characterize his status in society, including any official positions (and at what level), levels of education, wealth, donations, and honors. It is important that the same status classification applies to all five types of males in the analysis: the son, his father, and his paternal grandfather, as well as the father of his wife and the father of his mother. Table 1 shows the status groups distinguished in the analysis, as well as the sample shares for the paternal grandfather.

Table 1: Status Categories

Group	Descriptors	N (%)
1	No title, degree, and evidence of wealth; Honorary title; village head; Moderate wealth and status of 1st and 2nd degree family member	3,848 (77.0)
2	Official Student	239 (4.8)
3	Wealthy farmer, landowner, or merchant; set up lineage estates, large donations,	182 (3.6)
4	Military *sheng-yuan*	0 (0.0)
5	Imperial Academy Student	408 (8.2)
6	Civil *sheng-yuan*	70 (1.4)
7	Expectant official, no degrees	77 (1.5)
8	Expectant official, lower degree	0 (0.0)
9	Military *juren*, *jinshi*	5 (0.1)
10	Civil official, no or purchased degree	39 (0.8)
11	*juren*, *gong-sheng* with no office	25 (0.5)
12	*juren*, *gong-sheng* with expectant office	66 (1.3)
13	*jinshi*, no office	0 (0.0)
14	*jinshi*, official provincial post or expectant official	14 (0.3)
15	*jinshi* with top-level position in Imperial bureaucracy	27 (0.5)
All		5,000 (100.0)

Notes: Sample is based on all married men that can be linked over three generations. Classification based on Telford (1986), Chang (1955, 1962), Ho (1967), and Eberhard (1962). Distribution of status in final column is for paternal grandfathers.

The descriptors of Table 1 indicate that status in China during the sample period was strongly related to have passed the civil service exam and obtained official position. Scholar-officials had the highest status in this society. A graduate of the local civil service exam would be a *sheng-yuan*,

someone who passed the provincial exam a *juren*, and graduates of the national exam had the *jinshi* degree. Generally, the higher the degree, the more important would be the official position that a man would be able to attain. The monetary rewards of passing the civil service exam were high, although there were also rich merchants or landowners that did not have a degree or official position. Often, rich merchants or landowners would invest into the education of their sons to pass the civil service exam, because it was "the ultimate gateway to power" (Ho 1967) in China during our sample period.

Table 1 indicates that 77 percent of all paternal grandfathers belong to the lowest status group. While it would be interesting to be able to distinguish the lowest status category further, a fraction of around three quarter in the lowest status category is consistent with other evidence for China during this period. Furthermore, we see that graduates of the national civil service exam (*jinshi*) rank in the top 0.8% of the status distribution. The distribution of status for the other four types of men (sons, fathers, father-in-law, and maternal grandfather) are broadly similar to that in Table 1.

In the regression analysis below, status is employed as percentile rank (following Dahl and DeLeire 2008), which has the advantage over the status categories on the left in Table 1 that it accounts for the size of the group, among other advantages. For example, with 77.0% of men belonging to the lowest status class, each of them would be assigned 0.385 (the midpoint between 0 and 0.77). The focus on percentile ranks implies that by construction, the mean of every status variable is equal to 0.5.

Furthermore, because the distribution of status changes over time, we compute each man's percentile rank relative to the subperiod-specific distribution. Three subperiods are distinguished by birth year, namely (i) before 1700, (ii) between 1700 and 1800, and (iii) 1800 and later.[12] We also include son birth year fixed effects to capture shocks, such as the fall of the Ming dynasty (around 1644), as well as trends in the data.

Table 2 provides information on the lifetimes of the three-generation triplets in this multigenerational analysis based on the middle, the father generation.

Table 2: Lifetimes of Fathers

	Min	5%	50%	95%	Max
Birth Year	1518	1650	1759	1815	1859
Death Year	1567	1707	1814	1870	1894

Notes: Sample of all males linked over three generations for whom father-in-law status is available; N = 5,000

[12] There is a secular decline over time in the typical status that men would obtain because the return to passing the civil service exam fell towards the end of the Qing dynasty, see Shiue (2024).

Notice that while the lifetimes of fathers almost span the period 1500 to 1900, the lifetimes of the central 90 percent lie between the years 1650 to 1870. Furthermore, the median birth year of 1759 indicates that our analysis is mostly informed by mobility during the Qing than the Ming dynasty.

3.3 Representativeness

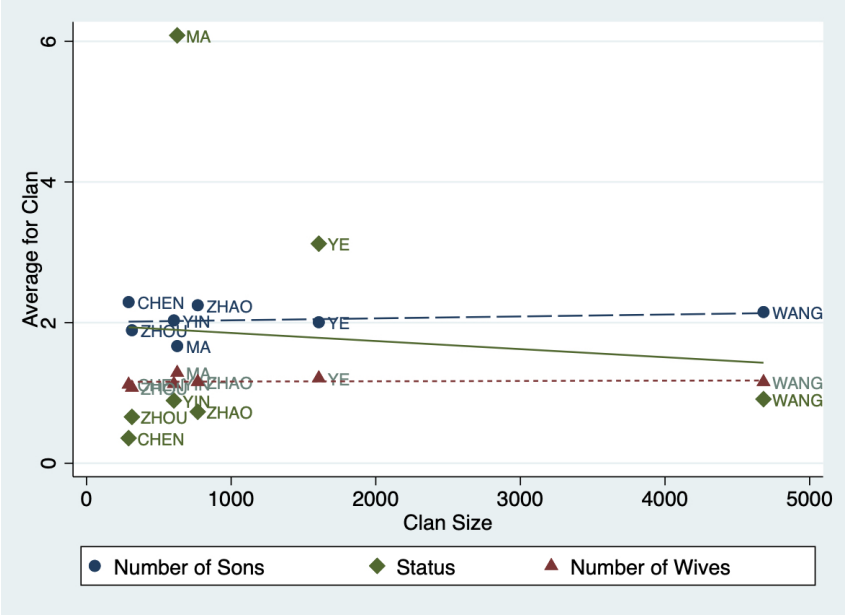
Genealogical information is self-reported, and there are no penalties for misrepresentation, which may give rise to the concern that it is unreliable. At the same time, there is only sparse official data available during the sample period, and it does not appear to be as reliable as the clan records (see Figure A.2).

Beyond ancestral worship Chinese genealogies served a number of important economic functions for the clans themselves that would require accurate information. First of all, property rights turn on information recorded in genealogies, because they establish and sustain village settlement rights for specific clans. Second, genealogies are critical as a means of defense, including war, because by determining who is member of the clan and who is not it defines allegiances, rights, and responsibilities in times of conflict (both versus other clans and versus the government). Third, genealogies provide information on taxation and public goods provision. On the one hand, the state delegates to local clans the right to tax as well as the responsibility to fund public works such as irrigation. On the other hand, a clan's genealogy would specify assessments (essentially taxes) on their members to found and maintain common clan property.

The sample was created using a targeted approach by considering more than three dozen genealogies from Tongcheng county and selecting a subset with the goal of generating a broadly representative sample. Seven clans have been chosen, with the following last names: Chen, Ma, Wang, Ye, Yin, Zhao, and Zhou. Conditional on choosing the genealogy of a given clan, all entries of the genealogy become part of our sample. Given these clans' characteristics and size, we have about 20% of the sample belonging to the upper class as defined by Fei (1946), and the fraction he gives nationally for that class is also 20%. Also, the fraction of local and provincial civil service examination graduates in the sample is comparable to Chang's (1955) figure of 2%. Thus, groups with different level of status account for similar shares in the sample and in various analyses of China overall.

Figure 2 shows a flat relationship between clan size and average status. Thus, it is not the case that high-status clans are overrepresented in the sample. This reflects the genealogical principle that all male clan members are included in the genealogy irrespective of status. In addition, Figure 2 considers other indicators of resources. In particular, wealthier men could afford multiple wives and maintain larger families (Harrell 1985). The figure shows the relationship between clan size and the number of wives is essentially flat. Furthermore, fertility as measured by the number of male children is not strongly related to clan size either, see Figure 2.

Figure 2: Clan Size versus Status, Fertility, and Female Partners per Man



Notes: Figure shows three relationships, between average (i) number of sons, (ii) husband status, and (iii) number of females per man, each with the number of clan members across seven male clans.

Section A.3 examines these issues further by considering a range of forms of selection and biases, including recall bias, progenitor bias, and survivor bias, finding little evidence that the sample would not be broadly representative of the Tongcheng area.

Our estimation sample with three linked generations reflects two important sample constraints. First, the men included in the sample are all married. While this is natural for an analysis of the role of marriage for mobility, it implies that the sample is upward selected compared to China's population at the time. In particular, it is estimated that between 10 and 20% of all males failed to marry in this period, for two reasons: (i) sons died before reaching adulthood, and (ii) even though they reached adulthood they were not able to marry given their resources and the prevailing sex ratio.[13] In the present analysis we take the focus on married men as given; an analysis of single versus married men is provided in Shiue and Keller (2024) in the context of the fall of the Ming dynasty.

The second sample constraint is availability of status information on the son's wife's father as well as the son's mother's father. Such in-law information across clans is available to a somewhat lesser extent than information on the biological family. Observations for which in-law information is available tend to be for relatively rich families, although quantitatively the effect is small: the average percentile rank status is 0.51 when in-law information is available, compared to 0.49 when it is not.

[13]Other reasons for non-inclusion in the genealogy such as outmigration into far-away areas were not quantitatively important.

Table 3 shows each clan's average status as well as sample share once the sample constraints have been imposed.[14]

Table 3: Status and Sample Share by Clan

	Chen	Ma	Wang	Ye	Yin	Zhao	All
N	108	415	3,285	926	28	238	5,000
(%)	(2.2)	(8.3)	(65.7)	(18.5)	(0.6)	(4.8)	(100.0)
Average Status	0.461	0.662	0.469	0.556	0.463	0.459	0.5

Notes: Given is the average son percentile rank status by clan.

Table 3 confirms that the clans differ in terms of size, with the Wang and the Ye clan being the largest. Also note that members of the Ma clan have on average the highest status. If rich men were overrepresented in the same, one would expect that the Ma clan accounts for a large fraction of the sample overall–and yet there are two clans with higher numbers in the sample. Across the six clans, the correlation between sample share and average status is low (and negative, -0.11).

We now turn to the empirical results.

4 The Role of Marriage for Intergenerational Mobility

4.1 Marriage with Two Generations

To analyze the role of marriage for intergenerational mobility, we specify the long-run component of the status of a daughter d of family i, $stat_{di}$, as in Lam and Schoeni (1994) and Chadwick and Solon (2002) by

$$stat_{di} = \alpha_d + \beta_d stat^p_{di} + \varepsilon_{di}, \tag{1}$$

where $stat^p_{di}$ is the long-run component of the status of her parents, and ε_{di} is the error term capturing factors affecting daughter status that are orthogonal to parent status. The slope coefficient β_d captures the intergenerational relationship between parent and daughter status. It is positive if daughters of high-status parents tend to have high status themselves.

We assume that assortative mating is summarized by a correlation γ between the daughter's and the son's status:

$$\gamma = Corr\left(stat_{di}, stat_{si}\right). \tag{2}$$

The relationship between son's status, $stat_{si}$, and the status of the daughter's parents is then given by (Lam and Schoeni 1994, Chadwick and Solon 2002)

$$stat_{si} = \alpha_s + \beta_s stat^p_{di} + \varepsilon_{si}, \tag{3}$$

[14]Note that the sample does not include members of the Zhou clan anymore.

where

$$\beta_s = \beta_d \gamma \frac{\sigma_s}{\sigma_d}, \tag{4}$$

with σ_s being the standard deviation of the status of the son and σ_d the standard deviation of the status of the daughter. Equation (4) shows that if the variances of daughter and son status are the same, the coefficient in the regression of son status on the status of the daughter's parents is equal to the product of the intergenerational coefficient of the daughter and the degree of assortative mating, $\beta_d \gamma$. In particular, if there is no assortative mating, β_s is equal to zero, and the higher is the degree of assortative mating, the stronger is the relationship between in-law status and son status. Equation (4) also shows that given an estimate of γ, for example from the correlation of status of father and father-in-law, regressing son status on father-in-law status can help to quantify the effect of parent income on the status of the daughter, β_d.

A second way to shed light on the role of marriage for mobility starts from the relationship between father and son status, which is given by

$$stat_{si} = \widetilde{\beta}_0 + \widetilde{\beta}_1 stat^p_{si} + \epsilon_{si}, \tag{5}$$

and parent status $stat^p_{si}$ is proxied by the status of the father, $stat^f_{si}$:

$$stat_{si} = \widetilde{\beta}_0 + \widetilde{\beta}_1 stat^f_{si} + \epsilon_{si}.$$

If instead due to marriage both biological and in-law parents matter for son status so that the true relationship is given by equation (6)

$$stat_{si} = \beta_0 + \beta_1 stat^p_{si} + \beta_2 stat^p_{di} + u_{si}, \tag{6}$$

where the daughter's parents' status $stat^p_{di}$ is approximated by her father's status, $stat^f_{di}$:

$$stat_{si} = \beta_0 + \beta_1 stat^f_{si} + \beta_2 stat^f_{di} + u_{si}, \tag{7}$$

then omitted variables analysis says that the OLS estimate $\widetilde{\beta}_1$ in equation (5) is equal to

$$\widetilde{\beta}_1 = \beta_1 + \beta_2 \widetilde{\gamma}, \tag{8}$$

where

$$\widetilde{\gamma} = Corr(stat^p_{di}, stat^p_{si}) = Corr(stat^f_{di}, stat^f_{si}), \tag{9}$$

the correlation between the daughter's and son's parents' status. Both equation (3) and equation (6) indicate that assortative mating crucially affects the interpretation of intergenerational mobility findings.

They differ in the definition of assortative mating, which is measured between biological and in-law parents according to equation (9) while it is measured between husband and wife according

to equation (2). The main reason why in the present case the parent measure is preferred is that both status of the biological and status of the in-law father are observed, whereas the status of the daughter is not observed. In addition, parent status has the advantage that it is predetermined at the time of marriage of husband and wife. We henceforth assume that $\gamma = \widetilde{\gamma}$.

We begin the empirical analysis by estimating the following equation by OLS:

$$stat_{si} = \beta_0 + \beta_1 stat^f_{si} + \beta_2 stat^f_{di} + \mu_{t(i)} + \varepsilon_i, \tag{10}$$

where $stat_{si}$ is the percentile status rank of the son from family i, and $stat^f_{si}$ is the percentile rank status of the son's father, and $stat^f_{di}$ is the percentile rank status of the son's father-in-law (or, the daughter's biological father). Equation (10) is identical to equation (7) except for the inclusion of birth year fixed effects of the son, $\mu_{t(i)}$. Given the relatively long sample period, they are potentially important controls to account for shocks and secular trends. Conditional on these variables, ε_i is assumed to be a mean-zero error term. Table 4 shows the results.[15]

Table 4: Marriage and Parent-Child Mobility

	(1)	(2)	(3)	(4)	(5)
Father Status	0.417**	0.410**		0.281**	
	(0.018)	(0.018)		(0.020)	
Father-in-Law Status			0.423**	0.271**	
			(0.020)	(0.022)	
Average (Father Status, Father-in-Law Status)					0.553**
					(0.020)
Husband Birth Year Fixed Effects	N	Y	Y	Y	Y
N	4,996	4,967	4,967	4,967	4,967

Notes: Dependent variable is son status as percentile rank. Estimation of equation (10); robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

The first specification includes only father status as a right hand side variable ($\beta_2 = \mu_{t(i)} = 0$). The coefficient β_1 is estimated at 0.417, indicating that a ten percentage point difference in father status translates on average into just over a four percentage point difference in son status (Table 4, column (1)). As typical in intergenerational mobility analysis, this is not a causal coefficient estimate. The implied degree of mobility is plausible in the light of other estimates. In particular, Shiue (2024) estimates a value of 0.52 for an earlier sample, which is consistent because mobility increased over the Ming-Qing period (i.e. lower β_1). Shocks do not matter strongly for this estimate, as can be seen from the similar results with birth year fixed effects ($\beta_1 = 0.410$, column (2)).

[15]The analysis includes a relatively small number of concubinages (unmarried relationships); results without unmarried relationships are similar.

The next specification relates son status to the status of his father-in-law, as in equation (3). The estimated coefficient is equal to 0.423, which is similar in size to the coefficient of 0.410 on the status of the son's biological father (compare Table 4, columns (2) and (3)). Thus, if the goal is to predict son status with a single measure of status in the parent generation, it does hardly matter whether one employs father status or father-in-law status. The main reason is that there is assortative mating. Recall that in the absence of assortative mating, the coefficient on father-in-law status in the son status regression is zero (equation (4)). The estimate of 0.423 in column (3) indicates that the degree of assortative mating is not zero but positive and sizable.[16]

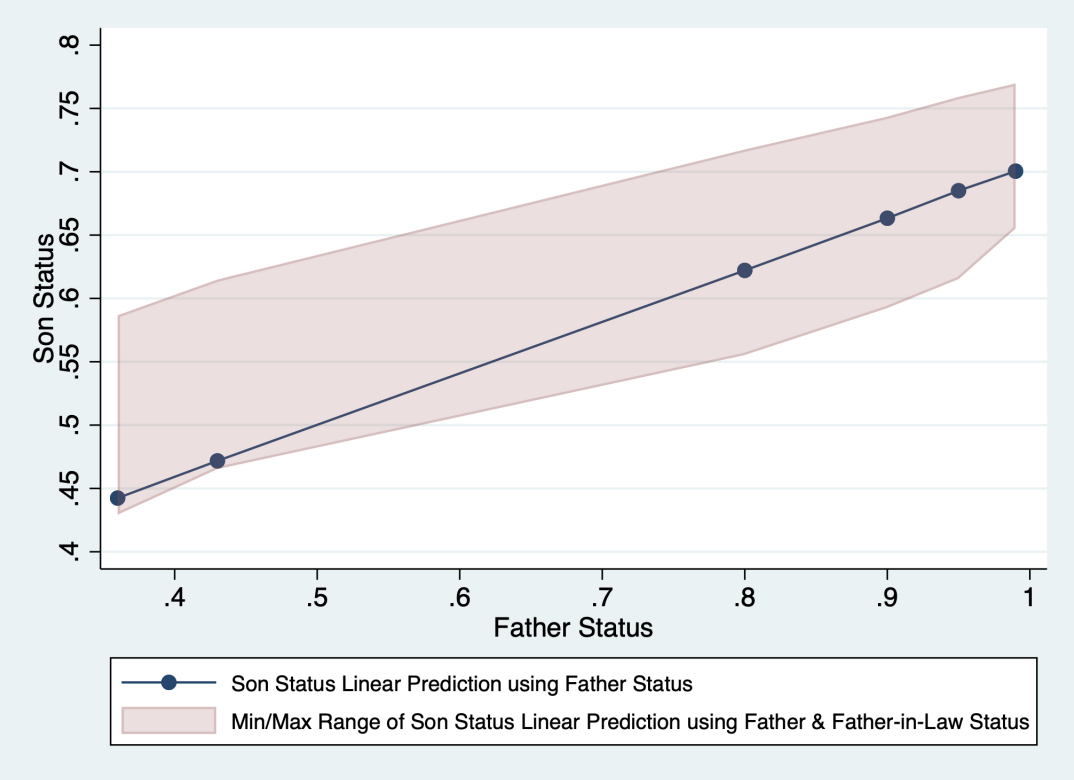
Column (4) shows results when both father and father-in-law variables are included. The father coefficient is now estimated at 0.281, while the father-in-law coefficient is equal to 0.271. Given that the two father variables have the same mean and similar variance, this provides new evidence from China that marriage can be comparable in importance for child status as the status of the biological parent.

By employing equation (8) we can estimate the degree of assortative mating γ from the results of columns (2) and (4). In particular, $0.410 = 0.281 + \gamma\, 0.271$, so that $\gamma = 0.48$. For comparison, in the analysis of a sample of 18th century Murcia, the coefficient of the regression of son on father status is with 0.78 higher, and an estimate of γ for 18th century Murcia is 0.73, see Espin-Sanchez, Gil-Guirado, and Vickers (2022). There may be several explanations, including differences in status inheritability in Spain versus China, however, the results are consistent with higher social mobility and lower marriage matching in Central China compared to Southern Spain.

Figure 3 presents additional evidence on the role of marriage for mobility by comparing the linear prediction of son status based on his biological father alone with the range of son status predictions based on both biological father and in-law father status.

[16]Furthermore, the estimate also indicates that to the extent that the variance in father and father-in-law status is similar, the role of parent status for daughter status, β_d, must be substantial.

Figure 3: Mobility with and without Marriage



Notes: Son status prediction as a function of father status based on columns (2) and (4) of Table 4. Status in terms of percentile rank.

While for the lowest father status the best linear prediction of son status is 0.44, accounting for the role of marriage shows that son status can in fact range between 0.43 (for the lowest in-law status) and 0.58 (for the highest in-law status). Thus, a top marriage is associated with a 14 percentage points son status, or around one third of the typical son status with such a low-status father. That the best linear prediction is close to the lower end of the range reflects that top marriages are unlikely for sons of fathers who have low status.

This is different for sons of higher status fathers. For example, a son of a father at the 80th percentile can rise *or* fall in status through his marriage by roughly equal amounts; the linear prediction is 0.63 while the range of high versus low status is 0.72 to 0.56, respectively. Further, while the range in predicted son status determined by marriage of about 15 percentage points is similar for most levels of father status, at the very top of the status distribution the range is more compressed (about 11 percentage points). This reflects a relatively high level of positive marital sorting at the top: daughters of the highest status, that is, *jinshi* daughters, tend to marry their equal in status–*jinshi* sons–at a higher rate than assortative mating occurs at lower status levels. Figure 3 provides new evidence not only that marriage plays an important role for intergenerational mobility but also that this role varies across the distribution of status in society.

In addition to mobility patterns, marriage also affects the overall level of mobility. Notice that

the sum of estimated coefficients of father status and father-in-law status is 0.55 (column (4)). If son A has both a father with 10 percentage points higher status than another son B and a wife that comes from a family with a 10 percentage points higher status than the family of the wife of husband B comes from, son A can expect to attain a 5.5 percentage point higher status, in contrast to 4.1 percentage points based only on father status (column (2)). This is confirmed by employing the average of father and father-in-law status as the single regressor, see column (5). Thus, accounting for marriage and the extent of marital sorting implies a considerably lower level of intergenerational parent-child mobility compared to an analysis based on biological family only.

Further analysis shows that the results of Table 4 are robust to accounting for clan effects, both of the seven male clans as well as the more than 120 clans of the in-marrying women; see Table A.2. Table A.3 shows that while overall the role of father and father-in-law status are similar in subperiods before and after 1750, it is possible that their influences change over time in opposite directions. In particular, there is evidence that the pro-mobility decline of β_1 (father status) is offset by an increase in β_2 (father-in-law status), see Table A.3.

4.2 Evidence from Brothers

Our analysis of the role of marriage for parent-child mobility so far has exploited variation both across biological and across in-law families. Some husbands are the sons of higher status fathers while others are not, at the same time when certain men become the husbands of daughters from high-status families while others do not. Additional insight comes from focusing on brothers, which is interesting because brothers share to some extent the family's biological and other resource endowments.[17] Restricting the analysis to brothers sharpens the focus on the role of marriage for intergenerational mobility.

We limit our sample in two ways. First, we focus on sons who have one or more brothers who are also recorded with their marriage in the clan's genealogy. Thus, brothers who do not grow up to adulthood and marry are excluded from the analysis as we do not observe their wife. Second, we include only sons who were married exactly once in their lifetime. That way we do not have to make an additional assumption on which marriage matters– is it the first, the average, or the marriage with the maximum in-law status, for example. With these sample restrictions, there are more than 2,400 observations, of which most are for the case of two to four brothers.[18]

The analysis employs a variant of equation (10) above, given by

$$stat_{si} = \beta_0 + \beta_2 stat^f_{di} + \mu_{t(i)} + \zeta^f_i + \varepsilon_{t(i)}, \tag{11}$$

which differs from equation (10) by the inclusion of father fixed effects, ζ^f_i. Given the inclusion of father fixed effects, the analysis exploits only "within father" variation, that is, variation among

[17] Of course, the genetic endowment of brothers is not the same unless they are monozygotic twins, and even then these brothers might not be treated identically by the family for various reasons.

[18] This includes also half-brothers in the sense that the sons have the same father but not the same mother. Half-brothers are distinguished from full brothers in Table A.4.

brothers. At the same time, the inclusion of father fixed effects implies that the coefficient on father status, β_1, is not separately identified. Table 5 shows the results, with different columns reporting results for an increasingly larger sample: two brothers in column (1) all the way to up to nine brothers in column (8). Specifically, there are N = 202 observations for which the number of brothers is exactly two, there are N = 884 observations with up to three brothers, and so on.

Table 5: Marriage and Brother Mobility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
No. of Brothers	2	< 4	< 5	< 6	< 7	< 8	< 9	< 10
Father-in-Law Status	0.462*	0.234**	0.176**	0.121**	0.083*	0.080*	0.078*	0.078*
	(0.186)	(0.079)	(0.055)	(0.045)	(0.039)	(0.039)	(0.038)	(0.038)
N	202	884	1,527	2,015	2,271	2,378	2,421	2,429

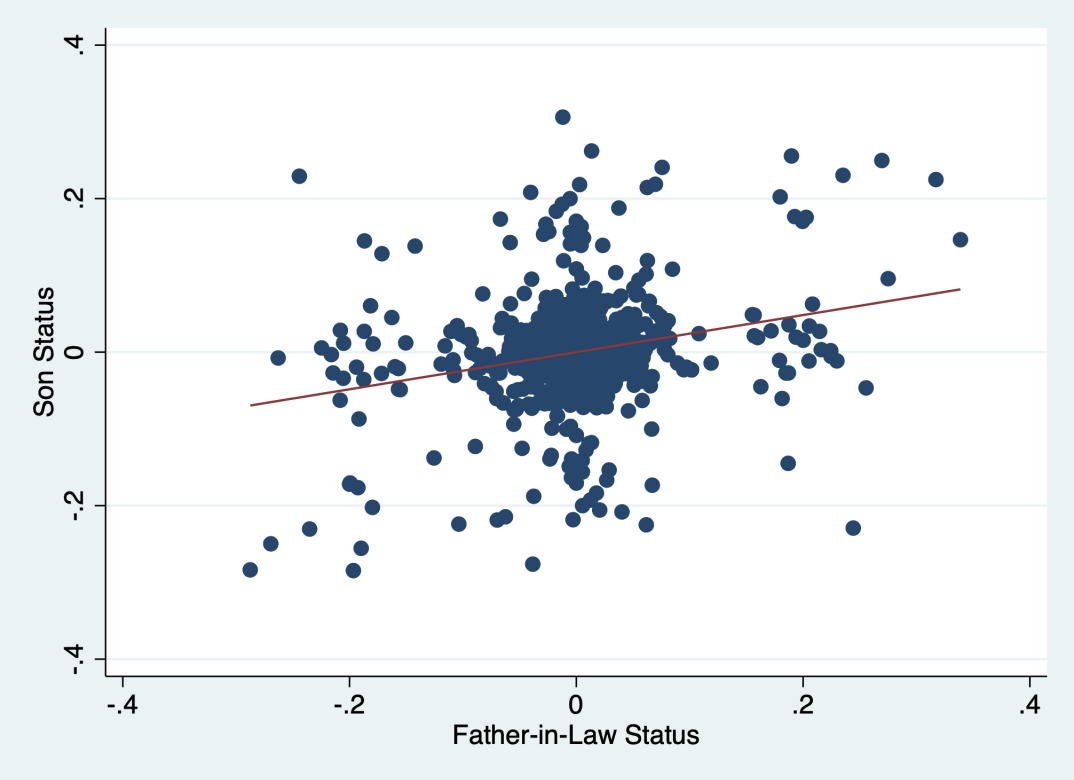
Notes: Dependent variable son status as percentile rank. Estimation of equation (11); robust standard errors clustered on biological father reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

The father-in-law status coefficient sheds light on the status difference between sons of the same father for any given difference in the status of the wives they marry (based on the status of their fathers). We estimate $\beta_2 = 0.462$ in the case of two brothers (column (1)). Thus, if two brothers marry wives whose fathers have a ten percentage points status difference, the brother with the higher status wife can expect to have 4.62 percentage points higher status than his brother marrying the lower-status wife. Put differently, the status of the in-law family accounts on average for close to half of the status difference between two brothers.

It is worth noting that the estimated father-in-law status coefficient does not fall, rather, it increases when we focus on within-father effects (compare Table 4 with Table 5, respectively). Status differences among brothers tend to be smaller than status differences in the population overall because among brothers, there is less biological variation. In particular, the standard deviation of percentile rank status among two brothers is less than one sixth of the standard deviation of status of all sons. Of the remaining variation in the status between two brothers, marriage differences account for almost half of the husband status differences, in contrast to a quarter in the sample overall, see Table 4.

Figure 4 shows the relationship between son status and father-in-law status with up to three brothers (based on column (2) in Table 5). There is a clear positive relationship.

Figure 4: Which Son Rises? Intergenerational Mobility and The Marriages of Brothers



Notes: Son and Father-in-Law Status in percentile rank. Values demeaned by (1) father and (2) son birth year; based on specification shown in column (2), Table 5.

The relationship in Figure 4 may not be causal. It could be that the reason why one brother marries better is that he is smarter, for example. Alternatively, the parents could invest more strongly into the marriage of the first-born son, which may translate into a higher-status wife as well.[19] As in the typical intergenerational analysis, causal statements would require additional empirical strategies. Figure 4 simply shows that the brother who marries a high-status woman tends to have high-status himself, while the brother who marries a low-status woman has typically low status. The correlation between the two variables in Figure 4 is 0.25.

Table 5 shows also that as the number of brothers is increased in the analysis, the coefficient on father-in-law status falls. One explanation is that the influence of idiosyncratic shocks increases as the number of brothers goes up. In the case of two brothers, the question is simply whether the brother with the wife of higher status has higher status himself. In contrast, with a rising number of brothers and given some idiosyncratic shocks, the chance of a positive correlation of father-in-law and son status falls. We have also distinguished half-brothers from full brothers, finding that it does not lead to different results (see Table A.4). Heterogeneity analysis is performed to examine whether the role of marriage for intergenerational mobility varies with characteristics of (1) the son,

[19]Examining whether the son's birth order affects the role that marriage plays for husband status, we find little empirical evidence for it, see Table A.5.

(2) his father, or (3) his mother. Results are in line with historical evidence and robust, see Table A.5, Table A.6, and Table A.7.

To summarize these results, from the point of view of the son's family it is useful to think of father-in-law status as a measure of marriage quality. We have seen that a 10 percentage point difference in marriage quality is typically associated with a 2.5 percentage point difference in son status, and a 10 percentage point difference in marriage quality is associated with an up to 5 percentage point difference in status among brothers. These findings provide new evidence on the importance of marriage for intergenerational mobility. In the following section, we extend the analysis to multiple generations.

4.3 Mobility through Marriage over Multiple Generations

Mobility analysis has increasingly moved beyond the two-generation parent-child analysis by considering the influence of grandparents. The following performs such an analysis in the context of the study of the role of marriage for multigenerational mobility. Consider the following specification:

$$stat_{si} = \beta_0 + \beta_1 stat^{f}_{si} + \beta_2 stat^{f}_{di} + \beta_3 stat^{gf}_{si} + \beta_4 stat^{gfm}_{si} + \mu_{t(i)} + \varepsilon_i, \tag{12}$$

where $stat^{gf}_{si}$ is the percentile rank status of the son's paternal grandfather, and $stat^{gfm}_{si}$ is the percentile rank status of the son's maternal grandfather. As before, the coefficient β_2 provides evidence on the importance of marriage because it quantifies the influence of the in-law family's status for the status of the son. At the same time, equation (12) estimates the role of a second marriage for son status, namely that of his father to his mother, by including the percentile rank status of the son's maternal grandfather.

Thus, one can think of equation (12) as shedding light on the role of marriage in generation (0) and in generation (-1), which are respectively between son and daughter on the one hand, and between father and mother on the other hand. While both coefficients β_2 and β_4 shed light on the role of marriage for intergenerational mobility, there is a difference. Son status may not only be affected by the resources of his maternal grandfather but also through genetic transmission through his mother, in contrast to the son's own marriage which does not affect his genetic endowment. In that sense, the generation-0 coefficient β_2 estimates the role of a more narrow aspect of marriage for mobility compared to the generation (-1) coefficient β_4. Results are shown in Table 6.

Table 6: Marriage and Multigenerational Mobility

	(1)	(2)	(3)	(4)
Father Status	0.409**	0.318**	0.240**	0.226**
	(0.018)	(0.020)	(0.021)	(0.021)
Paternal Grandfather Status		0.154**	0.096**	0.074**
		(0.015)	(0.015)	(0.016)
Father-in-Law Status			0.236**	0.214**
			(0.022)	(0.023)
Maternal Grandfather Status				0.075**
				(0.019)
N	4,972	4,972	4,972	4,972

Notes: Dependent variable is son status; all status variables in percentile rank; all specifications include son birth year fixed effects. Estimation of equation (12); robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

We introduce additional variables into equation (12) one at a time for comparability with the existing literature even though we have seen above that the status of the father in law matters for son status. The first specification of Table 6 presents the earlier finding that in a regression of son status on the status of his biological father, the coefficient is about 0.41. Next, we turn this into a multigenerational mobility analysis by including a measure of grandfather status. The coefficient for the paternal grandfather is $\beta_3 = 0.154$, which confirms other research that grandfathers play a role for intergenerational mobility even condition on father status (Table 6, column (2)). Note that the inclusion of the paternal grandfather reduces the size of the father coefficient, which reflects the fact that some differences that account for variation in son status are common to both his father and paternal grandfather. In addition, the sum of father and grandfather point estimates in column (2) exceeds the point estimate in the simple parent-child specification ($0.472 = 0.318 + 0.154$, versus 0.409), or by 15% ($= \frac{0.472}{0.409} - 1$).

The next specification adds the father-in-law status variable, and it is estimated with a coefficient of 0.236 (see Table 6, column (3)). As before, the point estimate for father-in-law status is similar in size to the point estimate of the biological father, confirming that marriage plays an important role for mobility. Moreover, the point estimate on father-in-law status is more than twice as large as the paternal grandfather coefficient. In the final specification of Table 6 we add the status of the mother's father to quantify the role for son mobility of marriage in the previous generation. Mother's father status enters with a coefficient of $\beta_4 = 0.075$. This indicates that not only current marriages but also past marriages play a significant role for mobility.

How important is the marriage in the previous generation compared to the marriage in the present generation? Based on the coefficients in column (4), the marriage of the past generation to

current generation matters roughly by a factor of 1:3 ($\approx \frac{0.075}{0.214}$). It is interesting to note that in the final specification, the extent to which paternal grandfather affects son status relative to the influence of his father is also approximately 1:3 ($\approx \frac{0.074}{0.226}$). Although our marriages take place in generations (0) and (-1) while fathers and grandfathers belong to generations (-1) and (-2), respectively, this suggests that the intergenerational decay of biological and marriage influences on mobility may be comparable. We have noted earlier that instead of capturing the influence of marriage in generation (-1), one can also see coefficient β_4 as quantifying the influence of another member of generation (-2), the maternal grandfather. Along these lines, note that in this sample the roles of paternal and maternal grandfather are estimated to be very similar (0.074 versus 0.075, respectively).

Overall, the results on mobility through marriage over multiple generations shown in Table 6 suggest that incorporating marriage in the analysis of intergenerational mobility is important. We have confirmed that not accounting for past generations beyond the father generation leads to an overestimate of mobility by 15% (columns (1) and (2) of Table , see Table 6). If we add to this the extent to which son mobility is affected by past and current marriage, mobility based on the simple father-son relationship is overestimated by 44%.[20]

This suggests that accounting for the role of marriage for mobility maybe more important than including higher-order generations of the own biological family.[21] Even if one treats the maternal grandfather entirely as an influence of the generation (-2) – as noted above, coefficient β_4 captures also the influence of past marriages– the sum of the point estimates of paternal and maternal grandfather is not as large as the point estimate capturing the role of current marriage ($\beta_2 + \beta_4 = 0.149$, versus $\beta_3 = 0.214$, column (4) of Table 6). In sum, for estimating mobility in a society, our results indicate that more attention should be given to account for unions between different families through marriage.

Broadly similar results are obtained in specifications that include clan fixed effects and that allow for results by cohort, see Table A.8 and Table A.9. We now turn to some conclusions.

[20] Computed as $\frac{0.226+0.074+0.214+0.075}{0.409} - 1 = \frac{0.589}{0.409} - 1 = 0.44$, see columns (4) and (1).

[21] In line with this are results of studying mobility with up to five paternal generations (Shiue 2024).

5 Conclusions

We have examined the role of marriage for multigenerational mobility based on a sample of three generations of married men with their families Anhui province's Tongcheng county. We find that father-in-law status is highly predictive of son status, with a coefficient that is similar in size to that of the son's biological father. Key to this is assortative mating, with a correlation of biological and in-law family status of just under 0.5.

Evaluating the roles of biological and in-law family side by side, we find that both enter with a positive coefficient that is similar in size. The role that marriage plays for son status is to some extent independent of that of the son's biological father, and it is quantitatively important. For most sons, the typical range of status for most successful to least successful marriage is about fifteen percentile ranks. Marriage influences even the mobility of brothers, who have a relatively large part of their genetic and resource endowment in common. We find that half of any difference in marriage quality of two brothers translates into status differences of these brothers over their lifetimes. This finding highlights that marriage matters for intergenerational mobility even conditional on the resources of the family.

We also find that not only a man's current marriage but also the marriage of his parents plays a role for his mobility. Often, the in-marrying women of different generations will not be genetically related; only their children will have genetic endowments from both parents. Marriage of father to mother affects father status only through non-genetic endowments of the in-law family, whereas for their son the past marriage plays a role both through genetic and non-genetic endowments of the mother's in-law family. Our finding that the role of marriage for mobility is larger than the influence of the grandparents suggests to investigate the role of other groups for mobility, not only the clan but also networks based on geographic proximity.

References

[1] Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. Perez (2021), "Automated Linking of Historical Data", Journal of Economic Literature Vol. 59(3): 865-918.

[2] Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos (2011), "Marrying Up: The Role of Sex Ratio in Assortative Matching", American Economic Journal: Applied Economics Vol. 3(3), pp. 124-157.

[3] Ahern, Emily M. (1974), "Affines and the Rituals of Kinship", in Religion and Ritual in' Chinese Society, ed. Arthur P. Wolf. Stanford: Stanford University Press.

[4] Althoff, Lukas, Harriet Brookes Gray, and Hugo Reichardt (2023), "The Missing Link(s): Women and Intergenerational Mobility", Yale University, December.

[5] Bailey, Martha J., Connor Cole, Morgan Henderson, and Catherine Massey (2020), "How Well Do Automated Linking Methods Perform? Lessons from US Historical Data", Journal of Economic Literature 58(4): 997-1044.

[6] Bailey, Martha J., and Peter Z. Lin (2022), "Marital Matching and Women's Intergenerational Mobility in the Late 19th and Early 20th Century U.S.", UCLA working paper, November.

[7] Bailey, M., Lin, P. Z., Mohammed, A. R. S., Mohnen, P., Murray, J., Zhang, M., & Prettyman, A. (2023), "The Creation of LIFE-M: The Longitudinal, Intergenerational Family Electronic Micro-Database Project", working paper, June.

[8] Bavel, Jan Van, Hilde Peeters, and Koen Matthijs (1998), "Connections between Intergenerational and Marital Mobility: A Case Study: Leuven, 1830-1910, Historical Methods Vol. 31 (3), pp. 122-134.

[9] 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.

[10] Becker, G. S. (1974). A Theory of Marriage: Part II. Journal of Political Economy, 82(2, Part 2), S11–S26.

[11] Becker, G. S. (1973). A Theory of Marriage: Part I. Journal of Political Economy, 81(4), 813–846.

[12] Braun, Sebastian Till, and Jan Stuhler (2018), "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany", *Economic Journal* 128, pp. 576-611.

[13] Buckles, Kasey, Adrian Haws, Joseph Price, and Haley Wilbert (2023), "Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project", NBER # 31671.

[14] Buckles, Kasey, Joseph Price, Zach Ward, and Haley Wilbert (2023), "Family Trees and Falling Apples: Intergenerational Mobility Estimates from U.S. Genealogy Data", paper presented at the NBER Summer Institute, July 2023.

[15] Bull, Hans Henrik (2005), "Deciding Whom to Marry in a Rural Two-Class Society: Social Homogamy and Constraints in the Marriage Market in Rendalen, Norway, 1750–1900", International Review of Social History Vol. 50, Supplement 13, pp. 43-63.

[16] Cao Shuji. 2000. *Zhongguo Renko Shi Mingqing Shiqi* (A Demographic History of China, the Qing Period), vol. 5. Shanghai: Fudan University Press.

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

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

[19] Chadwick, Laura, and Gary Solon (2002), "Intergenerational Income Mobility Among Daughters", American Economic Review 92(1), pp. 335-344.

[20] 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", The Quarterly Journal of Economics 129 (4): 1553-1623.

[21] Clark, Gregory (2014), The Son Also Rises, Princeton University Press.

[22] Choo, Eugene, and Aloysius Siow (2006), "Who is Marrying Whom and Why", Journal of Political Economy Vol. 114, No. 1, pp. 175-201.

[23] Craig, Jaqueline, Katherine Eriksson, and Gregory T. Niemesh (2019), "Marriage and the Intergenerational Mobility of Women: Evidence from Marriage Certificates 1850-1910", working paper.

[24] Curtis, Matthew (2022), "The her in inheritance: how marriage matching has always mattered, Quebec 1800-1970", ECARES working paper, September.

[25] Dahl, Molly, and Thomas DeLeire (2008), "The Association between Children's Earnings and Fathers' Lifetime Earnings: Estimates Using Administrative Data", University of Wisconsin-Madison, August.

[26] Dong, Hao, and Yu Xie (2023), "Trends in Educational Assortative Marriage in China Over the Past Century", *Demography* Vol. 60(1), pp. 123-145.

[27] Dribe, Martin, Bjorn Eriksson and Francesco Scalone (2019), "Migration, marriage and social mobility: Women in Sweden 1880-1900", Explorations in Economic History 71(C), pp. 93-111.

[28] Dribe, Martin, and Christer Lundh (2005), "Finding the Right Partner: Rural Homogamy in Nineteenth-Century Sweden", International Review of Social History Vol. 50, Supplement 13, pp. 149-177.

[29] Ebrey, Patricia Buckley (2023), "Women in Traditional China", Asia Society, New York, NY, https://asiasociety.org/education/women-traditional-china, accessed Aug 17, 2023

[30] Ebrey, Patricia Buckley (1991), "Shifts in Marriage Finance from the Sixth to the Thirteenth Century." In Watson, Rubie S. and Patricia Buckley Ebrey, Eds. *Marriage and Inequality in Chinese Society.* Los Angeles, CA: University of California Press.

[31] Eberhard, Wolfram (1962), *Social Mobility in Traditional China.* Leiden: E.J. Brill.

[32] Espin-Sanchez, Jose-Antonio, Joseph Ferrie, and Chris Vickers (2023), "Women and the Econometrics of Family Trees", NBER working paper # 31598.

[33] Espin-Sanchez, Jose-Antonio, Salvador Gil-Guirado, and Chris Vickers (2022), "La "Doña" è Mobile: The Role of Women in Social Mobility in a Pre-Modern Economy", Journal of Economic History 82(1): 1-41; doi: 10.1017/S0022050722000018

[34] Fang Xu-ling. 2010. "Qingdai Tongcheng Jinshi Shuliangkao" (A study of the number of successful jinshi candidates from Tongcheng during the Qing Period), *Journal of Anqing Teacher's College*, 29(5).

[35] Fei, H.T. 1946. "Peasantry and gentry: an interpretation of Chinese social structure and its change. *American Journal of Sociology.* 52(1).

[36] Fernández, Raquel (2013), "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century", American Economic Review 103: 472-500.

[37] Ferrie, J. P. (1996), "A New Sample of Males Linked from the 1850 Public Use Micro Sample of the Federal Census of Population to the 1860 Federal Census Manuscript Schedules", Historical Methods 29(4): 141-156.

[38] Freedman, Maurice (1979), The Study of Chinese Society, ed. G. William Skinner. Stanford: Stanford University Press.

[39] Goldin, Claudia (1983), "The changing economic role of women: A quantitative approach", The Journal of Interdisciplinary History 13 (4), 707–733.

[40] Guell, Maia, Jose V. Rodriguez Mora, and Christopher I. Telmer (2015), "The Informational Content of Surnames, the Evolution of Intergenerational Mobility, and Assortative Mating", Review of Economic Studies Vol. 82(2), pp. 693-735.

[41] Haller, Max (1981), "Marriage, Women, and Social Stratification: A Theoretical Critique", American Journal of Sociology Vol. 86(4), pp. 766-795.

[42] Hamilton, Gillian, and Aloysius Siow (2007), "Class, Gender and Marriage," Review of Economic Dynamics, vol. 10(4), pages 549-575, October.

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

[44] Hess, Christoph (2023), "Family and Community Constraints on Migration in Pre-Industrial China", Association for Asian Studies 2024 Annual Conference.

[45] Ho, Ping-ti (1967*), The Ladder of Success in Imperial China: Aspects of Social Mobility (1368-1911)*. New York: Columbia University Press.

[46] Hu, Sijie (2023), "Survival of the literati: Social status and reproduction in Ming–Qing China", Journal of Population Economics https://doi.org/10.1007/s00148-023-00960-2 .

[47] Jácome, Elisa, Ilyana Kuziemko, and Suresh Naidu (2022), "Mobility for All: Representative Intergenerational Mobility Estimates over the 20th Century", NBER # 29289, September.

[48] Jäntti, Markus, Bernt Bratsberg, Knut Roed, Oddbjorn Raaum, Robin Naylor, Eva Osterbacka, Anders Bjorklund, and Tor Eriksson (2006), "A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States", IZA Discussion Paper # 1938, January.

[49] Kalmijn, Matthijs (1994), "Assortative Mating by Cultural and Economic Occupational Status", American Journal of Sociology, Vol. 100, No. 2 (Sep., 1994), pp. 422- 452.

[50] Keller, Wolfgang, and Carol H. Shiue (2022), "Marriage Matching over Five Centuries in China", NBER # 30625.

[51] Ko, Dorothy (1994), *Teachers of the Inner Chambers: Women and Culture in Seventeenth-century China.* Stanford University Press.

[52] Lam, David, and Robert F. Schoeni (1994), "Family Ties and Labor Markets in the United States and Brazil", Journal of Human Resources Vol. 29(4), pp. 1235-1258.

[53] Long, Jason, and Joseph Ferrie (2018), "Grandfathers Matter(ed): Occupational Mobility across three Generations in the US and Britain, 1850-1911", *Economic Journal* Vol. 128, F422-F445.

[54] Mann, Susan. (1997), *Precious Records: Women in China's Long Eighteenth Century.* Stanford University Press.

[55] Mann, Susan (1991), "Grooming a Daughter for Marriage: Brides and Wives in the Mid-Qing Period", in Watson and Ebrey (eds.), *Marriage and Inequality in Chinese Society.* Los Angeles, CA: University of California Press.

[56] Mare, Robert (2011), "A Multigenerational View of Inequality", Demography Vol. 48, pp. 1-23.

[57] Mare, Robert (1991), "Five decades of educational assortative mating", American Sociological Review 56: 15-32.

[58] Mazumder, Bashkar (2005), "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data", Review of Economics and Statistics Vol. 87(2), pp. 235-255.

[59] Nunn, Nathan (2020), "The historical roots of economic development", Science 367, eaaz9986 (2020). DOI: 10.1126/science.aaz9986

[60] Olivetti, Claudia (2013), "The Female Labor Force and Long-run Development: The American Experience in Comparative Perspective", November 2013.

[61] Olivetti, Claudia, and M. Daniele Paserman (2015), "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850 –1940", American Economic Review 105(8): 2695-2724.

[62] Olivetti, Claudia, and M. Daniele Paserman, Laura Salisbury, and E. Anna Weber (2022), "Who married, (to) whom, and where? Trends in marriage in the United States, 1850-1940", working paper, Boston University, October.

[63] Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Issac Riley (2021), "Combining family history and machine learning to link historical records: The Census Tree data set", Explorations in Economic History 80, https://doi.org/10.1016/j.eeh.2021.101391

[64] Santavirta, Torsten, and Jan Stuhler (2024), "Name-Based Estimators of Intergenerational Mobility", CEPR Discussion Paper # 18740; forthcoming, Economic Journal.

[65] Shan, Chun (2013), "Marriage law and Confucian Ethics in the Qing dynasty", Frontiers of Law in China, Vol. 8, Issue 4, December.

[66] Shiue, Carol H. (2024), "Social Mobility in the Long Run: A Temporal Analysis of Tongcheng, China from 1300 to 1900", forthcoming, Journal of Economic History.

[67] Shiue, Carol H. (2017), "Human Capital and Fertility in Chinese Clans before Modern Growth", Journal of Economic Growth Vol. 22(4).

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

[69] Shiue, Carol H., and Wolfgang Keller (2024), "Human Capital Strategies for Big Shocks: The Case of the Fall of the Ming", February, University of Colorado-Boulder.

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

[71] Stelter, Robert, and Diego Alburez-Gutierrez (2022), "Representativeness is crucial for inferring demographic processes from online genealogies: Evidence from lifespan dynamics", Proceedings of the National Academy of Sciences of the USA, 2022 Mar 8; 119(10): e2120455119. doi: 10.1073/pnas.2120455119

[72] Telford, Ted A. (1992), "Covariates of Men's Age at First Marriage: The Historical Demography of Chinese Lineages", Population Studies Vol. 46, pp. 19-35.

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

[74] Wang, Jing (2017), "Qing dai Tongcheng jinshi yanjiu" (Research on the jinshi from Tongcheng), Ph.D. Dissertation. Zhejiang University.

[75] Ward, Zach (2023), "Intergenerational Mobility in American History: Accounting for Race and Measurement Error", January; forthcoming, American Economic Review.

[76] Watson, Rubie S. (1981). "Class Differences and Affinal Relations in South China." *Man* 16: 593-615.

[77] Watson, Rubie (1985), *Inequality among Brothers, Class and Kinship in South China.* Cambridge University Press.

[78] Watson, Rubie S. and Patricia Buckley Ebrey, Eds. (1991), *Marriage and Inequality in Chinese Society.* Los Angeles, CA: University of California Press.

[79] Weiss, Yoram (1997), "The formation and dissolution of families: Why marry? Who marries whom? And what happens upon divorce", Chapter 3 in *Handbook of Population and Family Economics*, Volume 1A, (eds.) M. Rosenzweig and O. Stark, Elsevier Publishers, pp. 81-123.

[80] Wolf, Arthur P. and Chieh-Shan Huang (1980), *Marriage and Adoption in China, 1845-1945.* Stanford: Stanford University Press.

[81] Zhao, Zongwei (1997) "Long-Term Mortality Patterns in Chinese History: Evidence from a Recorded Clan Population." *Population Studies*, 51(2): 117-127.

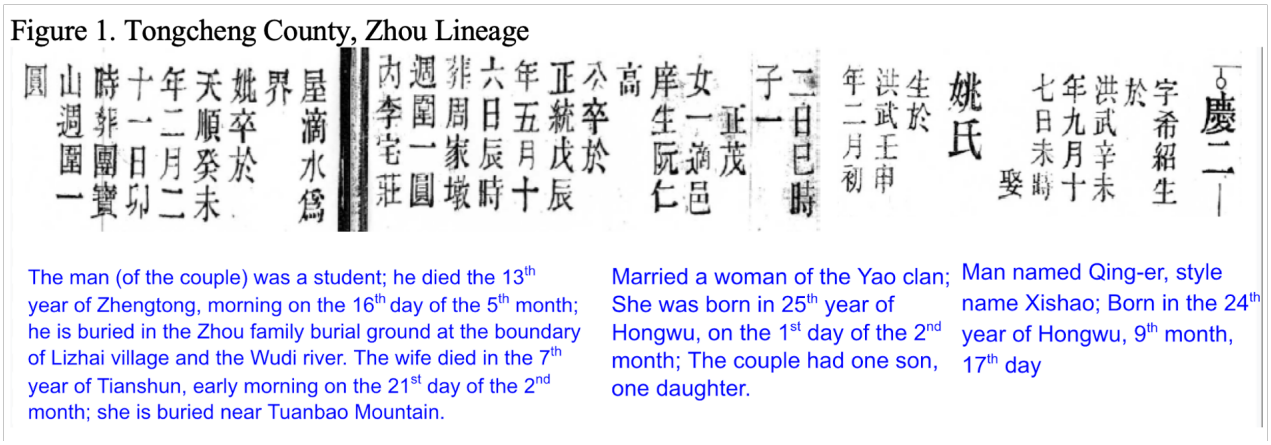
Appendix

A Data

A.1 Example of Tongcheng Genealogical Data

Figure A.1 presents part of one of the genealogies employed in this paper.

Figure A.1: Zhou Clan of Tongcheng: Example



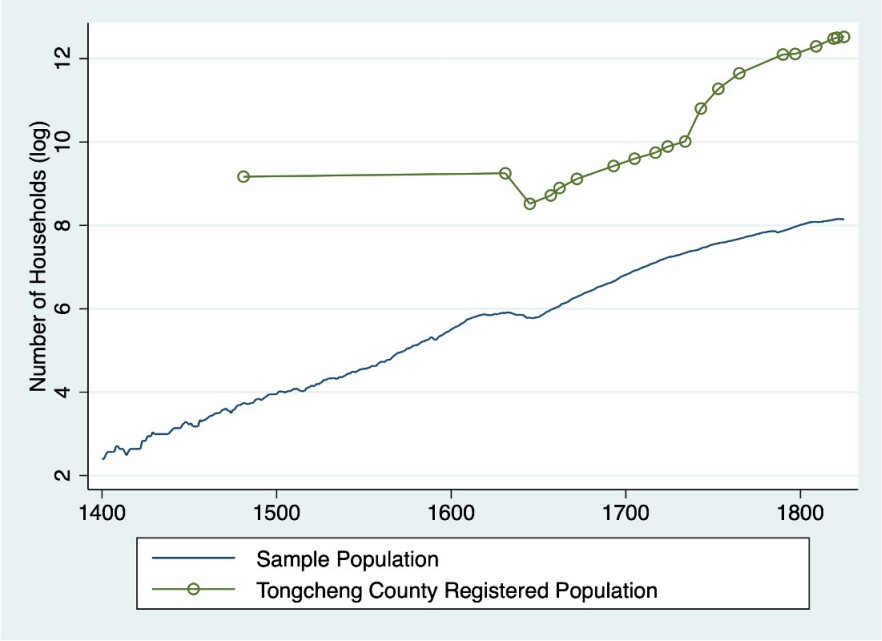
Notes: Part of the source data, see text for additional information.

A.2 Sample Construction and Characteristics

The sample was created with a targeted approach that selected a subset of genealogies from Tongcheng with the explicit goal of obtaining a broadly representative sample. Seven clans have been chosen for inclusion into the sample, with the following last names: Chen, Ma, Wang, Ye, Yin, Zhao, and Zhou. Conditional on choosing the genealogy of a given clan, all entries of the genealogy become part of the sample. The seven clans' genealogies are also relatively complete in terms of recording vital information. It is well-known that Chinese genealogies vary in the degree of completeness. Among the 75 genealogies from Anhui provinces surveyed by Telford (1986), for example, there is vital information on both husband and wife in 76% of the cases. In the present estimation sample, there is vital information for 79.5% of the husbands and 79.1% of the wives, and we have confirmed that our results do not depend on estimation of vital statistics.

Why would one use private genealogical data instead of official data? The reason is simply that there is no high-quality official data yet for our sample period. Figure A.2 shows data for Tongcheng county's overall population that is available from official sources ("registered" population from tax registers). Official data is only available for 21 years. Given this sparsity, as well as virtually unchanged levels for more than a century during a period for which genealogies indicate steady population growth suggests that official population data for this period in Tongcheng is, at a minimum, incomplete.

Figure A.2: Sample and Total Tongcheng Population over Time



Notes: Lower series is sample household population, measured as log number of male heads-of-households (N = 426); upper series is log number of households, Beattie (1979), Table 3 (N = 21). Also shown the fraction of the sample based on these series; scale on the right.

The limited availability of reliable official data for this period means that one cannot establish the representativeness of the sample by making comparisons to official data for all of Tongcheng. Instead, we rely on broader comparisons wherever data is available, at the national or regional level, and for subsets of our sample period.

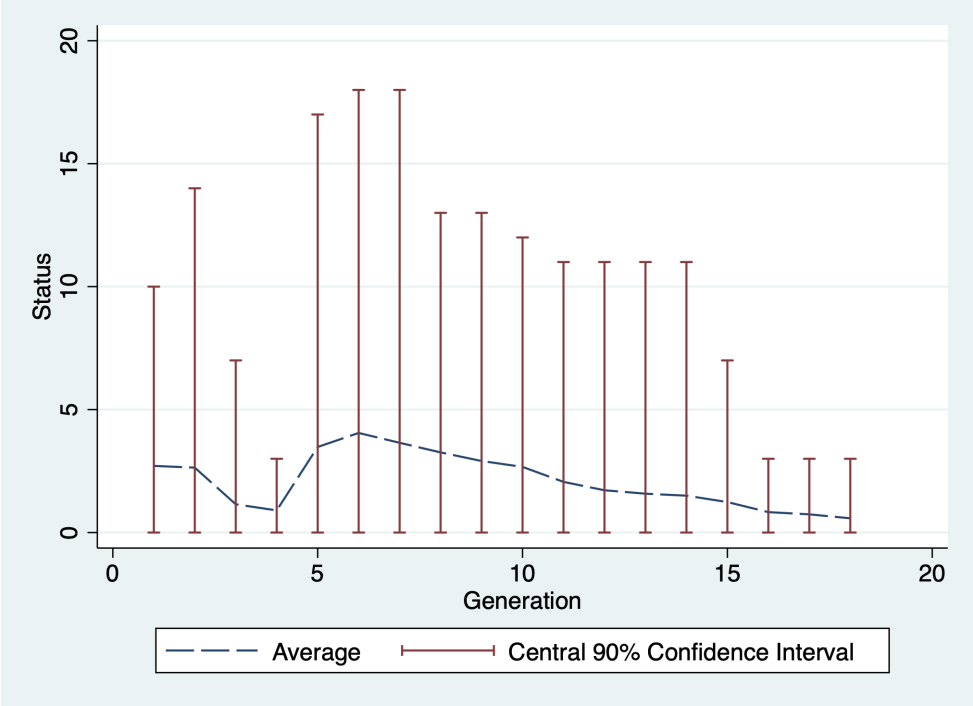
A.3 Sample Representativeness

A.3.1 Comparisons to Other Statistics

A first check is to consider mortality rates by age group. Population figures at the regional level are typically based on gazetteers (local histories about a certain place). Telford (1990) compares demographic patterns in the Tongcheng genealogical data and the Eight Banner populations for 1774 to 1873, when the latter starts to become available. He finds a very similar variation in the probability of dying for different age categories across the two sources (see Telford 1990, Figure 2).

The sample consisting of these seven clans matches national shares of upper status population as defined by Fei (1946) and Chang (1955). One might also ask how the present sample compares in terms of the very top status groups to other evidence. In his classic study based on national lists of *jinshi*, which are extremely reliable, Ho (1967) reports that during the Qing in Anhui there were 41 *jinshi* per one million population, or, 0.0041 percent. There were regional variations, and the province of Anhui was below the provincial average in terms of *jinshi* per capita in Qing China (Ho 1967, p. 228). In the Tongcheng sample, there were a total of 14 *jinshi* during the Qing, which is

Figure A.3: Clan Status by Generation



Notes: Shown are clan average status and the central 90% confidence interval of clan status for the first to eighteenth generation.

about 0.045 percent of the population in the data. Thus, there are about ten times more *jin-shi* in the sample than in Qing Anhui overall.

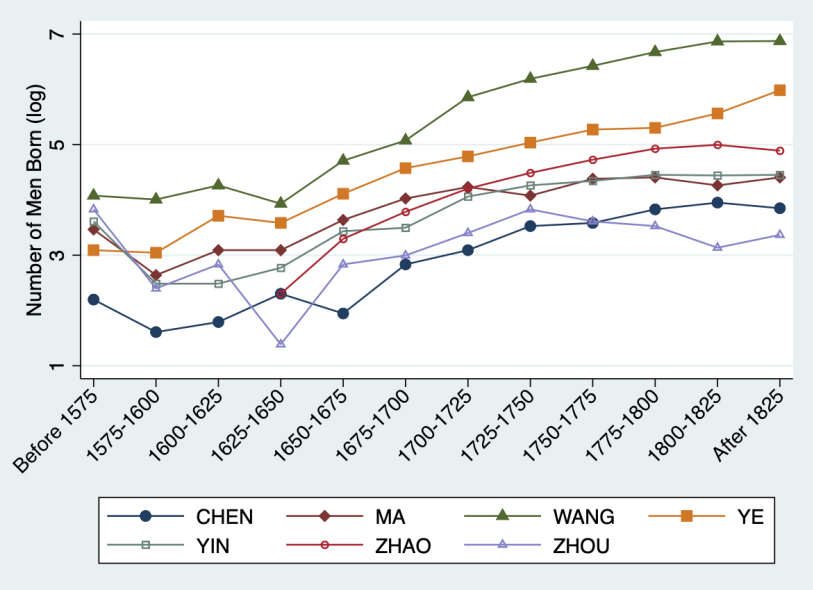
At the same time, *jinshi* were rare, with many parts of Anhui province not producing a single *jinshi* over centuries. Also, other regions of China, in particular in Zhejiang and Jiangsu province, produced *jinshi* by an order of magnitude higher than Tongcheng (see Ho 1967). Overall, Tongcheng was a noteworthy place at a local, perhaps provincial level, but it was not an unusual Chinese region. In addition, the list of people who are recorded in the sample as *jinshi* can be compared against other lists of *jinshi* degree holders from the Chinese state (see Fang 2010; Cao 2016; Wang 2017). We have verified that the information on the *jinshi* in the sample is consistent with the information of these official lists.

A.3.2 Inception of the Genealogy

One might ask why the clan decides to establish its genealogy at a particular time. If the genealogy of a clan is established because one clan member had achieved extraordinarily high status, this might affect intergenerational mobility estimates. To examine this, Figure A.3 shows average clan status by generation.

Figure A.3 shows that average clan status does not monotonically fall across generations following the inception of the genealogy. In particular, average clan status from the 5th to the 10th generation is higher than in the first generation. Thus, the role of the primogenitor of each clan is limited in

Figure A.4: Sample Size over Time



Notes: Shown is the size of each clan as measured by the number of men born in each of the twelve birth cohorts shown.

our sample.

A.3.3 Clan Representation Over Time

Figure A.4 shows the size of each clan in terms of the number of men born in each birth cohort. All clans are present during the period from 1575 to 1825.

In general, size grows for all clans, although growth rates vary by period; for example, the growth rate is negative for three clans in the birth cohort 1625-1650, which includes the dynastic transition from the Ming to the Qing (1644). Some clans grow faster than others, however, sustained overtaking in terms of clan size is rare.

A.3.4 Recall Bias

Other concerns result from the retrospective nature of how genealogies are compiled, in particular recall bias. Genealogies are typically updated after every two or three generations. One might be concerned that the updating is correlated with the particularly high status of a clan member, and that it leads to the reporting of many new members. In that way, the clan would use its resources to confirm the significance of its achievements ex-post. Another channel for clan size growth might be that people who were not previously part of the clan might try to establish an ancestry relationship to the high status individual. We employ panel regression analysis to provide evidence on this and related hypotheses (seven clans by twelve birth cohorts, as shown in Figure A.4). Table A.1 shows the results.

There is a negative point estimate for lagged average clan status (not significant, see Table A.1,

Table A.1: Clan Success and Recall Bias

	Clan Size		
	(1)	(2)	(3)
Clan Status (-1)	−0.098	−0.026	−0.061
	(0.096)	(0.098)	(0.157)
Birth Cohort Fixed Effects		Y	Y
Clan Fixed Effects			Y
N	74	74	74

Notes: Dependent variable is clan size, measured by the number of men born in a cohort (in units of 100s). Clan Status (-1) is average clan status lagged by one birth cohort. Estimation by OLS. Standard errors in parentheses.

column (1)), which does not support the hypothesis that notable achievements in the recent past trigger the inclusion of new members into the clan's genealogy. This does not change with the inclusion of birth cohort and clan fixed effects, see Table A.1, columns (2) and (3). Overall, we do not find evidence that recall bias plays a major role for the composition and characteristics of the sample.

A.3.5 Using Cross-Clan Variation to Assess Potential Bias

Because the creation and maintenance of a genealogy requires resources and literacy, the key selection concern for the sample is wealth bias. We can shed direct evidence on this using cross-clan variation, given that the seven clans differ substantially in their average wealth as proxied by status, see Table **??**. It was shown in the text that clan size does not simply reflect a clan's resources, measured either by status, by wives per man, or by number of children (Figure 2). There is no evidence that richer clans systematically report more new members than poorer clans.

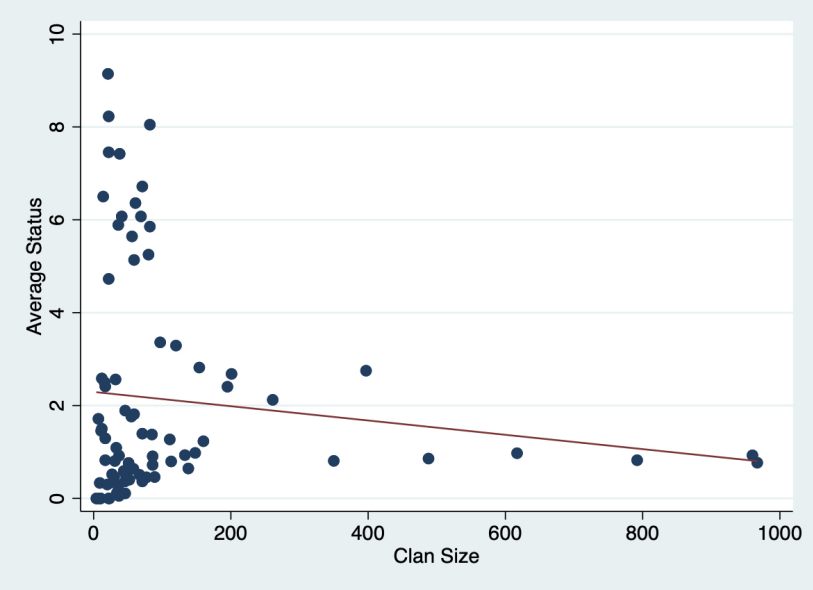
Wealth Bias: Additional Results One might be concerned that even though there is no cross-sectional correlation between a clan's size and its resources, there could be such a correlation in the temporal dimension. To provide evidence on this, the following examines such temporal correlation between status and clan size across twelve birth cohorts.[22] We begin with a scatterplot, see Figure A.5.

Figure A.5 shows that there is no strong relationship between status and clan size across birth cohorts, and to the extent that there is a relationship it is negative.

Survivorship Bias A related concern is survivorship bias: over time, this type of bias could result in a disproportionately fraction of high-achieving (high skill) individuals compared to low-achieving (low skill) individuals. One implication is that towards the end of the sample period the distribution becomes skewed towards relatively high-achieving clans. As Figure A.6 shows, however, there is

[22]The birth cohorts are those of Figure A.4.

Figure A.5: Clan Size and Status: Birth Cohort Analysis



Notes: Shown is the size of each clan as measured by the number of men born in each of twelve birth cohorts, versus average clan status in that birth cohort; N = 81.

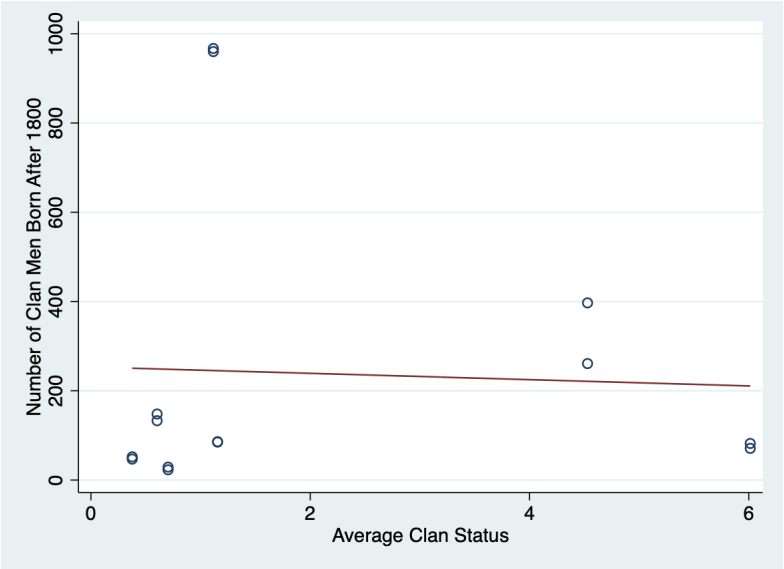
little evidence that the distribution of the sample towards the end of the sample period is becomes skewed towards high-status clans.

B Supplementary Empirical Results

B.1 Parent Child Intergenerational Mobility and Marital Sorting: Results with Clan Fixed Effects

Table A.2 shows results on the role of marital sorting for intergenerational mobility when clan fixed effects are included.

Figure A.6: Evidence on Survivor Bias



Notes: Figure shows number of clan men born in periods 1800-1825 and post-1825, both versus average clan status.

Table A.2: Marriage and Parent-Child Mobility with Clan Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	0.410**	0.281**	0.340**	0.253**	0.322**	0.244**			
	(0.018)	(0.020)	(0.018)	(0.020)	(0.018)	(0.019)			
		0.271**		0.222**		0.208**			
		(0.022)		(0.022)		(0.022)			
							0.553**	0.477**	0.454**
							(0.020)	(0.021)	(0.020)
	Y	Y	Y	Y	Y	Y	Y	Y	Y
			Y	Y	Y	Y		Y	Y
					Y	Y			Y
	4,967	4,967	4,967	4,967	4,925	4,925	4,967	4,967	4,925

tile rank; estimation of equation (10); robust standard errors reported in parentheses; **/*/+ indicates significance

B.2 Parent-Child Mobility Over Time

Table A.3 shows results on the role of marital sorting for intergenerational mobility separately for cohorts before and after the year 1750.

Table A.3: Marriage and Parent-Child Mobility over Time

	(1)	(2)	(3)
Father Status $<$ 1750	0.425**		0.295**
	(0.020)		(0.027)
Father Status $>=$ 1750	0.395**		0.269**
	(0.021)		(0.027)
Father-in-Law Status $<$ 1750		0.426**	0.261**
		(0.022)	(0.029)
Father-in-Law Status $>=$ 1750		0.420**	0.278**
		(0.023)	(0.021)
N	4,972	4,972	4,972

Notes: Dependent variable son status as percentile rank; estimation of equation (10); son birth year fixed effects included; robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

First, we see that in both subperiods, father-in-law status has a similarly large role for son status as the status of the son's biological father. There is a difference between the roles of biological and in-law family insofar as the point estimate on father status decreases over time while the point estimate on wife status does less so (column (3)). This indicates that trends in the role of marriage can at least partially offset trends in the role of the biological family for mobility.

B.3 Half-Brothers versus Full Brothers

Table A.4 reports results on the role of marriage for brothers that distinguish half-brothers from full brothers. Adding mother fixed effects to father fixed effects means that specifications in columns (2) and (4) exploit variation between brothers who have the same father and mother, that is, full brothers.

Table A.4: Biological versus Half-Brothers, Marriage, and Mobility

	(1)	(2)	(3)	(4)
	Two or Three Brothers		Two or More Brothers	
	Half and Full Bros	Full Bros	Half and Full Bros	Full Bros
Father-in-Law Status	0.265**	0.265**	0.078*	0.079*
	(0.084)	(0.084)	(0.038)	(0.038)
Father Fixed Effects	Y	Y	Y	Y
Mother Fixed Effects		Y		Y
N	860	860	2,389	2,370

Notes: Dependent variable son status as percentile rank; estimation of equation (11); son birth year fixed effects included; robust standard errors clustered on father reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

Results indicate that separating half-brothers from full brothers does not lead to major changes in the role of marriage for mobility of brothers, in part because the number of half-brothers is small.

B.4 Heterogeneity Analysis

B.4.1 Son Differences

Table A.5 examines heterogeneity of the role of marriage for mobility depending on characteristics of the son.

Table A.5: Son Heterogeneity: Marriage and Parent-Child Mobility

	(1)	(2) Migration	(3) Birth Order	(4) Burial Distance	(5) Lifespan
Father Status	0.281**	0.281**	0.281**	0.282**	0.279**
	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)
Father-in-Law Status	0.271**	0.274**	0.294**	0.280**	0.042
	(0.022)	(0.023)	(0.039)	(0.027)	(0.065)
Father-in-Law Status x Migration		−0.029			
		(0.069)			
Migration		0.021			
		(0.031)			
Father-in-Law Status x Birth Order			−0.010		
			(0.015)		
Birth Order			0.003		
			(0.007)		
Father-in-Law Status x Burial Dist				−0.004	
				(0.007)	
Burial Distance				0.002	
				(0.003)	
Father-in-Law Status x Lifespan					0.004**
					(0.001)
Lifespan					−0.001*
					(0.001)
N	4,972	4,972	4,972	4,972	4,972

Notes: Dependent variable son status as percentile rank; estimation of extended equation (10); son birth year fixed effects included; robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

The analysis considers heterogeneity with respect to the son's migration activity, measured as living in another town or village compared to his father, as well as the son's birth order, burial distance, and lifespan (death year minus birth year). Results indicate that there is only significant heterogeneity with respect to the son's lifespan. Specifically, higher lifespan increases the role of marrying a high-status woman, at the same time when the linear wife status coefficient turns insignificant (column (5)). One explanation is that lifespan is correlated with status, which is the left

hand side variable. Along these lines, the results in column (5) largely reflect that husband and wife status are complements. Interestingly, a lower husband birth order (earlier-born son) is not correlated with the role that father-in-law status plays for husband status (column (3)).

B.4.2 Father Heterogeneity

Table A.6 examines heterogeneity of marriage's role for son mobility depending on characteristics of the son's father.

Table A.6: Father Heterogeneity: Marriage and Parent-Child Mobility

	(1)	(2) Birth Order	(3) Generation	(4) Lifespan	(5) Status	(6) Clan Status
Father Status	0.281**	0.280**	0.269**	0.281**	0.187**	0.241**
	(0.020)	(0.020)	(0.020)	(0.020)	(0.052)	(0.019)
Father-in-Law Status	0.271**	0.313**	0.703**	0.265**	0.166**	−0.117
	(0.022)	(0.041)	(0.118)	(0.084)	(0.058)	(0.103)
Father-in-Law Status x Birth Order		−0.019				
		(0.015)				
Birth Order		0.008				
		(0.007)				
Father-in-Law Status x Generation			−0.034**			
			(0.009)			
Generation			0.003			
			(0.004)			
Father-in-Law Status x Lifespan				0.000		
				(0.001)		
Lifespan				−0.000		
				(0.001)		
Father-in-Law Status x Father Status					0.161+	
					(0.085)	
Father-in-Law Status x Av. Clan Status						0.604**
						(0.181)
Average Clan Status						0.014
						(0.101)
N	4,972	4,972	4,972	4,972	4,972	4,968

Notes: Dependent variable son status as percentile rank; estimation of extended equation (10); son birth year fixed effects included; robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

The results indicate that father lifespan (an indicator of health) and father birth order, perhaps an indicator of differential access to parental resources, do not influence the role of wife status for son intergenerational mobility (columns (2) and (4), respectively). The negative interaction coefficient with father generation likely reflects the secular downward trend in status due to the decline of China's civil service examination (column (3)). Results on father status and average clan father

status in column (5) and (6) indicate that a good marriage is particularly strongly associated with higher son status when the son's father, or his clan, are relatively rich. It also suggests that marriage search or other investments into a good marriage are complementary to parent's investments into their son to pass the civil service exam and obtain office.

B.4.3 Mother Heterogeneity

Table A.7 examines heterogeneity of marriage's role for son mobility depending on characteristics of the son's mother.

Table A.7: Mother Heterogeneity: Marriage and Parent-Child Mobility

	(1)	(2) Birth Order	(3) Lifespan	(4) Single Wife	(5) First Wife	(6) Second Wife First dead	(7) Second Wife First Alive	(8) Concubine
Father Status	0.281**	0.280**	0.279**	0.280**	0.280**	0.281**	0.281**	0.281**
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Father-in-Law Status	0.271**	0.313**	0.190**	0.283**	0.266**	0.270**	0.270**	0.271**
	(0.022)	(0.041)	(0.075)	(0.040)	(0.023)	(0.023)	(0.022)	(0.022)
Father-in-Law Status x Birth Order		−0.008						
		(0.030)						
Birth Order		0.005						
		(0.014)						
Father-in-Law Status x Lifespan			0.001					
			(0.001)					
Lifespan			−0.000					
			(0.001)					
Father-in-Law Status x Single Wife				−0.017				
				(0.044)				
Single Wife				0.004				
				(0.021)				
Father-in-Law Status x First Wife					0.039			
					(0.059)			
First Wife					−0.014			
					(0.030)			
Father-in-Law Status x Second Wife, First Dead						0.004		
						(0.058)		
Second Wife, First Dead						0.003		
						(0.028)		
Father-in-Law Status x Second Wife, First Alive							0.635**	
							(0.081)	
Second Wife, First Alive							−0.325**	
							(0.050)	
Father-in-Law Status x Concubine								−0.177**
								(0.046)
N	4,972	4,972	4,972	4,972	4,972	4,972	4,972	4,972

Notes: Dependent variable son status as percentile rank; estimation of extended equation (10); son birth year fixed effects included; linear Concubine coefficient not identified; robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

Results shown in Table A.7 indicate that the role of the mother in her household matters to some extent for the influence of father-in-law status on son mobility. The interaction point estimate with first wife is positive, which is evidence that the son of a first-wife mother benefits disproportionately more from the marriage to a high-status wife (not significant; column (5)). The son of a second wife where the first wife is alive benefits significantly more from the status of his wife (column (7)); one explanation is that such sons are particularly supported by their parents because an important reason why there is a second wife is that the first does not bear a son, and the fact that the first is alive indicates that the household is unusually rich. In contrast, the son of a concubine benefits disproportionately less from marrying a high status wife (column (8)); this may be due in part to the fact that concubines themselves have relatively low status.

Overall, results in these heterogeneity analyses of the role of marriage for husband status are both interesting and broadly in line with the historical evidence.

B.5 Marriage and Multigenerational Mobility with Clan Fixed Effects

Table A.8 shows results for multigenerational mobility when accounting for marriage and clan fixed effects.

Table A.8: Marriage and Multigenerational Mobility: Clan Effects

	(1)	(2)	(3)	(4)
Father Status	0.322**	0.267**	0.215**	0.207**
	(0.018)	(0.020)	(0.021)	(0.021)
Grandfather Status		0.110**	0.073**	0.060**
		(0.015)	(0.015)	(0.016)
Father-in-Law Status			0.186**	0.174**
			(0.022)	(0.023)
Maternal Grandfather Status				0.048**
				(0.018)
N	4,926	4,926	4,926	4,926

Notes: Dependent variable son status as percentile rank; father and grandfather status (paternal and maternal) in percentile rank; all specifications include husband birth year fixed effects, male clan effects, and fixed effects for the clans of the in marrying wives in the husband generation. Estimation of equation (12); robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

Overall, these results are broadly similar to the results without clan fixed effects shown in Table 6.

B.6 Marriage and Multigenerational Mobility over Time

Table A.9 shows results on the role of marital sorting for multigenerational mobility separately for cohorts before and after the year 1750.

Table A.9: Marriage and Multigenerational Mobility over Time

	(1)	(2)	(3)	(4)
Father Status $<$ 1750	0.295**	0.255**	0.269**	0.245**
	(0.027)	(0.030)	(0.029)	(0.031)
Father Status $>=$ 1750	0.269**	0.226**	0.232**	0.210**
	(0.027)	(0.028)	(0.027)	(0.028)
Father-in-Law Status $<$ 1750	0.261**	0.226**	0.229**	0.211**
	(0.029)	(0.031)	(0.032)	(0.033)
Father-in-Law Status $>=$ 1750	0.278**	0.244**	0.232**	0.218**
	(0.021)	(0.030)	(0.027)	(0.030)
Paternal Grandfather Status $<$ 1750		0.094**		0.078**
		(0.026)		(0.027)
Paternal Grandfather Status $>=$ 1750		0.098**		0.071**
		(0.019)		(0.020)
Maternal Grandfather Status $<$ 1750			0.083**	0.055^{+}
			(0.028)	(0.030)
Maternal Grandfather Status $>=$ 1750			0.114**	0.089**
			(0.022)	(0.024)
N	4,972	4,972	4,972	4,972

Notes: Dependent variable son status as percentile rank; estimation of extended equation (12); son birth year fixed effects included; robust standard errors reported in parentheses; **/*/+ indicates significance at 1%/5%/10% level.

As in the analysis of trends in the role of marriage for parent-child mobility, we see that marriage in the previous generation can offset the trend towards a smaller role of the son's paternal grandfather for son status (column (4)).