DISCUSSION PAPER SERIES

DP13589
(v. 3)

SOCIAL MOBILITY IN THE LONG RUN: A TEMPORAL ANALYSIS OF CHINA FROM 1300 TO 1900

Carol Hua Shiue

DEVELOPMENT ECONOMICS

ECONOMIC HISTORY

ISSN 0265-8003

SOCIAL MOBILITY IN THE LONG RUN: A TEMPORAL ANALYSIS OF CHINA FROM 1300 TO 1900

Carol Hua Shiue

Discussion Paper DP13589
First Published 15 March 2019
This Revision 10 January 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Development Economics
- Economic History

SOCIAL MOBILITY IN THE LONG RUN: A TEMPORAL ANALYSIS OF CHINA FROM 1300 TO 1900

Abstract

This paper studies intergenerational mobility with a population of families in central China over twenty generations. Employing genealogical data on individual lifetime achievements, I first find that while mobility was low initially there was a striking increase in mobility starting in the late 17th century. Through the lens of a Becker-Tomes (1979) model I explain this through a falling human capital earnings elasticity because the return to passing the civil service examinations, China's most important pathway to office and income at the time, declined over time. Second, as predicted by the model times of high human capital earnings elasticity are times of high cross-sectional inequality. Third, parent human capital affects child income for any given nonhuman parental investment and is estimated to have 2/3 of the effect of nonhuman investments on child income. Moreover, educational inequality is even more strongly correlated with social persistence than income inequality. Finally, much of the observed increase in mobility is accounted for by lower differences in average clan income, consistent with the hypothesis that part of the earnings elasticity of human capital is group-specific.

JEL Classification: N35, N45, J62

Keywords: N/A

Carol Hua Shiue - carol.shiue@colorado.edu
University of Colorado-Boulder and CEPR

Acknowledgements
Thanks to Miles Corak, Raquel Fernandez, Joseph Ferrie, Murat Iyigun, Wolfgang Keller, Naomi Lamoreaux, and Debraj Ray, as well as audiences at the CEPR Economic History Symposium, the Cliometrics Society Conference, the Conference on Deep Rooted Factors in Comparative Development at Brown, Iowa State, Kiel, Michigan, the NBER Summer Institute Development of the American Economy, Nebraska, Northwestern, NYU, the Paris School of Economics Conference on China and India, the Shanghai University of Finance and Economics (Shanghai), the Southwestern University of Finance and Economics (Chengdu), and the World Bank for useful comments. This project was supported by a grant from the National Institutes of Health (NICHD) Award HD 042731-01 "Intergenerational and Interfamily Economic Links". Ted Telford provided part of the data employed in this paper. Thanks to Matt Buttner and Peiyuan Li for excellent research assistance. Email: shiue@colorado.edu.

Social Mobility in the Long Run
A Temporal Analysis of China from 1300 to 1900*

Carol H. Shiue

University of Colorado, NBER, and CEPR

December 30, 2019

Abstract

This paper studies intergenerational mobility with a population of families in central China over twenty generations. Employing genealogical data on individual lifetime achievements, I first find that while mobility was low initially there was a striking increase in mobility starting in the late 17th century. Through the lens of a Becker-Tomes (1979) model I explain this through a falling human capital earnings elasticity because the return to passing the civil service examinations, China's most important pathway to office and income at the time, declined over time. Second, as predicted by the model times of high human capital earnings elasticity are times of high cross-sectional inequality. Third, parent human capital affects child income for any given nonhuman parental investment and is estimated to have 2/3 of the effect of nonhuman investments on child income. Moreover, educational inequality is even more strongly correlated with social persistence than income inequality. Finally, much of the observed increase in mobility is accounted for by lower differences in average clan income, consistent with the hypothesis that part of the earnings elasticity of human capital is group-specific.

JEL N35, N45, J62

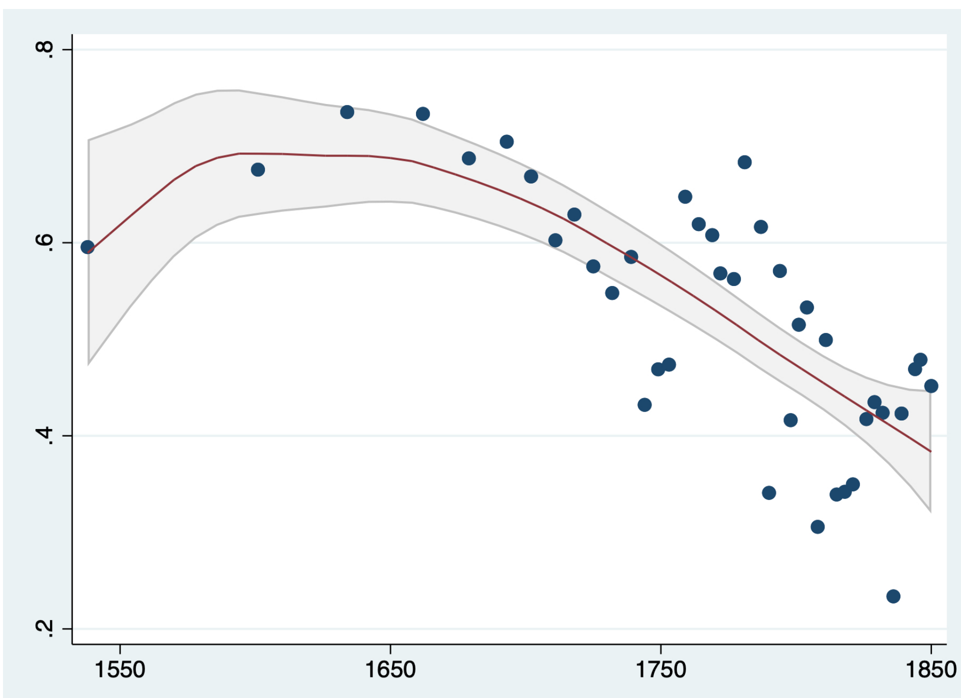
*Thanks to Miles Corak, Raquel Fernandez, Joseph Ferrie, Murat Iyigun, Wolfgang Keller, Naomi Lamoreaux, and Debraj Ray, as well as audiences at the CEPR Economic History Symposium, the Cliometrics Society Conference, the Conference on Deep Rooted Factors in Comparative Development at Brown, Iowa State, Kiel, Michigan, the NBER Summer Institute Development of the American Economy, Nebraska, Northwestern, NYU, the Paris School of Economics Conference on China and India, the Shanghai University of Finance and Economics (Shanghai), the Southwestern University of Finance and Economics (Chengdu), and the World Bank for useful comments. This project was supported by a grant from the National Institutes of Health (NICHD) Award HD 042731-01 "Intergenerational and Interfamily Economic Links". Ted Telford provided part of the data employed in this paper. Thanks to Matt Buttner and Peiyuan Li for excellent research assistance. Email: shiue@colorado.edu.

1 Introduction

Many factors affect social mobility, from parental investments over public expenditures to cultural values (Becker and Tomes 1979, Solon 2004, Chetty, Hendren, Kline, and Saez 2014, and Greif and Tabellini 2017). Although social mobility is a central characteristic of any economy, we know little about the extent of mobility in past periods and how mobility changed over time. This paper examines the patterns of social mobility by tracing out intergenerational mobility of individuals and families living in one area in a central province of China over twenty generations, spanning approximately the years 1300 to 1900. By studying a representative subset of the Chinese society, we can begin to understand changes of mobility in China, home to close to 20% of the human population.

The study uses genealogical data from biographies of individuals living in Tongcheng County, China. I begin by establishing the level of mobility over these centuries and showing that the trend was towards greater mobility. Employing a regression approach that relates son income to father income, the intergenerational coefficient is about 0.5, showing that any advantage that one parent has over another is, on average, cut in half in the generation of their children. This value, however, masks important temporal variation, as Figure 1 shows. Estimates for different periods indicate that mobility in China during my sample period was low until the middle of the 17th century before it increased substantially from the late 17th century to the 19th century.

Figure 1: Intergenerational Mobility: Income of Son Relative to Income of Father



Notes: Shown are β_c coefficients from OLS regressions of income rank son on income rank father for forty birth cohorts. High values of β_c indicate low mobility. Horizontal axis gives median birth year of son in cohort. Earliest son birth year is 1330. See section 5 for details.

For a population that was genetically similar and resided in the same location over the entire period, the changes in mobility are large–varying from over 0.6 to below 0.4. Moreover, the evolution of mobility over time is U-shaped. Therefore, one must look for reasons outside of population heterogeneity and cultural difference, which may be important when comparing cross-country differences in mobility.

One of the key factors that is related to mobility is the extent to which skills and education are rewarded in the economy. While large changes in human capital are associated with the period after the Industrial Revolution of the 18th and 19th centuries in Europe, changes also took place in China. Occupational barriers were loosened between the 14th and 16th centuries, and the civil service examinations became an institutionalized feature of society with the 17th century. To better understand the temporal changes, I present a simple version of the Becker-Tomes model (1979, 1986) model, where mobility depends on human capital investments (see Mogstad 2017 for broader overview). Here, not only do high-income parents invest more than low-income parents in the human capital of their children, but times during which the earnings elasticity is high are times in which parents choose to invest more intensively into their children's human capital. Consequently, intergenerational mobility tends to be low when the earnings elasticity is high and *vice versa.*

I employ data on seven clans living in Tongcheng county.[1] There are approximately 40,000 individuals in the sample (of which there are around 10,000 married couples). The data is used to examine the degree of intergenerational mobility between father and son for different birth cohorts, in order to shed light on factors that covary with the pattern of mobility over time, and to investigate the role of clan membership for mobility. The data reflect not only a wide range of high and low income groups but also a substantial degree of variation in mobility, inequality, and other factors, across birth cohorts.

There are several findings in addition to the change in intergenerational mobility shown in Figure 1. First, I show that there was a long-run relationship between mobility and the earnings elasticity of human capital. The 14th to late 17th centuries not only exhibited low intergenerational mobility but also a high earnings elasticity of human capital. Moreover, as the earnings elasticity started to fall in the 18th century, the lowering of the human capital earnings elasticity was accompanied by a relatively higher level of intergenerational mobility from the 1700s into the 1800s. The temporal relationship between mobility and earnings elasticity seen in China over the long-run is consistent with one of the key predictions of the model.

A second key prediction of my model is that a high earnings elasticity of human capital leads not only to intergenerational persistence but also to high cross-sectional inequality.[2] I examine this prediction in the Chinese context by measuring the degree of inequality for different birth cohorts between 1300 and 1900. The empirical evidence strongly confirms the prediction: cross-sectional inequality is highly correlated with intergenerational mobility for birth cohorts between the years 1300 to 1900, reinforcing the evidence

[1]Chinese clans are also described in the literature as lineages, common-descent groups, or dynasties. They are based on the male surname, mirroring the imperial dynasty.

[2]More generally, Becker, Kominers, Murphy, and Spenkuch (2018) show that intergenerational persistence and inequality need not be positively related.

that human capital investments play an important role for mobility. The relationship between mobility and inequality is not causal but rather both are endogenous to parental incentives to invest in child human capital and the effectiveness of such investments.

Third, my results provide evidence that parental human capital has a direct effect on mobility in addition to the parental income channel (Cunha and Heckman 2007). Specifically, augmenting the intergenerational regression, I find that conditional on the income of his father, a son's income is increasing in his father's education. Moreover, cross-sectional inequality in *education* is a more powerful predictor of intergenerational mobility than is cross-sectional inequality in income, providing further evidence for the important role of parental human capital.

I also present new results on how clans shape the patterns of intergenerational mobility. Decomposing total inequality into a between- and a within clan part, I show that much of the increase in mobility towards the 19th century can be accounted for by smaller between-clan differences in income. This indicates that part of the earnings elasticity of human capital is clan specific. One explanation, consistent with the historical evidence, is that members of the same clan transferred resources among themselves to maximize total clan income. Other groups effects (through inheritance of endowments, cultural values, etc.) are also possible. Together with the other findings this constitutes new evidence on the central role of human capital for intergenerational mobility based on long-run time-series variation.

To explain the timing of when the U-shaped empirical pattern turns, I argue that the most important factor was the increasing strains on the state's fiscal capacity in the face of a rapidly rising population. Decreasing per-capita tax revenues meant that the return to top official positions declined and civil service student stipends fell. Consistent with other historically documented conditions in the late 18th and 19th centuries, these factors imply a decline in the earnings elasticity of human capital.

This paper makes a number of contributions. First, my analysis of mobility focuses on parent and family incentives to invest into children. The historical setting of China provides in some ways an ideal setting in which to examine these fundamental dynamics of social mobility because government programs and redistribution are largely absent. While human capital investments were important after the Industrial Revolution when these investments increased globally, differences in government activity–laws on mandated schooling, tax redistribution, and other public welfare programs– may have also impacted the earnings elasticity of human capital and hence mobility. By contrast, in China during this period there was no income tax, no public schooling, and nor were there major social security programs (for instance for old age or unemployment). With few exceptions, redistribution was transacted privately and informally among individuals or groups of people who generally knew each other well. Thus, my analysis provides information about how human capital investments, mobility, and inequality are related in a simpler setting.

Second, I shed new light on social mobility in China, on which comparatively little is known. China was closer to being a 'land of opportunity' in the year 1800, compared to what it had been in the year 1500. We can be more certain of that because the sample that I use covers not only the elites but commoners as well. By contrast, the classic study of Ho (1967) saw evidence for relatively high mobility in China based on data from the top 3% of society. I complement Ho's work by combining the focus on top members with an analysis of the remaining 97% of society. Sociological research on mobility with data from the Eight-Banner system, a privileged class of Qing emperors, their relatives, and farmers who provided food for them, has yielded evidence for multigenerational and kin effects (Campbell and Lee 2003, Mare and Song 2014).[3] My analysis differs in that I study a more typical Chinese population, and I contribute by extending the period under study to the centuries before 1750.[4] My usage of genealogical information also pushes the analysis of individually-linked data back to the year 1300, which is important not least because social mobility changes only slowly over time.[5] Further, analyzing the speed of diffusion of distinctive surnames of elites into the general population, Hao and Clark (2012) estimate a low and largely constant level of social mobility for the Qing dynasty (intergenerational coefficients of about 0.85). My finding that mobility was higher and increasing is due in part to group effects that have not been identified in previous studies.[6]

Third, my analysis sheds new light on the role for mobility of parents' human and nonhuman investments (Becker and Tomes 1979, 1986). I provide new time-series evidence that intergenerational mobility and cross-sectional inequality are closely related to the earnings elasticity of human capital.[7] This analysis complements findings that the return to human capital has affected fertility in Ming-Qing China by changing the quantity-quality trade-off for children (Shiue 2017). Furthermore, my results present support from a long-term, historical perspective that parental human capital and investment in children are complementary (Heckman and Mosso 2014). Finally, my finding that the increase in China's social mobility is matched by a decline of between-clan inequality complements other work on the role of clans in economic development (de la Croix, Doepke, and Mokyr 2017).

Fourth, I contribute to a large literature studying social mobility using a two-generation regression approach (Solon 1999, Black and Devereux 2010). In the Chinese sample, I estimate an overall intergenerational coefficient of roughly 0.5 and a range between about 0.35 and 0.70 (see Figure 1). Estimates of log son on log father income for OECD countries during the late 20th century range from about 0.5 for

[3] Multigenerational effects are those beyond the influence of parent on child. See also Long and Ferrie (2018) for evidence on the influence of grandfathers for the US and the UK since 1850. For the Chinese case between 1300 and 1900, evidence on the influence of up to five generations (great-great-grandfather) is shown in section 6.5.

[4] The Eight-Banner system's task was to militarily protect the homeland of the Qing in North China. The region ranked highest among all regions of Qing China in terms of top degrees awarded (*jin-shi*) per capita, and given their high position members of the Qing lineage were impervious to upward mobility (Ho 1967 and Mare and Song 2014, respectively). The supporting population consisted mostly of hereditary tenants, serfs, and war captives (Campbell and Lee 2003, Mare and Song 2014). As a consequence, the Eight-Banner population is distinct from typical Chinese populations in ways that plausibly affect social mobility.

[5] For example, Chetty, Hendren, Kline, Saez, and Turner (2014) find that mobility for US children entering the job market now and in the 1970s is roughly the same. Even over four generations, Lindahl, Palme, Sandgren-Massih, and Sjogren (2015) find only small differences in educational mobility once average education is controlled for.

[6] Groups may help to determine individual success (see sections 2 and 5). One explanation for low mobility based on distinctive surnames is that distinctive surnames are indicative of income so that low mobility reflects the low mobility in group averages (Chetty, Hendren, Kline, and Saez 2014). Another possibility is that analysis of two generations underestimates the rate of persistence (Solon 2015, Long and Ferrie 2018). Employing five linked generations I show that while a multigenerational specification yields higher persistence than a two generation specification, additional generations do not explain mobility changes in my analysis (see section 6.5).

[7] A test of the Becker and Tomes (1979, 1986) model using information on grandfathers and great-grandfathers is Lindahl, Palme, Sandgren-Massih, and Sjogren (2014).

the United States, UK, and Italy to below 0.2 for Scandinavian countries (Corak 2013). Using tax income data, Chetty, Hendren, Kline, and Saez (2014) estimate a rank son-rank father correlation of 0.34 for the U.S. around the year 2000, documenting stark differences in social mobility across U.S. regions. While historical studies of income mobility are virtually non-existent, analyses with data from the mid-19th century onwards find intergenerational coefficients of about 0.25 in the case of US and UK occupational mobility (Long and Ferrie 2018) and 0.3 and 0.5 in the case of educational mobility in Sweden and Germany, respectively (Lindahl, Palme, Sandgren-Massih, and Sjogren 2015, and Braun and Stuhler 2018, respectively).[8] By estimating mobility differences over time for a homogeneous population rather than across countries that have their own populations, institutions, and laws, the temporal strategy in this paper provides an alternative approach for inferring changes in economic opportunity from changes in intergenerational mobility.

The remainder of the paper is as follows. Section 2 introduces a version of Becker and Tomes' (1979, 1986) model of intergenerational mobility and the distribution of income that will guide the interpretation of my empirical results. The following section 3 provides a synthesis of China's economy during the sample period while section 4 introduces key elements of the data underlying this study. Section 5 presents the results on how intergenerational mobility changed over my sample period, while section 6 explains these patterns in the light of changing human capital investments emphasized by the model based on historical evidence. Section 7 provides a concluding discussion. The Appendix contains additional information on the genealogical information employed in this study as well as important extensions and robustness checks.

2 Theoretical Framework

The framework introduced in this section relates human capital investments, inequality, and intergenerational mobility along the lines of Becker and Tomes (1979, 1986), Cunha and Heckman (2007), and Becker, Kominers, Murphy, and Spenkuch (2018).

Let a single parent p maximize the following Cobb-Douglas utility function:

$$U_p = (1-\alpha)\, log\, c_p + \alpha log\, y_c, \qquad 0 < \alpha < 1, \tag{1}$$

where U_p is parent utility, c_p is parent consumption, and y_c is the child's income. I assume that each parent has only one child.[9] The parameter α in equation (1) is a measure of the degree of parental interest in the child's income. Investing into the child's human capital raises child income but it reduces parent consumption:

$$c_p = y_p - i_p, \tag{2}$$

where i_p and y_p are parent's investment and income, respectively.[10] I assume that parental investment

[8]See also Ferrie (1996, 2005) as well as Long and Ferrie (2013) for earlier results.
[9]Shiue (2017) examines child quantity versus quality choice in Ming-Qing China.
[10]There are no bequests in this model; see Becker, Kominers, Murphy, and Spenkuch (2018) for an analysis of bequest.

positively affects the human capital of the child, H_c, according to

$$H_c = A_c i_p^{\gamma} H_p^{\kappa} L_p^{\mu}, \tag{3}$$

where $\gamma > 0$, $\kappa, \mu \geq 0$. The term A_c is child ability, originally modeled by Becker and Tomes (1979) as a stationary autoregressive process capturing a range of genetic and cultural endowments of the families. See Solon (2004, 2014) for a discussion of the implications of such endowment processes for intergenerational mobility. To increase the focus on parental investment I follow Becker, Kominers, Murphy, and Spenkuch (2018) and set $A_c = 1, \forall c$. [11]

The human capital of the parent, H_p, may positively impact the human capital of the child for a given level of investment (Cunha and Heckman 2007, Cunha, Heckman, and Schennach 2010). In addition, the child's human capital may also be affected by individuals other than the parents. Given the major role of kin in Chinese societies, I allow for clan effects, L_p in equation (3). One example of this could be resource transfers between clan members that serve to maximize clan income.[12]

Human capital translates into income according to:

$$y_i = rH_i^{\sigma} f(l_i), \qquad i = p, c \tag{4}$$

where l_i is luck, and $\sigma \geq 0$. In equation (4), r can be thought of as the "base price" of human capital (Becker, Kominers, Murphy, and Spenkuch 2018), while σ is the earnings elasticity of human capital, $\sigma \equiv \frac{\partial y_i / y_i}{\partial H_i / H_i}$. I assume that luck, which is stochastic, increases income, $f'(.) > 0$.

Using the budget constraint (2), the investment technology (3), and the relationship between human capital and income (4), the parent's objective function can be rewritten as

$$max \left\{(1-\alpha) \, log \, (y_p - i_p) + \alpha log \left[rH_c^{\sigma} f(l_c)\right]\right\} = max \left\{(1-\alpha) \, log \, (y_p - i_p) + \alpha log \left[r \left(i_p^{\gamma} H_p^{\kappa} L_p^{\mu}\right)^{\sigma} f(l_c)\right]\right\}. \tag{5}$$

The first-order necessary condition for the parent's optimal investment choice is

$$\frac{\partial U_p}{\partial i_p} = \frac{1}{y_p - i_p}(-1) + \alpha \frac{1}{\left[r \left(i_p^{\gamma} H_p^{\kappa} L_p^{\mu}\right)^{\sigma} f(l_c)\right]} rH_p^{\kappa\sigma} L_p^{\mu\sigma} f(l_c) \gamma \sigma i_p^{\gamma\sigma - 1} = 0 \tag{6}$$

which yields the following expression for parental investment in the child's human capital

$$i_p = \left[\frac{\alpha\gamma\sigma}{1 + \alpha\gamma\sigma}\right] y_p. \tag{7}$$

Note, first, that investment is increasing in parental income. Second, it is easy to verify that investment is higher the greater is the parent's taste for child income (α), the more productive is the investment in generating child human capital (γ), and the higher is the earnings elasticity of human capital (σ).

[11] Becker and Tomes (1979) and Solon (2004, 2014) note that a stationary AR(1) process for ability implies that child's income is affected by grandparent's income. I provide empirical evidence on the influence of grandparents in section 6.5.

[12] Section D.4 shows that clan rules were consistent with such activities.

To obtain a relationship between the child's income and the parent's income, substitute equation (7) into equation (3) and the result of that into equation (4) to obtain

$$y_c = r\left(\left[\frac{\alpha\gamma\sigma}{1+\alpha\gamma\sigma}y_p\right]^{\gamma} H_p^{\kappa} L_p^{\mu}\right)^{\sigma} f\left(l_c\right) = \xi y_p^{\gamma\sigma} f\left(l_c\right) \tag{8}$$

where $\xi \equiv r\left(\left[\frac{\alpha\gamma\sigma}{1+\alpha\gamma\sigma}\right]^{\gamma} H_p^{\kappa} L_p^{\mu}\right)^{\sigma}$.[13] Equation (8) shows that both higher parent income and luck increase child income. For positive values of κ and μ, child income is also increasing in parental human capital and clan contribution, respectively. Let $f\left(l_c\right)$ take the following form: $f\left(l_c\right) = e^{l_c}$. Taking logs of equation (8) yields

$$log\, y_c = log\, \xi + \gamma\sigma\, log\, y_p + l_c, \tag{9}$$

Equation (9) shows that in the case of $\kappa = \mu = 0$, the intergenerational coefficient in the familiar log son-log father income regression is an estimate of γ times σ. Thus, the better the investment technology for human capital (γ) and the higher the earnings elasticity of human capital (σ), the lower is intergenerational mobility. In particular, high values of σ mean that rich parents will invest a high fraction of their income into the human capital of their children (equation (7)), which tends to lead to persistence. Conversely, in the extreme case of $\sigma = 0$, parents' investment into the human capital of their children is zero (equation (7)), son income depends solely on luck (equation (9)), and intergenerational mobility is relatively high.

I now turn to the relationship between cross-sectional inequality and intergenerational mobility. Assume for now that $\kappa = \mu = 0$ and $\gamma\sigma < 1$. In this case, equation (9) is a stationary AR(1) equation. In steady state, when the variances of $log\, y_c$ and $log\, y_p$ are equal, one obtains

$$Var(log\, y) = \frac{1}{1-(\gamma\sigma)^2} Var\left(l_c\right). \tag{10}$$

Equation (10) shows that cross-sectional inequality, measured by the variance of log income, is increasing in γ and σ for a given variability in luck.[14] In particular, the model predicts that a high earnings elasticity of human capital not only increases intergenerational persistence but also cross-sectional income inequality, and *vice versa*. This is a central implication of this model that will be analyzed. The result depends on $\gamma\sigma$ being smaller than one, which is in line with my estimates below. Implications of other parameter values are discussed in Becker, Kominers, Murphy, and Spenkuch (2018).

3 Central Features of China's Economy, 1300 - 1900

This study covers the period from the late 13th to the late 19th century. There were three imperial dynasties in China during my sample period: part of the Yuan (1279-1368), all of the Ming (1368-1644), and the major part of the Qing Dynasty (1644-1911). Total population grew from about 100 million in year 1280 to 400 million by 1900 (Cao 2000; Maddison 2007). China's turn towards increased contact

[13] Note that for $\kappa \neq 0$ or $\mu \neq 0$, ξ depends on p; this is accounted for below (sections 6.3, 6.4).

[14] An analysis of transitional dynamics is provided in Becker, Kominers, Murphy, and Spenkuch (2018).

with Western countries took place during the 19th century, with some early signs of industrialization, but sustained growth in income per capita does not take place until the very end of the sample period.[15]

China's economy was largely based on agriculture, as was the case for most of the world during the sample period. Government administration followed a hierarchical structure from the emperor and upper ministries in Beijing down to provinces, prefectures, and counties, the outlines of which were inherited from previous dynasties. Thus, the Yuan Dynasty's rulers administered over China within the social and political structures inherited from the previous Song Dynasty (960-1279), a period during which many see a peak in Chinese economic development. While the governing structure varied, especially from the Yuan to the late imperial Ming-Qing period, the effective power of government was organized around an autocratic central authority and lower administrative regions. The state taxed lightly in international comparisons with other states. Moreover, the scope of markets for allocating scarce resources was more strongly limited by technology (e.g., transport technology, financial instruments) than by government regulation, in part because effective enforcement would have required more state resources than were in fact allocated. Recent research sees China's per-capita income at 95% of that of England in the year 1400, 77% by 1600, and 30% by 1800 (Broadberry, Guan, and Li 2017, Table 6).[16]

In this structure of governance by bureaucratic-scholar officials, education was a cornerstone of entry into higher income classes and a key aspect of the social contract between the state and local management. Consistent with low central taxation, the capacity of the state was limited. The provision of public goods was instead delegated to local (provincial and below) governments and the elites of local dynasties (Shiue 2004, 2005).

Local leadership was held by the elites who exercised paternal authority over their extended families (local dynasties in mirror to the imperial dynasties). At the same time, the participation of local elites in the political decision-making body of the state was legitimized through the state-run civil service examination system and its curriculum. Throughout the Ming-Qing period, participation in the tournament-style state exams and demonstration of literacy at the highest levels was the most direct path towards high income. Passing the civil service exam and obtaining a government office had substantial returns (see Appendix, section A). These returns were so attractive that even wealthy merchants tried to invest in their sons' educations in order to climb the social ladder. As Ho (1967) notes, social status derived from passing the civil service examinations was the "ultimate source of power" (p. 51). As a result, the earnings elasticity of human capital investments in China during my sample period is closely aligned to the civil service examination system.

Local leadership by elites who passed the civil exam strengthened and ensured the legitimacy of the central state on the one hand, while allowing local elites to earn high financial rewards on the other hand. There thus arose, in effect, a partnership between the central state and local elites, to mutual advantage. The result was a kind of equilibrium based on distinctively Chinese characteristics and different from, for

[15]See Keller, Li, and Shiue (2013), Brandt, Ma, and Rawski (2017).

[16]Due to incomplete data these figures come with a sizable margin of error. On the debate on the timing of the Great Divergence between China and Western Europe, see Pomeranz (2001, 2011) and Shiue and Keller (2007).

example, European forms of social organization (Mokyr 2002; Greif and Tabellini 2017).[17] While there is little controversy that local clans were an important underpinning of society, with the notable exception of de la Croix, Doepke, and Mokyr (2018) little is known at present about the economic impact of clans.

4 Data

4.1 Tongcheng Genealogies[18]

The data of this paper come from genealogies of individuals and households who lived in Tongcheng County of Anhui Province. Tongcheng County is approximately 30 miles by 60 miles and located just north of the Yangzi River, approximately 300 miles inland from the East China Sea. The distance is about 150 miles to Nanjing, the early Ming Dynasty capital, while Beijing, the later Ming and Qing capital, is about 650 miles away.

Tongcheng county had about 1.3 million inhabitants in the year 1790 (Beattie 1979), similar to the combined population of the newly independent states of Massachusetts, Connecticut, and New York at this time.[19] The county was representative of the more developed and densely settled regions of China, one of the many thriving economic regions near the lower Yangzi. The region was mainly a rice-producing area where the wealthiest families were typically landowners. Over the Ming and Qing Dynasties, the region gained some fame for having produced a number of the highest officials of the empire.

The dataset is created from genealogies of seven clans of Tongcheng County. Genealogies would primarily be compiled by the literate members of the clan, but government genealogical bureaus also assisted with the records (Pan 1929, Telford et al. 1983, and Zhao 2001). They originated as as a critical aspect of ancestral worship, explaining why e.g. burial locations of the deceased are carefully recorded in genealogies. The genealogy defined the membership of the clan, family rules of conduct, and played an important role in acknowledging the individuals who were entitled to the lands, the schools and other resources of the clan. Genealogies had key economic functions. Given these functions, it is plausible that they had an accurate record of their members, as that would have been in the interests of members. According to the genealogical principles of compilation, all male members of the clan were included.

I chose the Tongcheng genealogies because they are among those that combine a substantial level of demographic and socioeconomic detail with a high quality of compilation (Telford 1995). With 40,000 individuals including wives and children, the number of people is sizable. There is information on about 9,800 individual men in the sample. Around the year 1790, the size of my sample is 1.5% of the Tongcheng county population.[20] In the data, the clans' progenitor is recorded usually in the 14th century, with

[17]See also Hajnal (1965), De Moor and van Zanden (2010), and Dennison and Ogilvie (2014).

[18]Additional information in is given in the Appendix, section A. Shiue (2016) discusses Chinese genealogies more broadly. Earlier surveys on the content and scope of Chinese genealogies include Liu (1978) and Telford et al. (1983).

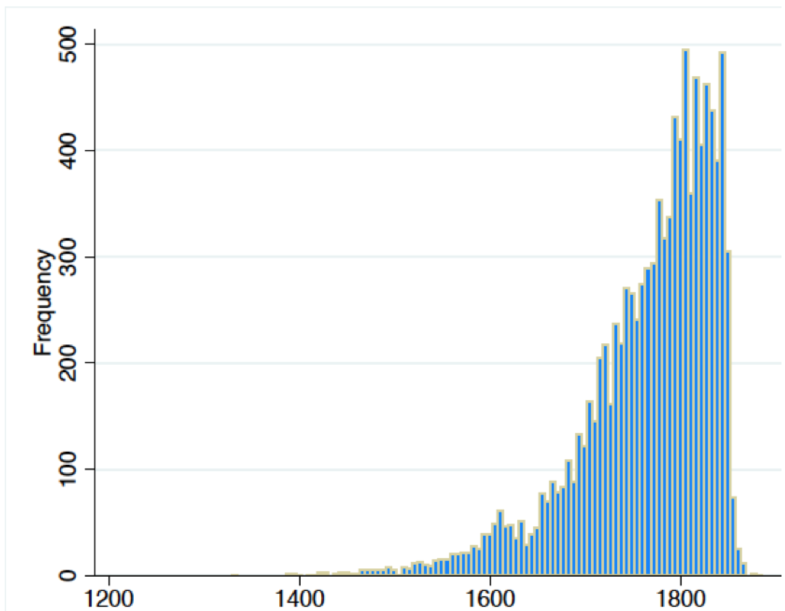
[19]The figure includes slaves for the US states; https://www.census.gov/library/publications/1793/dec/number-of-persons.html

[20]I observe about 3,600 men that would be alive in the year 1790 in my sample. These men had more than 4,200 wives,

the earliest date being the year 1298. The genealogies cover typically 18 consecutive generations, with a maximum of 21. The latest death recorded in the data set is 1925.

Figure 2 shows a histogram of the data over time for all seven clans in the Tongcheng genealogies. While my sample covers part of the Yuan and the Ming dynasties the large majority of observations are for the Qing (1644-1911).[21] The starting date depends on the birth date of the progenitor, but all seven clans are present throughout the sample period. Figure A.1 in the Appendix shows the relative size of each clan by sub-period.

Figure 2: Distribution of Data Entries



Notes: Number of father-son paired observations by birth year of son.

4.2 Income Distribution

Information in the genealogies includes the most noteworthy and highest achievements attained over a person's lifetime. These in turn determined a person's status in society, which is my measure of permanent income (income for short).[22] One important aspect of permanent income was whether a man had ever participated in the civil service examinations, and if so, whether he passed, whether he obtained an office, and at what rank. Other information relevant for a person's income includes various signs of wealth, for example whether a man ever set up large donations. The source does not give nominal amounts at the individual level, and my main analysis proceeds with six distinct income groups. I choose six groups because permanent income of these groups is clearly distinct and can be ranked. The six income groups are

and the data records more than 7,500 sons and 4,100 daughters, for a total of just under 20,000 persons. As noted above, Tongcheng had about 1.3 million inhabitants in 1790.

[21] Note also the dip in my sample surrounding the year 1644, consistent with the turmoil including deaths associated with the dynastic transition from Ming to Qing.

[22] Although there are a large number of women in the sample, women derived income from marriage.

summarized in Table 1. The role of these income groups for the results, including the level of aggregation into six groups, is discussed below (section 4 and Appendix, section A).

Table 1: Income Distribution

Group	N	%	Description	Annual Income
0	6,320	71.1	No signs of extraordinary income, wealth and status	5 *Taels*
1	1,766	19.8	Multiple consecutive marriage; educated w/o passing civil exam; family wealth	7.5 *Taels*
2	261	2.9	Landowners & merchants; multiple contemporaneous wives; official student	15 *Taels*
3	396	4.5	Official passing district/prefectural exam; student of imperial academy	30 *Taels*
4	108	1.2	Official passing provincial exam; individual passing national exam w/o office	60 *Taels*
5	18	0.2	Official passing national exam with top-level position	120 *Taels*

Notes: *Tael* is the silver currency in China during this time. Sample is all married males linked over three generations (son, father, grandfather); N = 8,892. Information is based on Chang (1955, 1962), Eberhard (1962), Ho (1967), and Telford (1986, 1992). See Appendix A for details.

Table 1 indicates that the data includes a relatively high fraction of persons with the lowest income, close to subsistence level. Men belong to this group when there is no sign of higher income, wealth, elevated status, or honors given in the data. The last column of Table 1 provides the annual income typical for these men, which I estimate at 5 *taels* per year (for more discussion, see Appendix, section A). Because the majority of China's population during the sample period in fact belonged to this group this is an advantage of using these genealogies for studying social mobility, compared to existing studies that typically only examine mobility among elites.

The second group is men whose annual income is above subsistence level (income group 1). The data indicates this in a number of ways. One is that a man has prepared for and participated in the civil service examination, even if he remained unsuccessful. Acquiring human capital for taking the civil service examination required spending a substantial amount of resources in addition to the opportunity cost of not working. Furthermore, in the context of this society it is plausible that men from families whose members had wealth or official position have above-subsistence income even if the men did not have exceptional careers themselves. Also men who married consecutively several times, typically due to the death of their first wife (divorce was not common) fall into this group. In China during this period marriage was typically a costly affair due to elaborate wedding festivities. I estimate the annual income of this group at 7.5 *taels* by calibrating to the average per-capita income in China.[23]

The next income group, group 2, belongs to the top 10% of society. For one, it includes official students. Because of their promise these men are released from other work and receive an annual government stipend. This rank of society includes also men who have contemporaneously multiple wives. While polygyny was not common (less than 1% in the sample), the income of these men is higher than that of men who marry

[23]See Appendix, section A. I examine the role of this and other income group determinations for my results below.

multiple times consecutively because the former must have for example a larger residence to accommodate the size of their household. Men with these means would typically have 15 *taels* of annual income. Also landowners, farmers, and merchants with substantial signs of wealth are in this group.[24]

The top 6% of society includes mostly men who became government officials at different levels. Since all of them had to participate successfully in the civil service examination, their permanent income can quite accurately be estimated by their degree and position. The highest level of the tournament-style government service exam a man had passed determined which official position he would be eligible for. Even those who did not hold office but were nonetheless educated at a certain level provided valuable managerial services that generated a comparatively high income. For those who gained official appointments, the position permitted a substantial accumulation of wealth. The lowest level of income which I designate as group 3 includes *sheng-yuan*, who are men who passed the lowest level of the civil entrance exam, as well as those who held a relatively minor office. Also included in this group are men who were students at the Imperial Academy in Beijing, typically with a higher level of income and stipend than official students elsewhere.

The income of the higher groups typically come from official salary, customary bonuses, and teaching activity, as well as landholding and merchant activities. At the same time, individuals of high rank were expected to aid and organize the provision of local public goods (irrigation, infrastructure, and other activities). Group 4, in the top 2% of the distribution, includes most importantly officials that have passed the provincial-level civil service examination (*ju-ren*), while income group 5 are top-level officials who have passed the national civil service examination (*jin-shi*). I estimate the typical annual income for groups 4 and 5 at 60 and 120 *taels*, respectively.

How do these figures line up with other information? Chang (1962) reports that the official salary of a provincial governor was 150 *taels* per year (his Table 1). In terms of my classification, a provincial governor would have a *jin-shi* degree and would have with 120 *taels* per year (group 5) a similar income. Below I also consider the possibility that income groups have changed over my sample period.[25] The figure of 5 *taels* per year in Table 1 for income group 0 is the low end of Chang's (1962) estimate of the per-capita income of a laborer (p.12). Overall, using the annual income figures of the last column of Table 1 I estimate that average income per capita in the Tongcheng sample was 7.8 *taels* per year, which is close to estimates for China as a whole (7.4 *taels* per year, Chang 1962).[26]

This definition of permanent income will underestimate social mobility to the extent that it misses mobility within each of the six groups. The influence of this for my results will be assessed by alternatively employing more than twenty instead of six income groups (see Table 5). In the baseline analysis I follow Dahl and

[24]Men included in group 2 are only those for whom there is substantial evidence of wealth and property. One might be concerned that based on their permanent income these men may belong into the top 5% or even top 1% of society, not only to the top 10% as assumed here. Exploring this I find that reassigning these men into a higher income category does not lead to very different results. This is in part because rich merchants and landowners that were not also high-ranking government officials were rare (see Table A.1 in the Appendix).

[25]See Table 8. Chang's (1962) figure is for the late 19th century.

[26]Both income and population figures rely on imperfect data; another estimate on annual per-capita income in the mid-19th century is 13 *taels* (Broadberry, Guan, and Li 2017). My estimate relies on the rank-specific information in Chang (1962) and as a result the average per-capita income is closer to his figure.

DeLeire (2008) and Chetty, Hendren, Kline, and Saez (2014) to employ the rank percentile of father and son in their respective income distributions. Ties are addressed by assigning all individuals of a given income level the mid-point of the percentile rank in the income distribution. For example, when 70% of all men have the lowest income level, each one of them is assigned the percentile rank of 0.35.[27] Alternative approaches to intergenerational mobility, including the use of transition matrices, are presented in section 5.

In addition to information on permanent income, the data provides information on birth year and birth month for each man, together with his clan, and generation in the clan. Furthermore, there is information on each man's wife (or wives) and their children. Recorded by couple in the raw data, the links between generations can be established with virtual certainty because male children re-appear in the genealogy as married men.[28]

The genealogies contain all male members of seven clans living in Tongcheng. For the purposes of studying intergenerational mobility, the main sample restriction is that a man's permanent income is only recorded for male clan members that not only appear as child but also as adult. Cases where clan members do not re-appear in the genealogies as adults may occur for three main reasons. The first is that during his lifetime the clan member may have migrated to a distant region of China so that his information is not included because it is not known in Tongcheng. However, migration over longer distances was quite rare (less than 2%, see Telford 1986). For the same reason, it is unlikely that out-migration has a major influence on my results. The second reason for unobserved permanent income is that the male clan member died as a child. The third reason why there is no income information for a male member is that he did not marry–single men's income is not recorded (due to the gender imbalance estimates of single men are about 10-20%). I will assess the robustness of my results in these respects by assigning the lowest income level to these men. This is reasonable because single men are quantitatively the most important reason for unobserved lifetime achievements and the main reason in this society why a man would not marry because of this is low income.[29]

I have compared the Tongcheng data with information from other sources, including official accounts by the Chinese state for a number of sub-populations, as there is no Census information during my sample period. These analyses confirm that the data is generally reliable, and several aspects of the data–the demographic characteristics including age-specific mortality rates in the Tongcheng data–are broadly in line with what we know for China during this period (see Appendix, section A). Also, the distribution of income groups is comparable to that (i) in other genealogies, (ii) in China-wide analyses of the higher parts of the income distribution, and (iii) in the Northern Eight Banner population (Fei 1946; Chang 1955, Elman 2000; and Campbell and Lee 2003, respectively). Furthermore, the information on the 18 *jin-shi* (highest income group 5) in the Tongcheng sample lines up with the official lists covering these top individuals from the Chinese state.

[27]In the same way researchers have addressed the frequent case of zero income (Chetty, Hendren, Kline, and Saez 2014).

[28]Vital statistics in form of birth year and month, as well as death year and month, birth order, clan generation and income group uniquely identify 99.4% of all men in the data.

[29]See Table 8 below for the analysis.

Overall, the evidence indicates that the Tongcheng data is suitable for studying intergenerational mobility during the sample period. In addition, my analysis below addresses a number of specific concerns such as the number of income classes as well as measurement error in income.

4.3 Summary Statistics

Summary statistics on status and basic demographic characteristics for N = 8,893 father-son pairs are given in Table 2.[30] Panel A of the table provides information on fathers, mothers, and sons. In the six group-classification of Table 1, the average income of fathers is 0.58, with a standard deviation of close to 1. The average birth year of fathers in the sample is 1732, which reflects the distribution of observations shown in Figure 2. The birth year of the first father in the data set is 1298, while the latest birth year of a son is in the year 1885. The average year of death of fathers is 1787.[31]

Table 2 also shows that about 11% of all fathers are educated (an indicator variable). Education is positively correlated with income but not identical. I define any man to be educated if at a minimum he prepared for the civil service examination (at any level), even if he did not succeed. These men would have had an extraordinary amount of human capital in this society, well above basic literacy.[32]

Turning to information on the wives and mothers, their average birth year is 1736, which is similar to that of fathers as one would expect. The earliest-born mother in my data is recorded for the year 1300, only two years different from the earliest-born father.

[30] There are 9,787 individual men in the Tongcheng genealogies for whom I have information on permanent income; lagging by one generation to identify grandfather income brings the number of observations down to N = 8,893.

[31] The difference, about 55 years, is higher than life expectancy in China during the sample period. This is because my sample of linked father-son pairs does not include those who died during childhood nor those who did not succeed to marry.

[32] More details on education by income class is presented in Table A.1. I show below that my results are robust to distinguishing additional levels of human capital (Figure A.15).

Table 2: Summary Statistics

Panel A.

		N	Average	Std. Dev.
Father	Income	8,893	0.58	0.99
	Year of Birth	8,893	1732.04	70.85
	Year of Death	8,659	1787.46	70.61
	Education	8,893	0.105	0.308
Mother	Year of Birth	8,893	1735.63	71.00
	Year of Death	8,251	1789.44	71.61
Son	Income	8,893	0.45	0.87
	Year of Birth	8,893	1763.90	71.30

Panel B. Clan Information

		Chen	Ma	Wang	Ye	Yin	Zhao	Zhou
N	Early	152	410	2,096	792	364	382	237
	Late	139	217	2,585	815	240	387	77
Income	Early	0.11	1.30	0.29	0.98	0.23	0.16	0.22
	Late	0.17	1.64	0.30	0.65	0.35	0.32	0.09
Distance to Tongcheng city	Early	49.11	22.98	79.91	46.82	71.55	128.36	51.72
	Late	74.64	22.99	76.25	49.76	66.14	122.21	62.79

Notes: Income levels 0 to 5, as shown in Table 1. Panel B. shows information on sons. Early is birth year son before 1780, Late is after 1780.

Sons have a somewhat lower average income level compared to fathers, and my analysis below accounts for such changes. The average birth years of fathers and sons implies that there are approximately 30 years between one generation and the next.

Panel B shows statistics separately for each of the seven clans by time period (first half versus second half of the sample). The most populous clan is the Wang, accounting for about half of the sample. Another large clan is the Ye, while the Chen and Zhou are relatively small. The Ma is the clan with the highest average income, followed by the Ye, while the lowest average income is obtained for the Zhou.

Which clan gained and which lost in terms of income over time? There is substantial variation across clans, with five clans gaining and two clans losing in terms of income. Average income for the large Wang clan hardly changed over time, in contrast to the also sizable Ye clan which experiences a substantial drop in income. In contrast, the other high-income clan, the Ma, sees an increase in income.[33] Overall, there is substantial variation in clan size and average income according to Table 2. Furthermore, between-clan differences in income are falling over time, with a standard deviation of 0.38 in the Early period, compared to 0.31 in the Late period. I will return to this below in the analysis of mobility over time in sections 5 and 6.

The last two rows shows information on the distance of the mens' residences in Tongcheng from the capital of Tongcheng. Figure A.3 shows the more than 200 residential locations of these seven clans in Tongcheng.

[33]The relationship between clan size and average income over time is mixed: for three clans the relationship is positive, while for four clans it is negative.

Rich men are more likely to live in the capital than poor men. For example, the men of the high-achieving Ma clan are on average 23 kilometers from the capital, in contrast to the Wang where the average distance is more than three times as high. While there is evidence that clans live in particular areas of Tongcheng–such as the Zhao, who tend to live further away than other clans–distance from the capital is correlated with changes in income. For example, the Ye and the Zhou whose average income falls over time also move further away from the capital.

In sum, while each clan on its own could hardly be expected to be representative of China, the seven clans taken together produce a relatively large and diverse sample. It also is key to to analyzing the role of groups for social mobility, see section 6.4 below.

5 Intergenerational Mobility in China

This section contains the main empirical results of the paper. I first examine examine intergenerational mobility in China over the period 1300 to 1900 as a whole before turning to changes in intergenerational mobility over time. As suggested by Fields and Ok (1999), multiple measures of intergenerational mobility are employed.

5.1 Mobility for the Period 1300 to 1900

I first analyze intergenerational mobility using transition matrices for the six income groups defined above. Among low-income families (Father group 0), almost 9 out of 10 of their sons remain in the lowest income group, see Table 3. In contrast, for sons from top-level *jin-shi* families (Father group 5), the chance of retaining *jin-shi* status is 7%, about 35 times the population share of *jin-shi*, and one in three of the *jin-shi* sons remains in the top 2% of the population (income groups 4 or 5). This points to a substantial level of intergenerational income persistence.

Table 3: Intergenerational Transitions between Income Groups

		Father Income						
		0	1	2	3	4	5	Total
Son Income	0	86.74	66.47	0	0	0	0	71.07
	1	11.39	23.95	66.33	61.43	28.12	10.71	20.20
	2	1.04	2.87	13.42	9.89	17.97	7.14	2.86
	3	0.75	5.16	15.70	21.88	35.16	46.63	4.45
	4	0.07	1.17	4.05	5.83	17.19	28.57	1.21
	5	0	0.37	0.51	0.97	1.56	7.14	0.21
	Total	100	100	100	100	100	100	100

Notes: Transition probabilities (%) between six income groups; columns sum to 100.

Moreover, the probability that a son rises to a given income level is generally increasing in the income of his father. For example, the chance of reaching status level 4 is 0.07% for a son of a group-0 father, and 1.17%, 4.05%, and 5.83% for a son of a group-1, group-2, and group-3 father, respectively. Conversely, the chance of dropping to the bottom of the income distribution is decreasing in father income. For example, the chance of falling to level 1 is about 11% for a son of a group-5 father, but 28%, 61%, and 66% for a son of a group-4, group-3, and group-2 father, respectively. These results shed a new quantitative light on intergenerational mobility in China.

Table 4: Percentile Rank Transition Matrix

		Income Father			
		[0,60]	[60,80]	[80,100]	Total
Income Son	[0,60]	$3,904$	$1,019$	412	$5,335$
		73.18	57.28	23.16	60.00
	[60,80]	963	404	412	$1,779$
		18.05	22.71	23.16	20.00
	[80,100]	468	356	955	$1,779$
		8.77	20.01	53.68	20.00
	Total	$5,335$	$1,779$	$1,779$	$8,893$
		100.00	100.00	100.00	100.00

Notes: Table gives N and transition probabilities (%) between the lower three, the fourth, and the fifth quintile.

Transition matrices based on income levels do not depend on the size of each group. An alternative that does capture the size of each group is a percentile rank approach, where every person in a given income group has the same percentile rank in the income distribution. For the following I have aggregated the lower three quintiles because this is roughly the size of the lowest income group.[34] Table 4 shows, for example, that there is an 8.8% probability that a son who is born to a father in the bottom 60% of the rank distribution is able to move into the top [80,100] quintile. The percentile rank approach underlying Table 4 is also employed in the following regression analysis.

Following a large literature, I estimate intergenerational mobility in a linear regression framework relating son income to father income:

$$R_S(i) = \alpha + \beta R_F(i) + \epsilon(i). \tag{11}$$

In equation (11), $R_S(i)$ and $R_F(i)$ are the percentile income rank of the son and the father, respectively, in pair i, i = 1,..., 8,893. The main difference between equation (11) and the model, equation (9), is that I employ the percentile ranks of father and son instead of their log incomes, since income is not observed. An alternative to percentile ranks is to employ the discrete group indicators, which leads to similar results (see below).[35] The mean-zero error term $\epsilon(i)$ captures all other influences affecting pair i, including parental human capital and clan effects; the latter will be addressed in sections 6.3 and 6.4 below.

A coefficient β close to one in equation (11) indicates low mobility because the son's rank in the income distribution is strongly determined by father rank, while there is relatively high mobility when β is close

[34]The men in the [0, 60] percentile group are randomly drawn from all individuals in income group 0.

[35]I compute percentile ranks of father and son separately for their respective birth cohorts, defined as a 25-year time window (such as years 1575 to 1600). Conditioning on the income distribution in each birth cohort means that changes in average income between birth cohorts will not affect the results. At the same time, my findings are robust to holding the income distribution fixed for the entire sample period, see Appendix Figure A.8.

to zero. I estimate β using OLS and obtain $\beta = 0.528$ (Table 5, column (1)). This means that a father with a 10 percentage points rank advantage over another father can expect that his son has about a 5 percentage point rank advantage over the other son. Put differently, the advantage in the father generation is approximately cut in half in one generation.

Table 5: Intergenerational Mobility in China, 1300-1900

	(1) Percentile Rank with 6 Groups	(2) Discrete Income with 6 Groups	(3) Percentile Rank with 23 Groups
β	0.528	0.552	0.570
(s.e.)	(0.009)	(0.010)	(0.008)
$E\left[R_S(i)\,\vert R_F(i) = 0.25\right]$	0.42	n/a	0.42
$E\left[R_S(i)\,\vert R_F(i) = 0.80\right]$	0.62	n/a	0.63

Notes: Estimation of β (equation (11)) by OLS; $E\left[R_S(i)\,|R_F(i) = 0.25\right]$ is the expectation of percentile rank of sons with fathers at percentile rank < 0.5. $E\left[R_S(i)\,|R_F(i) = 0.80\right]$ is the expectation of percentile rank of sons with fathers at percentile rank > 0.6. Robust s.e. given in parentheses. N = 8,893.

Table 5 reports also measures of upward and downward mobility. A relatively low-income father, at percentile rank 0.25, can expect that his son has rank 0.42 in the next generation's income distribution. In contrast, a high-income father, at rank 0.80, can expect that his son will be at rank 0.62 in the sons' income distribution. The direction of these moves confirm that there is regression to the mean while their magnitude provides information on the speed of intergenerational mobility.

The next columns in Table 5 present results for two alternative specifications. Employing six income groups valued at 0 to 5 yields a β estimate of 0.552 (column (2)). This is similar to the baseline estimate, see column (1), Table 5.[36] Column (3) in Table 5 reports results based on percentile ranks with 23 instead of 6 income groups, yielding a coefficient of 0.57.[37] The estimate implies somewhat more intergenerational persistence than the baseline (column (1)), indicating that a higher level of disaggregation does not necessarily yield more intergenerational mobility.

Overall, my estimates are in a relatively narrow range, with the intergenerational regression coefficient for the period 1300 to 1900 estimated to be somewhat above 0.5. A number of issues, including measurement error in permanent income and the influence of generations earlier than father, are examined in the Appendix, showing that the findings are quite robust (see section B). The value of 0.5, as noted in the Introduction, is higher than estimates for the United States around the year 2000 ($\beta = 0.34$, Chetty, Hendren, Kline, and Saez 2014) but similar to estimates for educational mobility in 19th century Germany (β around 0.5, Braun and Stuhler 2018). Keeping in mind the challenges for comparing β's across countries, intergenerational mobility does not seem to have been particularly low in China during the period 1300 - 1900. This is broadly in line with Ho's (1967) findings based on top-level officials.

[36]Employing annual income (last column of Table 1) in logs instead of permanent income yields a coefficient $\beta = 0.522$, robust s.e. of 0.012, which is also similar to the baseline estimate in column (1) of Table 5.

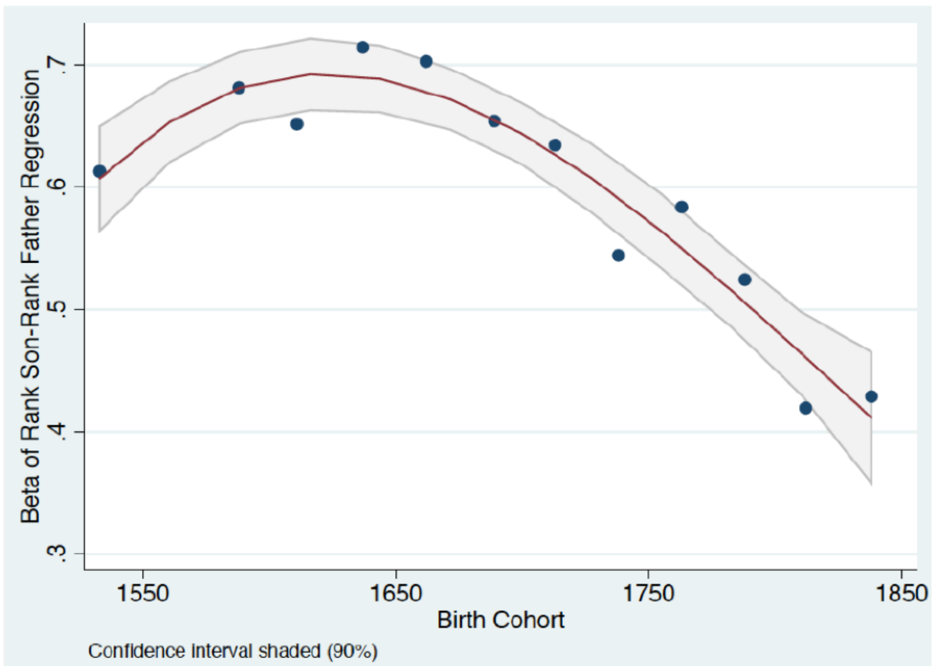
[37]See the Appendix, Table A.1, for the 23 groups.

5.2 Intergenerational Mobility over Time

This section provides evidence on intergenerational mobility over time. I begin by estimating the intergenerational coefficient β for different birth cohorts between 1300 and 1900, turning to alternative mobility measures and robustness checks towards the end of the section.

Evidence on changing mobility over time comes first of all from estimating equation (11) for successive birth cohorts (denoted by c) during the overall sample period of 1300 to 1900. Figure 1 above has shown the pattern of β_c for forty equal-sized birth cohorts. Figure 3 shows the evolution of β_c for twelve birth cohorts of twenty-five years each (such as 1575 to 1600; a varying number of observations in each cohort).[38]

Figure 3: China's Intergenerational Mobility Over Time - 12 Birth Cohorts



Notes: Shown are β_c coefficients from OLS regressions of percentile rank son on percentile rank father (equation (11)). Ten twenty-five year birth cohorts (1575 to 1825), plus birth cohorts before 1575 and after 1825. Median birth year of son in cohort given on horizontal axis. Earliest son birth year in sample is 1330.

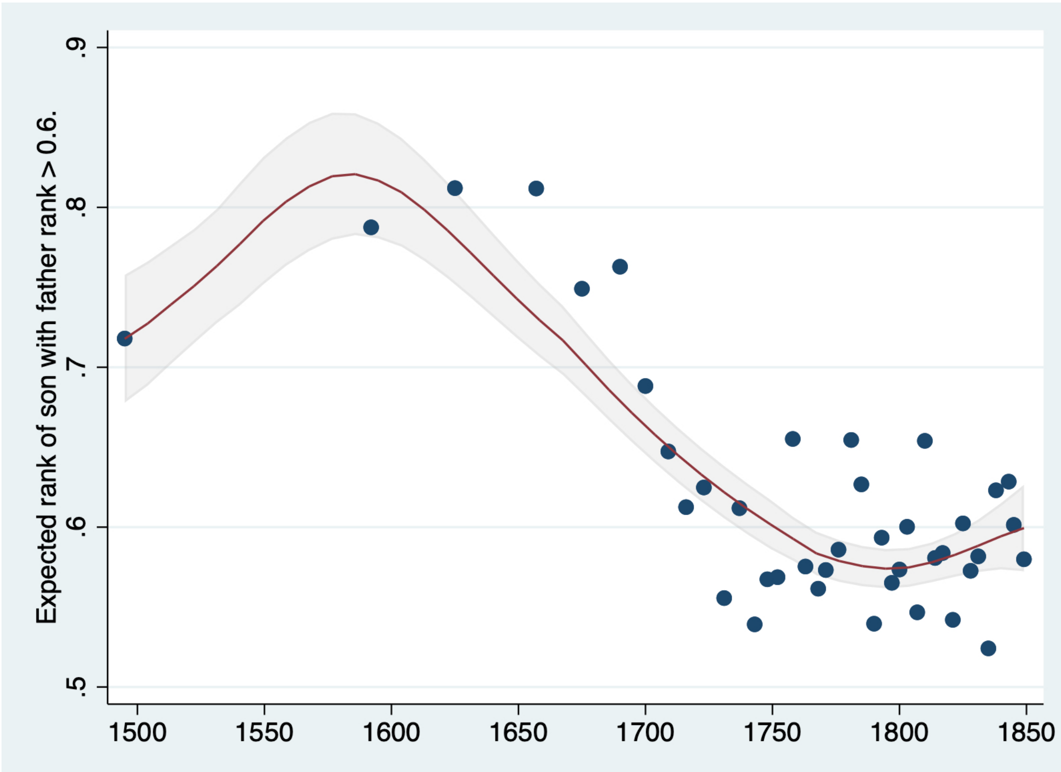
Figure 3 confirms the earlier pattern: mobility was low through the 17th century (high β_c) before it increased by the 19th century, with values of β_c close to 0.4. The finding of increased mobility over time holds irrespective of how I define a birth cohort.

Downward Mobility Figure 4 presents evidence on changes in downward mobility by showing the expected percentile rank of a son whose father is at the 80th percentile of the income distribution. Around the year 1550, there is relatively high persistence, as evidenced by a fall of less than ten percentile ranks for sons of these high-ranking fathers. Around the year 1625, such sons could even expect not to fall at

[38]Exceptions are the first and the last cohort, which due to data availability are defined as sons born before 1575 and after 1825.

all in the income distribution relative to their father. Starting around the year 1750, however, downward mobility has increased and the expected fall of sons with fathers at the 80th percentile is typically more than 20 percentile ranks.

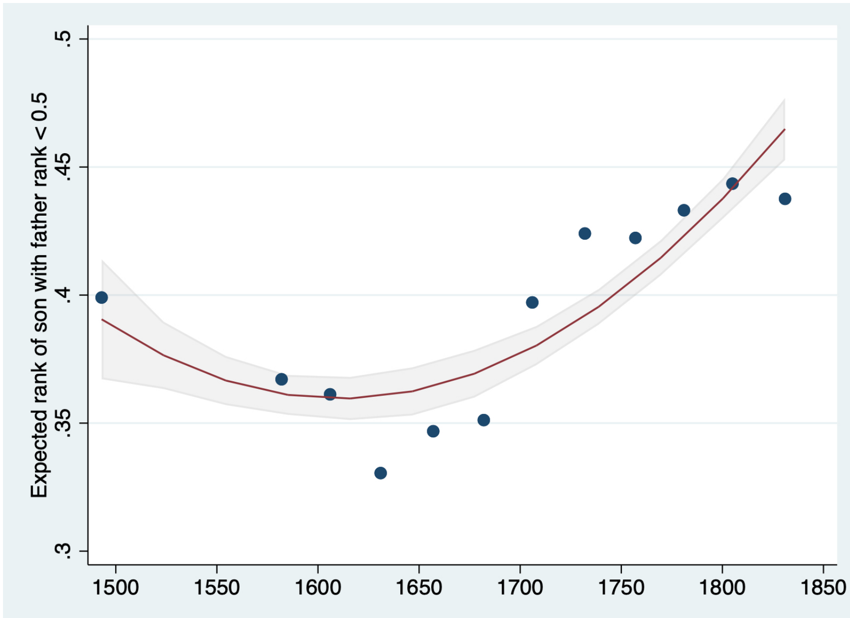
Figure 4: Downward Mobility over Time



Notes: Shown is expected rank of son given that father has rank > 0.6. Forty equal-sized birth cohorts. Earliest birth year of son is 1330.

Upward Mobility Turning to upward mobility, the expected rank of sons with fathers at the 25th percentile in the income distribution is shown in Figure 5. We see that around the year 1500, these sons would typically climb by about 13 percentile ranks in the income distribution, to about the 38th rank, while by the early 19th century the typical upward movement was closer to 20 percentile ranks (to about the 45th rank).

Figure 5: Upward Mobility Over Time



Notes: Shown is expected rank of a son with father at rank < 0.5. Ten twenty-five year birth cohorts (1575 to 1825), plus birth cohorts born before 1575 and born after 1825. Earliest son birth year is 1330.

The results on upward and downward mobility confirm the trend towards greater mobility over time that is evident from changes in the intergenerational coefficient β_c (see Figures 1, 3). Furthermore, while there are small differences, the period of highest persistence for all measures of mobility is generally found in the 17th century.

Transition Matrices The following results present evidence on changes in intergenerational mobility based on transition matrices. I distinguish between the earlier and the later half of the sample. Given the distribution of the data in Figure 2, this analysis sheds light on whether transition probabilities around the year 1825 are different from the transition probabilities in the earlier part of the sample.

Table 6: Transition Matrix for First Half of Sample

		Father Income						
		0	1	2	3	4	5	Total
Son Income	0	91.61	66.90	0	0	0	0	71.28
	1	6.22	22.07	62.32	60.99	28.89	0	17.87
	2	1.31	4.37	16.20	9.89	20.00	18.18	3.99
	3	0.75	5.75	16.90	24.45	37.78	36.36	5.55
	4	0.11	0.80	3.87	4.12	11.11	27.27	1.11
	5	0	0.11	0.70	0.55	2.22	18.18	0.20
	Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: Transition probabilities (%) between six status levels before the year 1780.

Table 6 shows that in the earlier period, more than 91% of sons from lowest-income families stay in the bottom group of the distribution. Also, a remarkable 18% of top-level *jin-shi* families (income group 5; top 0.2% of the distribution) stay in that income group from one generation to the next. As Table 7 indicates, the corresponding probabilities in the later period have fallen to 82% and 0%, respectively. Fewer sons from poor families stay poor, and fewer sons from very rich families stay very rich - upward and downward mobility, respectively, has increased. Importantly, this result is not driven by a change in share of these income groups in the population. As the last columns in Tables 6 and 7 indicate, the population shares of the poor (income group 0) and the very rich (income group 5) are with approximately 71% and 0.2% quite similar in the earlier and later period.

Table 7: Transition Matrix for Later Half of Sample

		Father Income						
		0	1	2	3	4	5	Total
Son Income	0	82.22	66.11	0.00	0.00	0.00	0.00	70.85
	1	16.19	25.57	76.58	62.06	26.32	17.65	22.51
	2	0.79	1.59	6.31	9.88	13.16	0.00	1.73
	3	0.76	4.66	12.61	18.18	28.95	52.94	3.36
	4	0.03	1.49	4.50	8.30	31.58	29.41	1.32
	5	0.00	0.59	0.00	1.58	0.00	0.00	0.22
	Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: Transition probabilities (%) between six status levels in the year 1780 and after.

Shorrocks (1978) has proposed a measure of the mobility of a transition matrix based on its trace, defined as $M = \frac{n - trace\, P}{n-1}$, where P is the square transition matrix and n its dimension. The higher is the trace, the lower is mobility. In the present case, M for the earlier period is 0.83 while M for the later period is 0.87. Thus, evidence from transition matrices confirms that mobility has increased in China during my sample period.

Additional evidence for the increase in mobility comes from a multigenerational approach that relates income of the son to the four generations (father to great-great-grandfather) before him, see Appendix, Figure A.9.

Aggregation and Definition of Income Table 8 reports the strength of the correlation between the baseline results on mobility over time and the corresponding results for five alternative specifications. The baseline results are for β_c, the birth-cohort specific intergenerational regression coefficient (shown in Figure 3), for the expected rank of a son whose father is at percentile rank 0.25 (denoted by Upward in Table 8), and for the expected percentile rank of a son whose father is at rank 0.8 in the income distribution (Downward in Table 8). The five alternative specifications are (1) a more disaggregated approach with 23 income groups, (2) a broader sample to include single men, (3) a lower income group for men from families with significant wealth or official positions but who did not achieve extraordinary income levels themselves, and (4) substantial measurement error in father income.[39] The fifth alternative is a lower income group for minor officials in the second half of the sample, in line with some accounts of this period (column (5)).

Table 8: Correlation of Baseline Results with Alternative Specifications

Mobility Measure	(1) 23 Groups	(2) Single Men	(3) Substantial Family Income or Wealth	(4) Measurement Error	(5) Lower Rank Officials > 1780	(6) Average
Reg. Coefficient	0.995	0.922	0.820	0.968	0.966	0.934
Upward	0.993	0.941	0.926	0.983	0.976	0.964
Downward	0.730	0.962	0.907	0.985	0.968	0.910
Average	0.906	0.942	0.884	0.979	0.970	

Notes: Correlation for three mobility measures with the baseline specification for five alternative specifications, as described in the text; twelve birth cohorts.

Table 8 indicates that while there are some differences, on the whole the results on the evolution of intergenerational mobility in China are confirmed in form of high correlations. In particular, column (6) of Table 8 shows that across five alternative specifications the pattern of β_c has an average correlation of 0.934 with the β_c's of the baseline shown in Figure 3.

These results on the changing pattern of mobility in China are confirmed by a range of extensions, see Appendix section C. Overall, these results show that my findings on the evolution of intergenerational mobility in China hold quite broadly under alternative specifications.

6 Explaining Temporal Variation in Intertemporal Mobility

6.1 Changes in the Earnings Elasticity of Human Capital

In the model, differences in the intergenerational regression coefficient β_c are due to differences in $\gamma\sigma$, where γ is the productivity of parental investment in human capital while σ is the earnings elasticity of human capital (see equation 9).[40] Because large-scale industrialization did not occur in China during the sample period, changes in σ will mainly be associated with changes in the civil service examination system. My estimates show that until 1700 mobility was relatively low, before it increased towards the end of the sample period (Figures 1, 3). Furthermore, there is some evidence that mobility was lowest around the year 1650 (see Figures 3, 4, 5). A number of key changes related to the civil service examination, I show in the following, affected the earnings elasticity of human capital in a way that explains these findings.

First, consider the period before the year 1700. Over time, there were fewer institutionalized and legal

[39] In column (3), groups 3, 4, and 5 in the disaggregated classification of Table A.1 are in group 0, not 1 in the baseline income distribution with six groups (Table 1). Column (4) introduces measurement error in father income with mean zero and a standard deviation of 0.5, as in Table A.5, column (2).

[40] This is as long as $\kappa = \mu = 0$, which will be relaxed below.

barriers to upward mobility. Occupational mobility during the Yuan dynasty was limited due to the way the state extracted labor services from the civilian population. To enforce the obligations of labor service, the state required households to be registered in segregated occupational groups — commoners, artisan, soldier, salt producer, miner, scholar, astrologer, and other categories. These occupational groups were hereditary and also largely determined the position of an individual in the income distribution. The practice of compulsory occupational registration was carried over from the Yuan to the Ming dynasty but by the latter half of the Ming the family histories of some high ranking officials who came to prominence in the 15th century reveal backgrounds in the artisan or other occupational groups, an indication that the hereditary nature of the categories had started to break down during the Ming (Ho 1967). While aspects of the older practice of hereditary occupations were still present, occupational mobility became more fluid as the Ming started to become laxer in enforcing labor obligations and instead resorted more frequently to paying wages for desired services. The Qing Dynasty discontinued the practice of family occupational registration altogether, by which time there were no effective legal barriers to mobility due to any pre-designated occupational status of the family (Ho 1967).

Hereditary aristocracies started to be eliminated sometime during the Tang dynasty (670-906 AD). By the start of the Qing in 1644, the only types of hereditary privileges and automatic status that remained belonged to the imperial dynasty, where the throne was passed from the emperor to one of his sons, and the leading families of the Eight-Banner system.[41]

The lifting of occupational restrictions and hereditary privileges meant that for many individuals investing into human capital became a viable strategy of upward mobility. Beginning in the Song dynasty (960 AD), some members of the elite and high-ranking officials of the state were selected on the basis of formal examinations. Sumptuary laws, which prevented commoners from taking the civil examinations, were eliminated in the later part of the Ming.[42] The new institutions, which involved relying more heavily on the civil service examinations to test a candidate's knowledge, became a central aspect of governance with the beginning of the Qing dynasty (1644). Institutions governing the higher ranks of society also gradually changed over time. National civil service examinations became more routine in nature, and over time, regional quotas for the various academic degrees that existed earlier became less discretionary starting in the year 1371. The shift to a more rules-based system lowered uncertainty regarding the return to human capital investments. Overall, changes between 1300 and 1700 lifted the return to human capital investments for a substantial part of China's population. As a consequence, σ was comparatively high and intergenerational mobility was low.

The trend towards lower intergenerational mobility starts around the year 1700, and by the 1800s β has fallen by about 20 percentage points to around 0.45. This is consistent with a falling earnings elasticity of human capital. The change in the officials' lifestyle and spending patterns, based on information from

[41]The Eight-Banner system was a hereditary institution that dominated certain military and command functions of the Qing state; banner families were eligible for special privileges.

[42]The biographies of the successful exam candidates of the Ming period show that the family backgrounds of the highest level *jin-shi* degree holders included also people from those special occupational status designations, such as the soldiers, army officers, horse breeders, medical officials, official cooks, and others (see Li Zhou Wang 1746).

biographies, clan genealogies, as well as local histories (gazetteers) indicate that the returns from official position fell during this period.[43]

Furthermore, the support for official students fell as well. During the Ming dynasty, students received generous allowances in the form of food (two bushels of rice), cotton and silk cloth, embroidered silk cloth, as well as sets of clothing, headgear, and boots, as well as travel money to go home to visit family. They were also given holiday money, money to help pay for their wedding, grain to support their wives and children, and other goods (Ho 1967). In contrast, during the Qing students received only a minimal grain stipend and tax exemptions from the state (Cong 2007). Moreover, income derived from teaching fell over time. Men who passed the lower level exams but repeatedly failed to pass the intermediate or higher exams could not obtain an official position, with the result that many of these individuals resorted to a livelihood as a schoolteacher. Towards the end of the Qing, the income from teaching was so low that in some cases teachers could not sustain themselves on teaching alone, forcing them to take up other tasks (Chang 1962, Ho 1967). Overall, teachers' real wages were lower in the 19th century compared to the earlier Qing period (Rawski 1979, Figure 1).

Ultimately, the decline in the return to human capital was driven by a decline in the fiscal capacity of the Chinese state, which guaranteed the return to official position and human capital more generally. China's population was growing from 140 to above 400 million from the years 1650 to 1850 and per-capita fiscal capacity fell, see Figure 6.[44]

[43] One example is the case of Chiang Chi, who spent 300,000 taels in the 18th century on constructing roads around his native place, after having served for ten years as official on the Board of Punishment. In contrast, comparably high sums for the 19th century are much less common (Chang 1962).

[44] The increase in population with a constant number of high-ranking official positions (1,385 in the Ming and 1,360 in the Qing) also meant that a given human capital investment had a lower expected return.

Figure 6: China's Fiscal Capacity, 1650 to 1850

Notes: Shown are log per-capita tax receipts. Source: Sng (2014).

Given the decline in fiscal capacity, as well as the lack of any major offsetting increase in human capital returns due to industrialization, the earnings elasticity σ fell during this period. This can explain the increase in intergenerational mobility.

How about changes in γ, the productivity of human capital investment technology, which is the other determinant of intergenerational mobility in equation (9)? There is not much evidence, that γ changed over time, or that there was a general decline in the productivity of human capital investments during the later Qing. At the same time, a lower earnings elasticity of human capital may have affected the contribution of parents' human capital or clan resources to human capital formation (see equation 3). The next section will consider the relationship between inequality and the earnings elasticity before I return to the role of clans for mobility.

6.2 Cross-Sectional Inequality and Earnings Elasticity

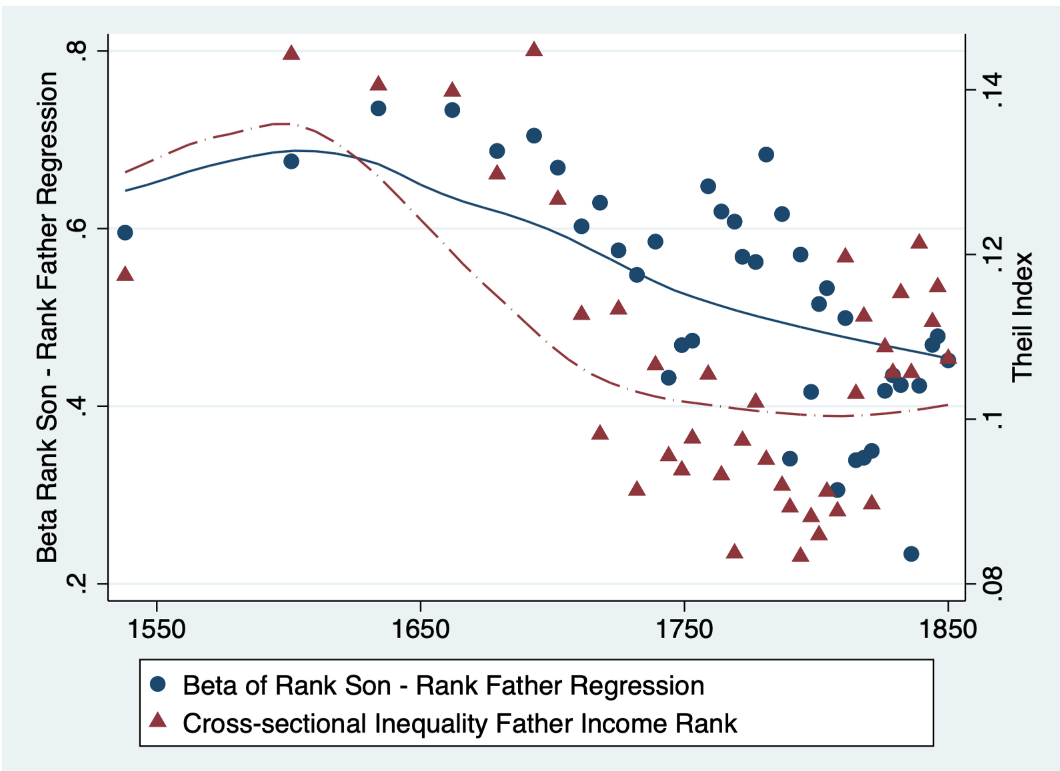
According to the model, changes in the earnings elasticity of human capital do not only impact intergenerational mobility but also cross-sectional inequality. Specifically, inequality is increasing in the earnings elasticity of human capital (see equation 10). To provide empirical evidence, the following examines the relationship between mobility and inequality over my sample period.

The measure of inequality I employ the Theil (1967) index.[45] It is defined as

$$T_c = \frac{1}{N_c}\sum_{i=1}^{N_c} \frac{R_F(i)}{\mu_c} ln\left(\frac{R_F(i)}{\mu_c}\right) \tag{12}$$

where $R_F(i)$ is the percentile rank status of the father in father-son pair i belonging to cohort c, μ_c is the average percentile father rank in cohort c, and N_c is the number of fathers in cohort c, $c = 1, ..., 40$. The circles (fitted with a solid line) in Figure 7 show intergenerational mobility of sons in the different cohorts; for each of these cohorts the figure also shows a triangle (fitted with a dash-dot line) that gives the cross-sectional income inequality observed in the father's generation.

Figure 7: Income Inequality and Intergenerational Mobility



Notes: Shown are β_c and the Theil index of income in the father generation for forty birth cohorts. Earliest son's birth year is 1330.

Figure 7 shows that times of relatively high inequality are times of high β_c (low mobility). Intergenerational coefficients of 0.6 and above before the year 1700 are found when the Theil index is equal to 0.12 and higher. By the year 1850, when β_c has fallen to 0.45, the Theil index is lower as well. The correlation between the intergenerational coefficient β_c and T_c is equal to 0.43. In addition, accounting for the geographic location of individuals, in particular how close from the administrative center of the region they live, strengthens

[45]The Theil index is a well-known member of the single parameter Generalized Entropy Class indices; I choose it in part because it is additively separable (see below). See Bourguignon (1979), Shorrocks (1980) for overviews of inequality measures. My main results are similar when I use other measures such as the Gini index, see below.

the relationship between intergenerational mobility and inequality; see Figure A.10, Appendix.

Not only the intergenerational regression coefficient but also my nonparametric measures of upward and downward mobility are correlated with cross-sectional income inequality. In particular, income inequality is strongly negatively correlated with upward mobility and strongly positively correlated with downward mobility (correlations of -0.80 and 0.75, respectively, see Figures A.12 and A.13, Appendix).

This section has provided evidence that the factors determining intergenerational mobility are closely related to those that determine cross-sectional inequality. This confirms the prediction of the model and strengthens the support for human capital investments being the driving force behind mobility changes in China.

6.3 Parental Human Capital

This section examines the role of parental human capital for intergenerational mobility. According to the model, human capital of the child is increasing in parental human capital for any given level of investment as long as $\kappa > 0$ (equation 3). I present two pieces of evidence for this. First, if parental human capital positively affects child human capital for a given investment, father education enters the intergenerational regression equation as an additional regressor (see equations (8), (9)). To examine this I extend equation (11) to include an indicator for father education, $E_F(i)$,

$$R_S(i) = \alpha + \beta_1 R_F(i) + \beta_2 E_F(i) + \epsilon(i). \tag{13}$$

Estimating equation (13) by OLS yields $\beta_2 = 0.24$ (s.e. 0.01) for father education, while β_1 now is 0.36 (s.e. $= 0.01$). The elasticity of son income with respect to parental human capital in the model is equal to $\kappa\sigma$, and $\beta_2 = 0.24$ provides evidence that $\kappa > 0$.[46]

Furthermore, the elasticity of child income with respect to parent income is equal to $\gamma\sigma$, so that $\beta_2/\beta_1 = \kappa/\gamma$ is an estimate of the role of human versus nonhuman investments. This suggests that parental human capital has about 2/3 of the effect of parental nonhuman investments. In fact, the relative role of human investments is even higher once I account for the difference in the variances of $R_S(i)$ and $E_F(i)$. The ratio of standardized regression coefficients (with mean zero and variance of one) $\widetilde{\beta_2}/\widetilde{\beta_1}$ is equal to 0.33/0.38, that is, the roles of human and nonhuman investments are similar in magnitude. This provides new evidence that parental inputs unrelated to monetary investments strongly affect child income as well (see also Mogstad 2017).[47]

Second, a positive impact of parental human capital on child income conditional on nonhuman investment (that is, $\kappa > 0$) means that educational inequality may be an even better predictor of intergenerational mobility than income inequality. To examine this, I obtain the Theil (1967) measure of educational

[46] A caveat is that neither $R_S(i)$ nor $E_F(i)$ enters equation (13) in logs.

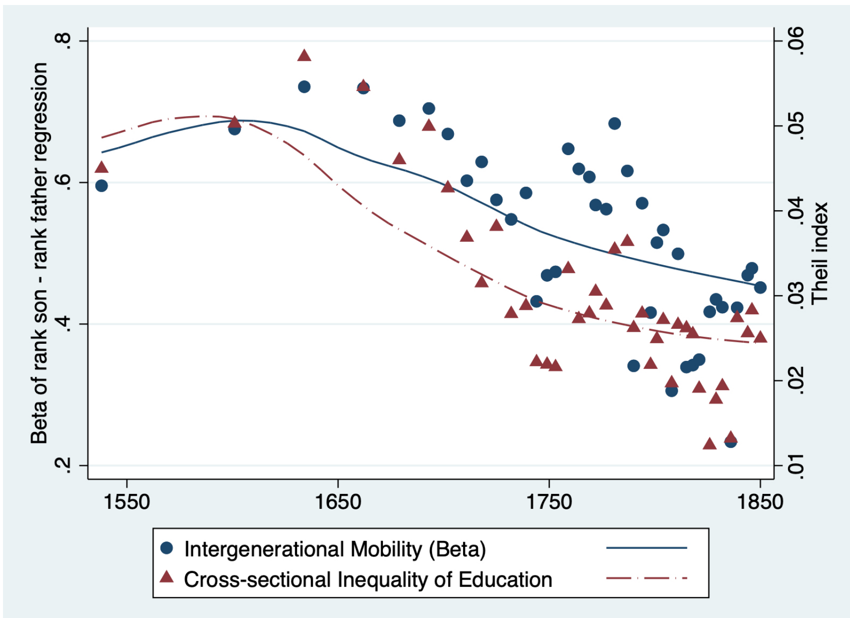
[47] I have also estimated equation (13) birth cohort by birth cohort to examine changes in the role of parental human capital over time. For most of the sample period β_2 is estimated to be about 0.2, except at the very end of my sample period when it is higher.

inequality by birth cohort c as

$$T_c^e = \frac{1}{N_c}\sum_{i=1}^{N_c}\frac{E_F(i)}{\mu_c^e} ln\left(\frac{E_F(i)}{\mu_c^e}\right) \tag{14}$$

where $E_F(i)$ is an indicator variable for the education of the father in pair i belonging to cohort c, and μ_c^e is the average father education level in cohort c. One is added to all values of the indicator. Figure 8 shows the relationship between the intergenerational regression coefficient β_c in equation (11) and educational inequality as measured by T_c^e.

Figure 8: Inequality in Education Intergenerational Mobility



Notes: Shown are β_c (equation (11)) and Theil index of educational inequality in the father generation. Horizontal axis shows median birth year in birth cohort; first birth year is 1330.

As Figure 8 indicates, the relationship between educational inequality and the intergenerational coefficient β_c is positive. Furthermore, with a correlation of 0.83 the relationship between mobility and educational inequality is stronger than between mobility and income inequality (the latter is 0.43, see Figure 7). Results for multiple levels of education (as opposed to an education indicator variable) and for alternative measures of inequality in the Appendix show that these results are robust (Figures A.15, A.16).

Overall, there is strong evidence that parental human capital influences intergenerational mobility in addition to the nonhuman investments made by parents.

6.4 Group Effects: The Role of Clans for Mobility

Another important question is the role of groups for mobility, specifically in this context, the role of clans. The importance of clans for mobility can be assessed in a number of ways. In the following, the starting point is the strong relationship between cross-sectional inequality and mobility. I examine the role of clans by asking whether inequality between clans matters for intergenerational mobility. One reason for this may be that members of the same clan support each other. For example, there may be resource pooling among clan members when supporting a small group of young clan members to prepare for the civil service examination. Alternatively, clan membership may play a role in the allocation of official positions and other financially rewarding endeavors once a clan member has obtained high office. Evidence on clan rules is entirely consistent with these possibilities, see Appendix section D.4.1.

Let $m_{F,c}^{D}$ denote the average percentile rank of all fathers of clan D in father-son pairs i belonging to cohort c. The Theil measure of between-clan inequality in cohort c is defined analogously to equation (12), with individual father percentile rank being replaced by the group (clan) mean, $m_{F,c}^{D}$:

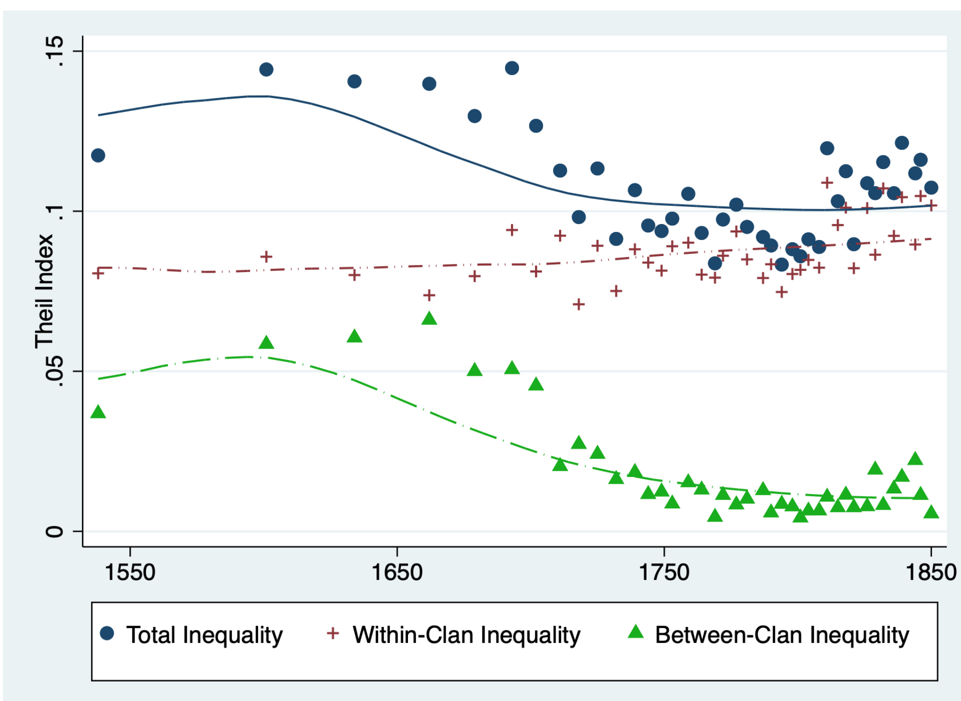
$$T_c^B = \frac{1}{N_c}\sum_{i=1}^{N_c} \frac{m_{F,c}^{D}}{\mu_c} ln\left(\frac{m_{F,c}^{D}}{\mu_c}\right) \tag{15}$$

The between-dynasty component is the inequality that would be observed if income levels were identical within each clan. Given the additive properties of the Theil index, within-clan inequality is total minus between-clan inequality:

$$T_c^W = T_c - T_c^B. \tag{16}$$

Figure 9 shows the decomposition into between- and within-clan inequality.

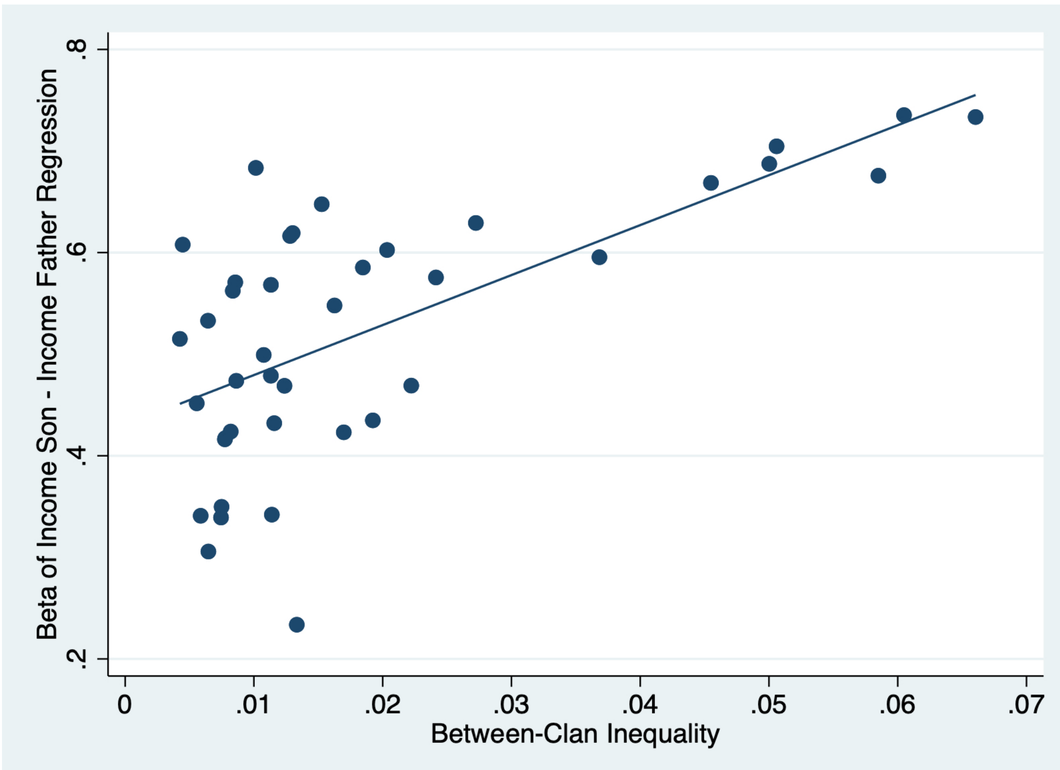
Figure 9: Decomposition into Between- and Within-Clan Inequality



Notes: Total, within-, and between clan inequality in forty birth cohorts; horizontal axis is median birth year of cohort. Earliest son birth year is 1330.

Figure 9 shows a number of results. First, between-clan inequality accounts for about 20% of total inequality in the sample. Thus, even though there are differences across clans, the majority of income inequality is due to individual variation. Second, the relative contribution of between-clan inequality to total inequality has fallen over time, with differences in income between clans accounting for only 5% of total inequality by 1850. Correspondingly, within-clan inequality has increased over time (equation 16). Third, the evolution of total inequality follows that of between-clan inequality. Indeed, seventy percent of the variation in total inequality can be explained by variation in between-clan inequality.[48] Thus, to understand changes in inequality one focus should be on between-clan inequality.

Figure 10: Intergenerational Mobility and Between-Clan Inequality



Notes: Shown are β_c and between-clan (Theil) inequality across forty birth cohorts. Horizontal axis shows median birth year in birth cohort; first son birth year is 1330.

Next, Figure 10 shows that the intergenerational coefficient β_c is positively correlated with between-group inequality T_c^B (correlation of 0.66). Times of high between-group inequality are times of social persistence. This is evidence that clans affect the degree of social mobility in China, and it indicates that part of the earnings elasticity of human capital that drives intergenerational mobility is clan-specific. Resource transfers between clan members to maximize clan income is a plausible explanation. At the same time, other explanations such as changing aspirations or cultural values at the clan level may play a role as well.

Another way of examining the role of groups for mobility is to add clan fixed effects to the regression equation (11). These fixed effects capture differences across clans that may affect son income, such as the average income of the clan. The intergenerational coefficient β is then estimated at 0.48 (s.e. 0.01), which

[48] From an OLS regression of total on between-clan inequality. In contrast, within-clan inequality changes account for only 8% of the variation in total inequality.

compares with an estimate of 0.53 without clan fixed effects (see Table 5). The result is consistent with a small positive clan effect. Do clan fixed effects help to explain changes in mobility? No. The evolution of the estimate β_c is very similar if one estimates the intergenerational equation cohort by cohort with and without clan fixed effects (correlation of 0.98, see Figure A.17). Thus, allowing for arbitrary time-varying clan-specific intercepts leaves the evidence on mobility changes unaffected. Rather, it is between-clan inequality that closely tracks changes in intergenerational mobility.

6.5 Extension: Multi-Generational Mobility

Another question is the role of considering only two generations, father and son, for the results. Recently it has been emphasized that mobility may be affected by earlier generations (multigenerational mobility, e.g. Solon 2014, Long and Ferrie 2018). Therefore, the following shows evidence on multigenerational mobility with five generations, son to great-great-grandfather. These quintuples cover typically about 150 years of time. See Table 9 for the results.

Table 9: Multigenerational Mobility in China, 1300-1900

	(1) Baseline	(2) Baseline 5-Gen Sample	(3) Multi-Generational
Father	0.528	0.526	0.467
	(0.009)	(0.009)	(0.014)
Grandfather			0.021
			(0.017)
Great-Grandfather			0.046
			(0.014)
Great-Great-Grandfather			0.048
			(0.012)
N	8,893	7,319	7,319

Notes: Estimation of extended versions of equation (11). Robust s.e. given in parentheses.

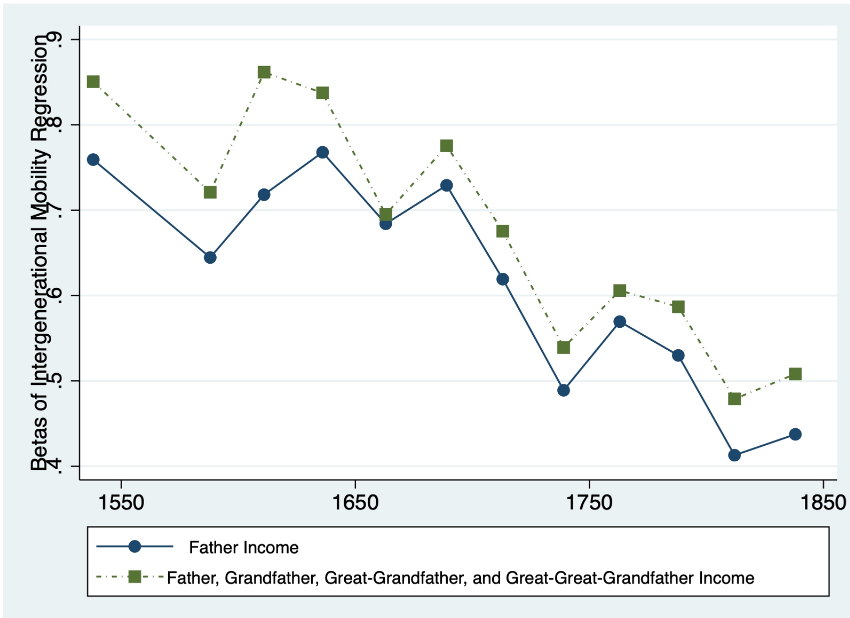
Column (2) of Table 9 shows the two-generation son-father regression with the five-generation sample.[49] I find that the coefficient on father income is similar to before. The next specification includes grandfather, great-grandfather, and great-great-grandfather income as additional regressors. The coefficient on father income falls to about 0.47 while results for additional generations are positive (column 3; grand-father marginally significant). One interpretation of this is that part of the father effect shifts to earlier generations.[50] Note that the sum of the coefficients on four generations, father to great-great-grandfather, is about 0.58 in column (3), which is above the point estimate when only father is included (column 2). This

[49] The number of observations is now lower due to intergenerationally lagging additional generations (N = 7,319 instead of N = 8,893).

[50] Another explanation is a serially correlated endowment process, which predicts a non-zero coefficient on grandfather (Solon 2014, Lindahl, Palme, Sandgren-Massih, and Sjögren 2014, and Becker and Tomes 1979). The fact that the point estimate on grandfather is lower than that for great-great-grandfather provides evidence that the coefficients do not reflect the strength of personal interaction of the son with earlier generations.

is additional evidence that a multigenerational approach may lead to less intergenerational mobility than the two-generation approach. At the same time, the change over time in mobility I find is similar to when a two-generational approach is employed, see Figure 11.

Figure 11: Intergenerational Mobility and Between-Clan Inequality



Notes: Shown are β_c on father income as well as the sums of coefficients on Father, Grandfather, Great-Grandfather, and Great-Great-Grandfather. Horizontal axis shows median birth year of son in birth cohort; first son birth year is 1330.

Based on Figure 11 there is little reason to believe that adopting a multigenerational approach will change this paper's key findings.

7 Conclusions

Employing a sample in central China between the years 1300 and 1900, with a median year of observation around 1780, I estimate an intergenerational regression coefficient of about 0.5. Parental advantage is halved in the generation of the children. This level of social mobility is not unusual in the light of what we know about mobility at later times in other countries. The value of 0.5 is found for countries such as Italy and the United Kingdom in the late 20th century, as well as for intergenerational mobility in education in Germany since the late 19th century.[51]

Intergenerational mobility in China has changed over time. Although mobility was relatively low until the late 17th century with intergenerational coefficients of around 0.7, I have documented a striking increase in mobility through the 18th and 19th centuries when intergenerational mobility coefficients dropped to below 0.4. To a large extent we are able to see these changes only because the horizon of this study is long enough so that such changes can be observed.

Since public schooling was largely absent and taxation as well as redistribution comparatively light, a model in the Becker and Tomes (1979, 1986) tradition where mobility and income distribution are determined by parents' investments in their children is a natural choice for interpreting the evidence. The historical evidence of civil service degrees and official positions available points to a decline over time in the earnings elasticity of human capital that can in turn explain the observed increase in mobility. The argument is strengthened by the finding that cross-sectional inequality and persistence are correlated, as predicted by the model.

I show that parent education was complementary to nonhuman investment into children, with an educated parent resulting in an income boost of more than 50% for the child. The importance of parental human capital conditional on nonhuman child investments explains why cross-sectional inequality in education tracks more closely intergenerational mobility than income inequality does. Finally, the determinants of mobility operate in part at the group level, as my finding that between-clan inequality is highly predictive of social mobility underlines.

It is worth emphasizing that the drastic change in intergenerational mobility that I document materializes in essentially the same population, where there are few changes in technology, public education, taxation, and cultural norms. Migration into and out of the region remained low as well. A simple model of parental investments in the face of a changing earnings elasticity of human capital turns out to be quite successful in explaining long-run trends in social mobility and inequality. While other factors surely play a role for mobility, the findings support the notion that parents' incentives to invest into the human capital of their children were as central to intergenerational mobility in societies of the past as they are today.

As a final note, the trend towards higher mobility and lower inequality in China from the 17th to the 19th century occurs at a time in which China's economic growth falls behind that of a number of countries due

[51]See Corak (2013) and Braun and Stuhler (2018), respectively. A regression of son education on father education in my setting yields a coefficient just under 0.3, in line of what Lindahl, Palme, Sandgren Massih, and Sjögren (2015) find for their Swedish sample starting in the year 1865.

to not only rising growth in the West but also non-rising or possibly declining incomes in China. While there is no consensus yet on the reasons, an important question for future research will be to learn more about the relationship of social mobility, inequality, and economic growth.

References

[1] Aaronson, Daniel, and Bashkar Mazumder. 2008. "Intergenerational Economic Mobility in the United States, 1940 to 2000." *Journal of Human Resources* XLIII(1), pp. 139-172.

[2] Abramitzky, Ran and Leah Platt Boustan. 2017. "Immigration in American Economic History." *Journal of Economic Literature* 55(4), pp. 1311-45.

[3] Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review*, 102 (5): 1832-56.

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

[5] Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey. 2017. "How Well Do Automated Linking Methods Perform in Historical Samples? Evidence from New Ground Truth." NBER Working Paper 24019.

[6] Barone, G., and S. Mocetti. 2016. "Intergenerational mobility in the very long run: Florence 1427-2011". Banca D'Italia working paper # 1060, April.

[7] 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.

[8] Becker, Gary S., and Nigel Tomes. 1979. "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of Political Economy* Vol. 87(6), pp. 1153-89.

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

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

[11] Bjorklund, A., M. Jantti, and J. Roemer. 2012. "Equality of Opportunity and the Distribution of Long-Run Income in Sweden." *Social Choice and Welfare* 39: 675-696.

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

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

[14] Bourguignon, François. 1979. "Decomposable Income Inequality Measures." *Econometrica,* 47(4), pp. 901–20.

[15] Bourguignon, François, and Christian Morrisson. 2002. "Inequality Among World Citizens: 1820-1992." *American Economic Review*, 92 (4): 727-744.

[16] Brandt, Loren, Debin Ma, and Thomas G. Rawski. 2014. "From Divergence to Convergence: Reevaluating the History behind China's Economic Boom." *Journal of Economic Literature*, 52 (1): 45-123

[17] Braun, Sebastian, and Jan Stuhler. 2018. "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany." *Economic Journal.*

[18] Campbell, Cameron, and James Lee. 2003. "Social Mobility from a Kinship Perspective: Rural Liaoning, 1789–1909", *International Review of Social History* 48, pp. 1–26.

[19] Cao, Guangling, 2016. "Ming Dai Anqing Fu Keju Rencai Yanjiu" (Research on Civil Service Examination Candidates from Anqing Fu in the Ming Dai). Ph.D. Dissertation, Anhui Normal University.

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

[21] Card, D., J. DiNardo, and E. Estes. 2000. "The More Things Change: Immigrants and the Children of Immigrants in the 1940s, the 1970s, and the 1990s. In *Issues in the Economics of Immigration*, edited by G. Borjas, University of Chicago Press for the NBER.

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

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

[24] Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, Jimmy Narang. 2017. "The fading American dream: Trends in absolute income mobility since 1940." *Science*, 356 (6336): 398-406.

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

[26] Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, Nicholas Turner. 2014. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review*, vol. 104(5), pages 141-47, May.

[27] Clark, Greg. 2014. *The Son Also Rises: Surnames and the History of Social Mobility.* Princeton University Press.

[28] Clark, Greg, and Neil Cummins. 2015. Intergenerational Wealth Mobility in England, 1858-2012: Surnames and Social Mobility." *Economic Journal.*

[29] Collins, William, and Marianne Wannamaker. 2015. "Up from Slavery", paper presented at University of Colorado, Boulder.

[30] Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives*, 27 (3): 79-102.

[31] Cunha, Flavio, and James Heckman (2007), "The technology of skill formation", American Economic Review 97(2), 31–47.

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

[33] Daqing luli (The Great Qing Code), 1740 edition.

[34] Daqing huidian shili (1899 edition) Ch. 370.

[35] De Moor, Tine and Jan Luiten Van Zanden. 2010. "Girl Power: the European marriage pattern and labour markets in the North Sea region in the late medieval and early modern period." *The Economic History Review*, 63(1).

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

[37] Durlauf, Steven., and Irina Shaorshadze (2014), "Intergenerational Mobility", mimeo, University of Wisconsin.

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

[39] Elman, Benjamin. 2000. *A Cultural History of Civil Examinations in Late Imperial China.* Univ. of California Press.

[40] 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).

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

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

[43] 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.

[44] Ferrie, J. P. 2005. "The End of American Exceptionalism? Mobility in the United States since 1850.".*Journal of Economic Perspectives* Vol.19(3): 199-215.

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

[46] Genicot, Garance and Debraj Ray, 2017. "Aspirations and Inequality." *Econometrica*, 85(2): 489-519.

[47] Greif, Avner, and Guido Tabellini, 2017. "The clan and the corporation: Sustaining cooperation in China and Europe." *Journal of Comparative Economics* 2017, 1–35.

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

[49] Hajnal, John. 1965. "European marriage patterns in perspective." In *Population in History*, ed. D.V. Glass and D. E. C. Eversley. London.

[50] Han, S., & Mulligan, C. 2001. "Human Capital, Heterogeneity and Estimated Degrees of Intergenerational Mobility". *The Economic Journal*, 111(470), 207-243.

[51] Hao, Yu, and Gregory Clark (2012), "Social Mobility in China, 1645-2012: A Surname Study", mimeo, UC-Davis.

[52] Heckman, James, and Stefano Mosso. 2014. "The Economics of Human Development and Social Mobility". Working Paper, University of Chicago, May.

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

[54] Hoen, M., S. Markussen, and K. Roed. 2018. "Immigration and Social Mobility". IZA Discussion Paper No. 11904, October.

[55] Keller, Wolfgang, Ben Li, and Carol H. Shiue. 2011. "China's Foreign Trade: Perspectives from the past 150 Years." *The World Economy*, 34(6).

[56] Krueger, A. B. 2012. "The Rise and Consequences of Inequality in the United States", speech at Center for American Progress, January 12. https://cdn.americanprogress.org/wp-content/uploads/events/2012/01/pdf/krueger.pdf

[57] Lee, C.I., and G. Solon. 2009. "Trends in Intergenerational Income Mobility." *Review of Economics and Statistics.*

[58] Li, Zhou Wang. Guochao like timing beilu chuji . 1746 edition.

[59] Lindahl, M., M. Palme, S. Sandgren Massih, and A. Sjögren. 2015. "Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations". Journal of Human Resources, Volume 50, Number 1, pp. 1-33

[60] Lindahl, M., M. Palme, S. Sandgren Massih, and A. Sjögren. 2014. “A Test of the Becker-Tomes Model of Human Capital Transmission Using Microdata on Four Generations”. Journal of Human Capital, Vol. 8, No. 1, pp. 80-96.

[61] Liu Hui-Chen Wang. 1959. *The Traditional Chinese Clan Rules.* New York: J.J. Augustin Incorporated Publisher.

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

[63] 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.

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

[65] Mare, Robert D., and Xi Song. 2014. “Social Mobility in Multiple Generations”, California Center of Population Research working paper, UCLA, September.

[66] Mogstad, Magne. 2017. “The Human Capital Approach to Intergenerational Mobility”. *Journal of Political Economy* Vol. 125(6), pp.1862-1868.

[67] Pan, Guangdan. 1929. “Zhongguo jiapu xue lueshi” (A historical outline of Chinese genealogical studies), Dongfang Zazhi (The Eastern Miscellany). 26(1): 107-120.

[68] Piketty, T. 2000. “Theories of Persistent Inequality and Intergenerational Mobility.” *Handbook of Income Distribution*, eds. A.B. Atkinson and F. Bourguignon.

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

[70] Pomeranz, Kenneth. 2011. “Ten Years After: Responses and Reconsiderations.” *Historically Speaking*, 12(4).

[71] Schultz, T. W. 1961. “Investment in Human Capital.” *American Economic Review* Vol. 51(1), pp. 1-17.

[72] Shiue, Carol H. 2002. “Transport Costs and the Geography of Arbitrage in Eighteenth Century China.” *American Economic Review*, 92(5): 1406-1419.

[73] Shiue, Carol H. 2004. “Local Granaries and Central Government Disaster Relief: Moral Hazard and Intergovernmental Finance in Eighteenth and Nineteenth Century China.” *Journal of Economic History*, 64(1): 101-125.

[74] Shiue, Carol H. 2005. “The Political Economy of Famine Relief in China, 1740-1820.” *Journal of Interdisciplinary History*, 36(1): 33-55.

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

[76] Shiue, Carol H. 2017. "Human Capital and Fertility in Chinese Clans Before Modern Growth." *Journal of Economic Growth* Vol. 22(4), pages 351–396.

[77] Shiue, Carol H., and W. Keller. 2007. "Markets in China and Europe on the Eve of the Industrialization." *American Economic Review*, 97(4):1189-1216.

[78] Shorrocks, A.F. 1978. "The Measurement of Mobility." *Econometrica*, 46(5): 1013-1024.

[79] Shorrocks, A.F. 1980. "The Class of Additively Decomposable Inequality Measures." Econometrica, 48(3): 613-625.

[80] Sng, Tuan-Hwee. 2014. "Size and dynastic decline: The principal-agent problem in late imperial China, 1700–1850", *Explorations in Economic History* Vol. 54, pp. 107-127.

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

[82] Solon, G., 2004. A model of intergenerational mobility variation over time and place. In: Corak, M. (Ed.), *Generational Income Mobility in North America and Europe.* Cambridge University Press, Cambridge, pp. 38–47 (Chapter 2).

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

[84] Solon, Gary. 2018. "What do we know so far about multigenerational mobility?" *Economic Journal* Vol. 128, pages F340-F352.

[85] Telford, Ted A., et al. 1983. An Annotated Bibliography of Chinese Genealogies on Microfilm at the Genealogical Society of Utah. Taipei: Cheng Wen.

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

[87] Telford, Ted A. 1992. "Covariates of Men's Age at First Marriage: The Historical Demography of Chinese Lineages." *Population Studies*, 46(1): 19-35.

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

[89] Theil, H. (1967). *Economics and Information Theory.* Amsterdam: North Holland

[90] Townsend, R. 1994. "Risk and Insurance in Village India." *Econometrica*, 62(3), 539-591.

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

[92] Watson, James. 1982. "Chinese kinship reconsidered: Anthropological perspectives on historical research." *China Quarterly* 92:589-622.

[93] Yang, C.K. 1959a. *A Chinese village in early communist transition*. Cambridge, MA: Technology Press, MIT.

[94] Yang, C.K. 1959b. *The Chinese family in the communist revolution*. Cambridge, MA: Technology Press, MIT.

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

Appendix - For Online Publication

A Data

B Social Mobility, 1300 - 1900: Supplementary Analyses

C Intergenerational Mobility over Time: Supplemental Analyses

D Explaining Temporal Variation in Mobility: Supplemental Analyses

E Extension: Inequality and Intergenerational Mobility: Multi-variate Results

A Data

A.1 Chinese Genealogies as Data Source

This section summarizes briefly Chinese genealogies as a source of information for research in Economics. A broader discussion is provided in Shiue (2016).

The tradition of compiling genealogies began around the time of the Song Dynasty (960-1279) (Zhao 1997).[52] While there are different forms of genealogies they all share some general principles, which can be seen by comparing genealogies with census data. Census data records information at a certain point in time, and in the ideal case, the information in the census gives a complete representation of all strata of the population. To study intergenerational mobility with census information, an individual must be observed in one census as a child living with his or her parents, and then in a second census when the child is old enough to have his or her own income level. Thus, records from different cross-sectional observations have to be linked, which can be difficult and sample attrition is a critical concern (Bailey, Cole, Henderson, and Massey 2017). Furthermore, a common issue is that the child's income is observed earlier in his or her life than the parent's income (Solon 1999). Chinese genealogies, in contrast, are conceptually similar to pedigree charts. Identifying the same individual in different cross-sections is unnecessary, as the data immediately captures intergenerational outcomes. Furthermore, genealogical biographies are summaries that record the highest position and other accomplishments for each individual, which yields comparable information on permanent income for the younger and the older generation.

Thus, while genealogical data has a number of advantages, the retrospective and non-administrative nature of the compilation may affect inferences. The accuracy and representativeness of the data are among the major concerns, and I will examine these issues below. At the same time, to date research on social mobility with survey or census administrative data –going back to the middle of the 19th century– is typically based on samples comparable in size to the present study. Analyses based on large, highly reliable individual-level datasets are the exception and limited to modern times.[53]

A.2 Income Categories

The analysis in the text employs six income groups. This section introduces the main reasons that classification. I begin by describing the most disaggregated information in the data, which distinguishes twenty-three groups before turning to the aggregation into six groups employed in the text. This classification has the advantage that the six groups have distinct income levels that can be clearly ranked. The categories draw on work by Chang (1955, 1962), Eberhard (1962), Ho (1967), and Telford (1986, 1992).

[52]Originally the genealogy was a valuable private record, stored for safekeeping in the hometown of the family — in the home of an elder or in the ancestral halls — they were never meant for public exhibition. Typically, genealogies start with the progenitor of the clan from which all following clan members descend. The members were related by birth or by marriage, where not only women, but men sometimes married into a clan and adopted the surname of their spouse's family.

[53]Selection and measurement error may matter. For example, recent research for the US using comprehensive tax records finds an intergenerational regression estimate of 0.34, versus an estimate of 0.47 using other data (Chetty, Hendren, Kline, and Saez 2014 and Corak 2013, respectively).

Table A.1 shows the baseline classification into six groups in column (1) and the more disaggregated classification in column (2). Columns (3) and (4) provide information on the groups' importance in the sample, while column (5) gives a short description of the individuals' sources of income.

The data entry includes information on the major stages of each man's life, evidence of high income and wealth, various aspects of elevated status, certain functions, and specific actions such as large donations or setting up ancestral estates. There is also information on a person's education and whether he was a government official (and if so at which rank). The entry also lists the man's wife (or wives) as well as each couples' children. If there is nothing other than vital statistics in the individual's biography, and this person had no titles, degrees, or evidence of wealth, then he is coded with income group 0 (see Table A.1).

Table A.1: Income Groups with High Level of Disaggregation

(1)	(2)	(3)	(4)	(5)	(6)
Income	Income-23	N	% of Sample	Description	Educated
0	0	6,320	71.08	No title, degree, and evidence of wealth	0
1	1	35	0.39	Honorary or posthumous title; village head; other honors	0
1	2	741	8.33	Multiple wives in consecutive marriage (two or more not living at the same time)	0
1	3	824	9.27	Evidence of moderate wealth of 1st degree family, incl. minor and expectant	0
				official, lower level degree (*sheng-yuan*, *jian-sheng*), and official student	
1	4	20	0.22	Wealthy family member 2nd degree, incl. offical, *ju-ren*, *gong-sheng*, and *jin-shi*	0
1	5	31	0.35	Wealthy family member 1st degree, incl. official, *ju-ren*, *gong-sheng*, and *jin-shi*	0
1	6	145	1.63	Educated, scholar, no degrees or office; editor of genealogy;	1
				refused office, or prepared but did not pass exam	
2	7	79	0.89	Concubinage (two or more wives or concubines at the same time)	0
2	8	11	0.12	Substantial evidence of wealth and property; set up ancestral estates,	0
				large donations, philantrophy; wealthy farmer, landowner, or merchant	
2	9	163	1.83	Official student	1
2	10	1	0.01	Military *sheng-yuan*, minor military office	0
3	11	133	1.50	Purchased *jian-sheng* and/or purchased office	0
3	12	93	1.05	Student of the Imperial Academy (non-purchased)	1
3	13	48	0.54	Civil *sheng-yuan*; minor civil office	1
3	14	95	1.07	Expectant official; no degrees	0
3	15	4	0.04	Expectant official one of the lower degrees	1
3	16	23	0.26	Military *ju-ren*, *jin-shi*; major military office	1
4	17	38	0.43	Civil official with no degree, minor degree, or purchased degree	0
4	18	23	0.26	*ju-ren*, *gong-sheng*, with no office	1
4	19	47	0.53	*ju-ren*, *gong-sheng*; with expectant office	1
4	20	0	0.00	*jin-shi*, no office	1
5	21	11	0.21	*jin-shi* with official provincial post or expectant official	1
5	22	7	0.08	*jin-shi* with top-level position in Imperial bureaucracy	1
				(Hanlin Academy, Grand Secretariat, Five Boards, Prime Minister)	1

Notes: Table gives information on a man's permanent income. Sample shown is all married sons linked over three generations (son, father, and grandfather; N = 8,892). Income groups based on Telford (1986, 1992), Chang (1955, 1962), Ho (1967), and Eberhard (1962).

Notice that 71% of the men belong to the lowest income group. This indicates that this data is not primarily a record of the lives of the elites. Furthermore, the classification reveals that higher income in society was correlated with passing the tournament-style civil service exam and obtaining an office in the government. There were a number of different degrees. The *sheng-yuan* degree was the lowest of the recognized categories of government education, conferred upon those who had passed the local (district or prefectural) degree threshold. The *sheng-yuan* who were scholastically more competent were awarded with the *gong-sheng*, "imperial student" title; above them in rank were the *ju-ren* (graduate of the provincial examinations), and above the *ju-ren* were the *jin-shi* (graduate of the national metropolitan examinations). The levels are building up on each other, that is, in order to have the *jin-shi* degree one must have the *ju-ren* and the *sheng-yuan*, and in order to be *ju-ren* one must have passed the *sheng-yuan* examination.[54] Not all *sheng-yuan* advanced to the next levels, and those who didn't may have given up and turned instead to working for officials in a secretarial capacity, or, helping to manage local affairs—settling disputes, organizing local public goods projects, improving welfare and security interests, or providing education in their community (Chang 1962).

What was the role of higher income not related to government office? Even though there were rich landowners and merchants in Tongcheng, as elsewhere in China, they would often seek to acquire official government positions due to the additional income it would generate. The fraction of wealthy landowners, farmers, or merchants in my sample is only 0.12 percent (Table A.1, column (4)); this is the group that did not also held official degrees or government positions. For example, several of the Cohong merchants engaged in the lucrative bi-lateral monopoly trading relationship with Western countries in Canton (Guangzhou) in the early 1800s sought to rise to the highest government offices requiring *jin-shi* degrees, without success, so they stayed at *ju-ren* level positions (Chang 1955).

Income Individual-level income is virtually never available in historical studies of mobility. Nevertheless we do know how rank in society translates into income differences. This section provides a brief synthesis (see Chang 1955, 1962, and Ho 1967 for additional discussions). First, the income of a man is determined by two factors, the level of civil service exam he passed and the type of government office he obtained (if any). These factors were the most important determinants of where a household stood in the societal income distribution. The paper matches the degrees earned and government positions to potential earnings based on historical evidence. As noted in Table A.1 above, there were local, provincial, and national (or metropolitan) exam levels, with corresponding degrees called *sheng-yuan*, *ju-ren*, and *jin-shi*, respectively.

Having passed a certain level of degree would make a person eligible for a certain level of official position. For example, there were nine levels of civil positions during the late 19th century (Chang 1962, Table 1). A district magistrate would be seventh-ranked civil official, while a provincial governor would be a second-level civil official. The mapping between degree and official position is not deterministic, however,

[54]There were no age requirements or limitations for advancement, but since the examinations required a high level of literacy and years of study, the earliest that one could attain the *jin-shi* degree would be in the early twenties, and it was not unheard of for a man in his fifties to still be a *sheng-yuan*.

the level of office is increasing in the degree that a man has obtained. Becoming a top-level official in the imperial bureaucracy with only a *sheng-yuan* (local) degree is almost impossible, and conversely, most *jin-shi* have better-paid positions than being a district magistrate. The level of degree is useful to study mobility because there is a relatively small number of them, and they are consistently mentioned in the data.

As Tables A.1 and Table 1 show, my income groups capture the relationship between civil service examination degree and official position. My ranking is also consistent with the fact that military offices generated lower incomes than civil offices. For example, the highest civil official had a salary of 180 *taels* per year, while the highest military office came with only 82 *taels* (see Chang (1962, Tables 1 and 2). The challenge in estimating the incomes of men in the higher ranks of Chinese society includes the fact that official salary, on which there are detailed records, is only one of several income sources. Passing the examination at any level also meant income because it would make the individual eligible for providing teaching and managerial services. Teaching income has been about half of the officeholding income in the late 19th century (Chang 1962, Table 26). Officials received customary bonuses as well as in-kind support, partly for running their offices. Customary bonuses were substantial; for example, the full dress of a *hsien* magistrate might have cost 3,000 *taels*, versus the official salary of 45 *taels* a year (Chang 1962). At the same time, officials and high-income individuals more generally were expected to contribute to public goods, and the higher their income the higher was the expected contribution.

Because individual-level income is unobserved my analysis allocates all men into six distinct income groups. While this is relatively coarse it amounts to more income differentiation than other work on historical mobility in China (only two groups, high versus low; Campbell and Lee 2003, Mare and Song 2014). Arguably the best information on differences across income groups comes from assessment schedules of clans to their members who have reached extraordinary income levels. This can be thought of a tax on the clan member who has achieved a significant amount of income. There exist multiple assessment schedules for different clans and times that broadly yield the same patterns (see Chang 1962). Furthermore, there is little reason to believe that the clans' assessment schedule would not be consistent with the income generated by each of these achievements.

My analysis is based on the assessment schedule of the Chang clan, with annual assessments are given in column (5) of Table A.2.

Table A.2: Chang Clan Assessment Schedule

(1) Group	(2) %	(3) Description	(4) Source	(5) Annual Copper Cash	(6) Annual Income *Tael*	*Tael* Table 1
0	71.1	No evidence of extraordinary income or wealth	Chang (1962), p. 12		5	5
1	19.8	Educated w/o passing exam	Chang (1962), Table 26	500		7.5
2	2.9	Official student (Government student)	Clan Assessment	1,000		15
3	4.5	Imperial Student	Clan Assessment	2,000		30
4	1.2	*Ju-ren* (provincial degree)	Clan Assessment	4,000		60
5	0.2	*Jin-shi* (national degree)	Clan Assessment	8,000		120

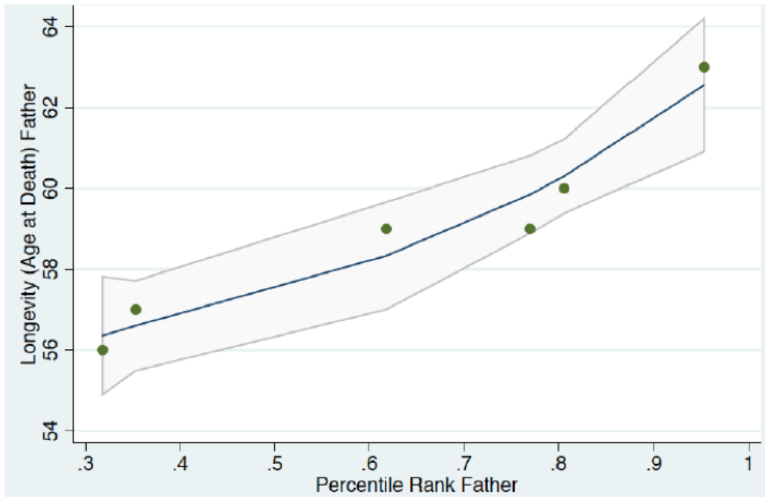
Notes: *Tael* is silver currency; clan Assessments for Chang clan, late 19th century; see Chang (1962), page. 28.

Two income groups, level 0 and 1, are not determined by the clan assessments. Income group 1 includes those individuals who are educated because they prepared for the civil service exam but they did not pass (neither did they obtain official position). The major income source of these individuals is teaching. I estimate their income to be one quarter of that of income group 3, that is, 500 copper cash per year (based on Chang 1962, Table 26). I estimate that income group 0 has 5 *taels* per year, on the low end of the range of annual income of a laborer (between 5 and 10 *taels*, Chang 1962, p.12). The income difference between groups 0 and 1 is pinned down by the average per-capita income reported by Chang (1962), which is 7.4 *taels* per year (page 328). With an annual income of 7.5 *taels* for group 1, the sample distribution (see column 2 of the table) yields an average of 7.8 *taels* per year. A 50% difference in annual income between groups 0 and 1 is plausible. The last column of the table shows annual income estimates for the six groups in my main analysis, see Table 1 in the text.

Correlation of Longevity and Income A simple check on the income groups is to examine the relationship between income and longevity. The latter can be computed from vital information on birth and death in the data source. One would expect that richer men tend to live longer. Figure A.1 shows the result of relating income and longevity across the six groups. Low-income men lived typically about 56 years, compared to men with the highest income who typically reached an age of 63 years.[55]

[55]These are statements about averages; the person with the highest lifespan in the sample, 91 years, was in the lowest income group. Also, these ages are higher than life expectancy at birth in China during my sample period. This is because my sample is based on men who survive past childhood and who succeed to marry.

Figure A.1: Permanent Income and Longevity



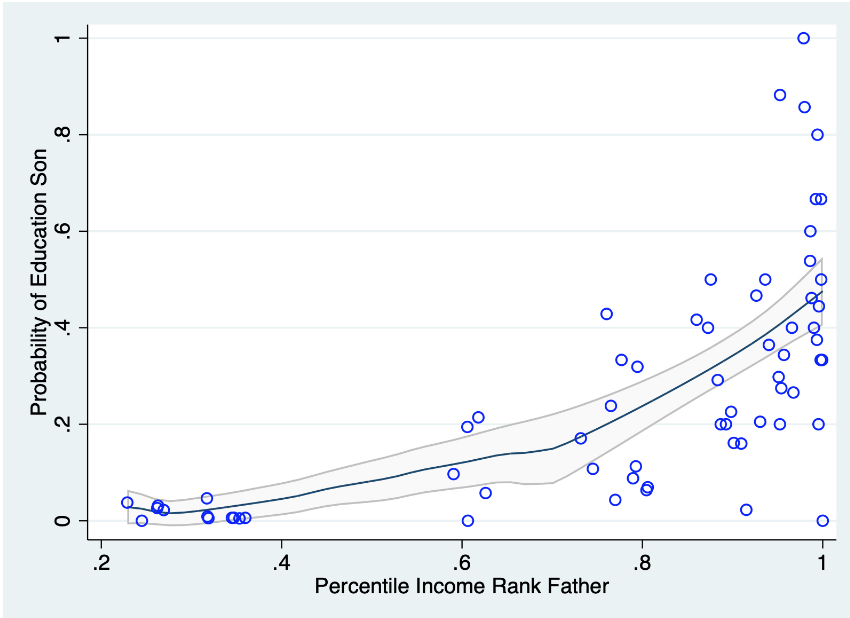
Notes: Figure shows median percentile rank and longevity for all men for whom there is information on birth and death year; 90% confidence interval shaded.

Income and Future Education A second check is to examine the relationship between income and education. Figure A.2 shows the probability that the son is educated by income group of the father. Here, education is measured with an indicator variable, shown in column (6) of Table A.1.[56] Sons from a high-income family have a substantial advantage over sons from low-income families. In particular, the probability to become educated for a son that is born to a lowest-income family is close to zero, while it is around 50% for a son of a *jin-shi*. Furthermore, the probability that the son is educated is 14% if the father is in the top two quintiles, compared to 1% when the father is in the lower half of the distribution. The relationship appears to be stronger at the high end of the income distribution.

Overall, higher father income translates into more education of the child. This provides also evidence in favor of the human capital investment equation (7).

[56]More heterogeneity in terms of education is considered in Figure A.15 below.

Figure A.2: Income and Next-Generation Education

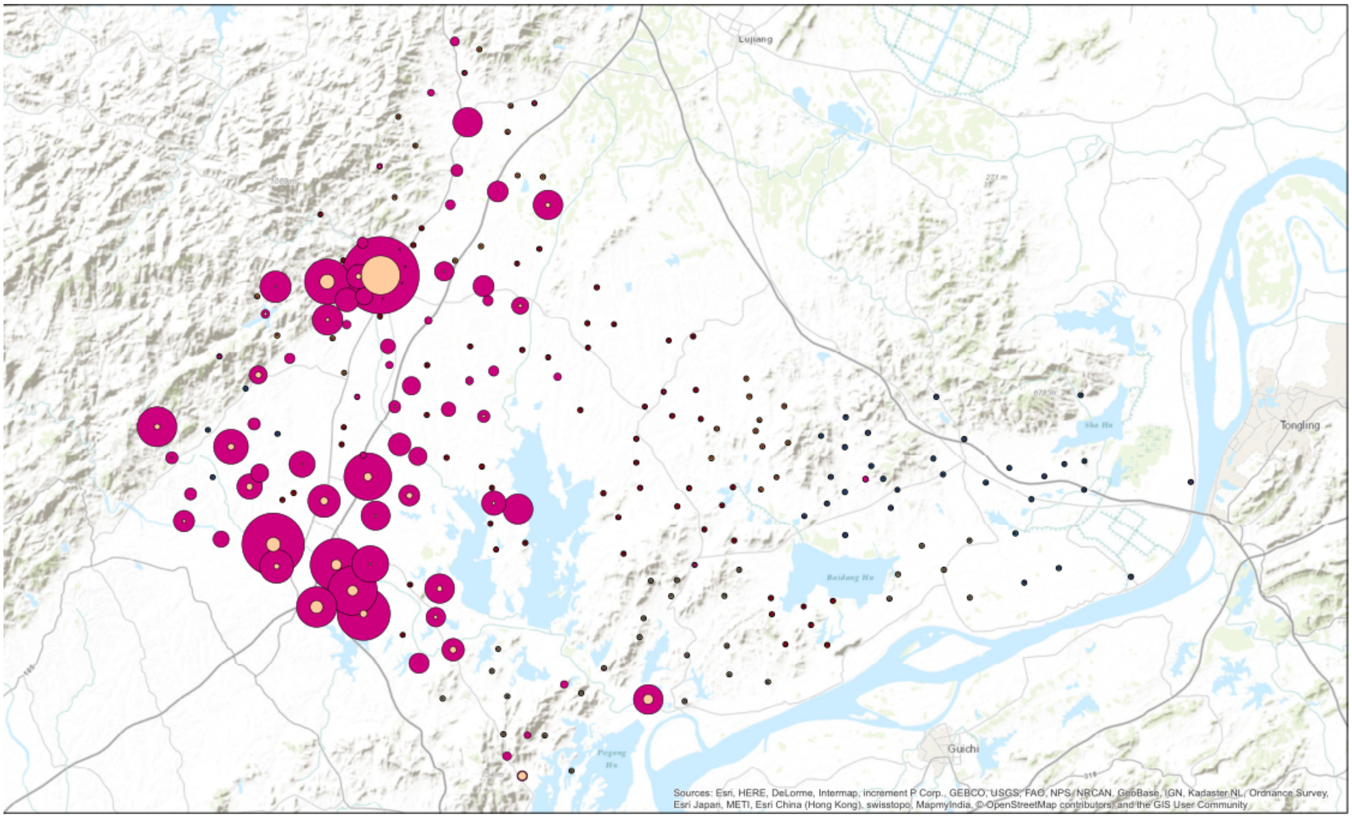


Notes: Figure shows relationship between income rank of father and son education across birth cohorts; 90% confidence interval shaded.

Residential Locations in Tongcheng County Figure A.3 shows a map of Tongcheng county, with the Yangzi River in the lower right of the figure. The location of villages and towns where the members of the seven clans resided are indicated. The lighter color shading indicates the number of men who are educated (see Table A.1, column (6)).

The capital of the county, Tongcheng city, is the largest point, located in the upper left of Figure A.3. Geographic location is an aspect of income, with richer men being more likely to live in the capital; see Table 2 in the text .

Figure A.3: Tongcheng County - Residential Locations



Notes: The frequency of observations for a village, 1300-1900, is proportional to the point's radius. Pink (darker) is the total observations while tan (lighter) reflects the number of educated.

A.3 Data Accuracy

This section examines the accuracy of the data, and whether there is evidence for certain biases, including recall bias. Additional results are presented in Shiue (2016). Table A.3 lists extended information on vital statistics by giving the men's birth and death months. The average for either is close to 6.5, which is what one might expect if the timing of births and deaths are random events and there is no artificial age heaping.

Table A.3: Summary Statistics on Father Month of Birth and Death

	N	Average	Std. Dev.
Year of Birth	8,893	1732.04	70.85
Month of Birth	8,893	6.92	3.49
Year of Death	8,658	1787.44	70.59
Month of Death	8,658	6.47	3.43

Notes: Information according to Chinese lunar dates has been converted to solar dates following Shiue (2002).

The result is not very surprising because dates in genealogies are in terms of traditional calendrical symbols, not in numerical format, and traditional calendrical symbols are difficult to falsify (Zhao 1997). This increases the reliability of the recorded dates, which are key to forming birth cohorts as well as calculating each man's longevity (see above).

I have also directly compared the list of people who are recorded to have obtained the *jin-shi* degree against other known lists of degree holders from Tongcheng County (Fang 2010; Cao 2016; Wang 2017). There were over 51,000 *jin-shi* degree holders from the Yuan, Ming, and Qing. Information on top degree holders can be cross-checked for accuracy by referring to known lists of *jin-shi* degree holders from the Chinese state, which give the name, the date on which someone received his degree, and his hometown. I have verified that the information on the 18 *jin-shi* in my Tongcheng sample is consistent with the information of these official lists.

Differences across Clans Table 2 in the text shows that the seven clans in my sample differ along several dimensions, including their average income levels. Variation across clans is useful for asking whether a particular person is more likely to be included in the genealogy than another. The following examines various forms of sample selection using differences across clans. While there were a number of different reasons why the genealogical tradition emerged, one relevant concern here is that genealogies begin with a particularly noteworthy man, who then becomes the progenitor of the clan. Part of his noteworthiness might come from a high level of education, which is one of the most important signs of income and one of the most consistently reported characteristics of noteworthy persons. Alternatively, perhaps later generations were more likely to select a noteworthy progenitor. In either case, the implication would be a strong trend of declining income over time.

In the Tongcheng genealogical sample, there are three clans whose records begin with an educated progenitor: the Chen (progenitor born in 1298), the Wang (1358), and the Ma (1408). However, the status of these three progenitors was not more than what might be considered an intermediate level, not the highest level (*jin-shi*). For the other four Tongcheng clans, the highest levels of income are typically found nine generations after the inception of the dynasty. Thus, the income patterns in the data do not simply reflect particularly noteworthy individuals that started the clan records as progenitors. Figure A.4 shows the size of the seven clans in three different sub-periods. We see that all clans are present in all sub-periods.

Figure A.4: Clan Size Over Time

Notes: Clan names are as follows: 1 is Chen, 2 is Ma, 3 is Wang, 4 is Ye, 5 is Yin, 6 is Zhao, and 7 is Zhou.

Selection would also arise if genealogical records contain more entries of success compared to failure. If a genealogy were more apt to record success, we might expect that it would do this on multiple measures. One way to check for this would be to see if clan size and average income are correlated. The correlation between clan size and average income group (column (1) in Table 1) is virtually zero. Thus, there is no evidence that on average, richer clans have included more entries in their genealogies. We can also check if this kind of effect might be present over time to see if it is true that periods during which a clan is successful are also those when relatively many clan members are recorded. Breaking down the overall sample period of 1300 to 1900 into twelve birth cohorts, as in the text, yields a correlation between income and the number of clan members of close to zero (and negative, -0.10).

Furthermore, because the updating of the genealogy was retrospective one might believe that it is in times right after a clan had been relatively successful when a relatively high number of clan members would be recorded. This might happen if the appearance of a high-income individual correlates with a better memory of all the family members who were related to this locally famous individual, compared to periods when no one in the clan was particularly successful or famous. However, there is no positive relationship between the number of clan members recorded and the average clan income (the correlation is insignificant at -0.09).

Additionally, it is possible to examine whether high-achieving clans tend to be overrepresented towards the end of the sample, as would be the case if there was survivor bias. If so, one would expect that clans with high income account for a relatively large share of the post-1800 observations in the Tongcheng data. Across clans, however, I do not find a strong relationship between average income and the share of post-1800 observations (insignificant correlation of 0.07).

Overall, these analyses provide evidence that the data is of high quality and that recall and other biases do not seem to play a major role.

A.4 External Validity

Another question is whether the Tongcheng sample can be considered as nationally representative for China. A first way to assess this is to consider mortality rates by age group. Population figures at the regional level are typically based on gazetteers, which are local histories about a certain place.[57] In addition, there are official accounts for subsets of the population, such as the Qing population registers, which are the product of the Eight Banner registration system.[58] Telford (1990) compares demographic patterns in the Tongcheng genealogical data and the 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).[59]

Another way to gauge the representativeness of the Tongcheng sample is to look at the number of top income individuals as a percentage of the population. Much of the available estimates focuses on the upper groups as a percent of the population. While there is no consensus on who should be considered to belong to the upper income groups there is wide agreement that education and success in the civil service examinations were important. Chang (1955) takes the view that *sheng-yuan* holders and above were in the upper class, and estimates that they were in the top 2% of the total population in the later half Qing period. In my analysis, the part of the population corresponding to Chang's (1955) definition account for 3.3% of the sample, which is comparable.[60] Fei (1946) presents another, wider estimate of the upper income groups, which he placed at 20%. In my analysis, groups 2 to 22 in Table A.1 correspond to Fei's definition of high income—and the share of these groups in my sample is 20.2%. Both these comparisons suggest that the Tongcheng genealogies are fairly representative of China's population as a whole with respect to the size of top income groups as well as the relative size of higher versus commoner groups.

Given the genealogy is a written document, if literate individuals only recorded information about themselves their and immediate kin, the percentage of top income people in the genealogy should be very high. Alternatively, if genealogies recorded extended family who were not of high income—rules of ritual say that all adult male members are eligible, regardless of education or income—the percentage should typically be low. How does the share of top income groups in my sample compare with other evidence? In his classic study based on national lists of *jin-shi*, which are extremely reliable, Ho (1967) reports that during the

[57] Three county-level gazetteers about Tongcheng cover the period under analysis: *Tongcheng xian zhi* (1490), *Tongcheng xian zhi* (1696), *Tongcheng xuxiu xian zhi* (1827).

[58] These data are available for areas in China's northeast, in today's Liaoning and Heilongjiang Provinces, these lands were organized under the Imperial Household Agency and the Jilin Military Yamen, an office in the General Office of the Eight Banner Command. See https://www.icpsr.umich.edu/icpsrweb/ICPSR/series/265. For the imperial household dynasty, there are observations going back to the seventeenth century (Lee et al. 1993).

[59] Campbell and Lee (2002) compare data from genealogies of Liaoning to the household registers. They find evidence of higher mortality rates in the genealogies compared to what was reported in the registers. This contradicts the idea that mortality in underreported in the genealogical sample. It also not what one would expect if more privileged and educated men would be more likely included in the genealogies.

[60] In Table A.1, column (2), they are groups 13 and above.

Qing in Anhui there were 41 *jin-shi* 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 *jin-shi* per capita in Qing China (Ho 1967, p. 228). In my sample, there were a total of 14 *jin-shi* during the Qing, which comes to 0.045 percent of the population in the data.[61] Thus, there are about ten times more *jin-shi* in the Tongcheng sample than in Qing Anhui overall. The reason for this is that Tongcheng was an important urban center in Anhui that had a well-known reputation for producing high-achieving individuals during the Qing (Beattie 1979).

At the same time, *jin-shi* were rare, and some parts of Anhui province did not produce a single *jin-shi* over centuries. Furthermore, Tongcheng was not among the areas of China where top individuals were most prevalent. Some areas had a number of *jin-shi* that was higher by an order of magnitude compared to Tongcheng.[62] Therefore, while the number of men in the highest income group in Tongcheng was higher than in the local surrounding area, Tongcheng was noteworthy at a local, perhaps provincial level, but it was not an unusual region in China.

Moreover, the variation in *jin-shi* across clans in the Tongcheng sample dwarfs differences in the population. At the top of the list, the Ma clan had 9 *jin-shi* relative to 627 men, a ratio of 1.4%, whereas other Tongcheng clans do not have a single *jin-shi*. Put simply, the potential sample selection in the genealogy, as a genre, is likely to be minor in comparison to the rather large and pronounced differences that we see in the achievements of different clans.

This analysis of the representativeness of top income in the sample can be supplemented by examining the representation of other income levels. There may be a tendency to exaggerate the income of people in the genealogy, or to drop the poorest segments of the clans—in both cases there would be many more officials or educated people in the genealogical sample than in the society at large. While there exists no generally agreed-upon income classification for China, another comparison to other sources is that in the Liaoning sample during the Qing dynasty starting in the mid-18th century, 98% of males had "No Status" while 2% were "Officials".[63] This compares to about 71% of men with the lowest income level in the Tongcheng sample, while about 1.4% have an official position. The relatively high fraction of "Officials" in the Liaoning data might be related to the fact that it was a less densely populated area in the North of China.

Apart from income groups, I have checked whether the percentage of successful examinees in the Tongcheng genealogies broadly lines up with national averages. The most systematic evidence on education in China

[61]There are 8,291 married men during the Qing in the sample. Telford (1986) finds that the proportion of unmarried men in Tongcheng during a somewhat earlier period of the Ming was above 20%. I assume that 20% of all men did not marry , and that the Qing population was composed of below-age-of-marry/men/women to one-third each. This gives a scaling factor of 3.75: 14 *jin-shi*/(8,291 x 3.75) = 0.045 percent. If there are 20% of men not marrying, and there is universal marriage of women, then there must be 20% fewer daughters than sons, and 20% fewer women than men.

[62]Zhejiang and Jiangsu were among the provinces with high densities of *jin-shi*. Ho (1967) reports that a single prefecture in China could have as many as 1,004 *jin-shi* during all years of the the Qing. With typically seven counties to a prefecture, this means that there could be as many as 1,004/7 = 143 *jin-shi* per county during the Qing. Compared to this, the 14 *jin-shi* in my sample are not exceptional.

[63]Source: Author's computations from the China Multigenerational Dataset, Liaoning 1749-1909, http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/27063.

during Ming-Qing is related to the civil service examinations. In particular, the number of *sheng-yuan*, the individuals that passed the local examination, was about 500,000 in the year 1700 (Elman 2000), or roughly 0.3% of the population. In the Tongcheng sample, about 0.76% of the men around the year 1700 were *sheng-yuan*. Accounting for women, children, and elderly indicates that the fraction of *sheng-yuan* in Tongcheng was similar, or perhaps somewhat lower than in China as a whole.

In summary, overall comparisons of degree holders, upper income groups, and mortality rates suggest that the income distribution in the Tongcheng sample is not very different from what one might expect from a randomly drawn sample. Genealogies were compiled for ancestral rituals, so there would not be an obvious incentive to systematically create false entries. There is no evidence that major biases exist in this Tongcheng sample of seven clans. Systematic checks of internal consistency and external validation of the Tongcheng data suggests that while measurement error is no doubt present the records appear to be fundamentally sound. Furthermore, information in the sample is consistent with what we know and expect based on other sources for larger parts of China. To a significant extent this is because the sample is based on seven genealogies that each describe rather different local clans.

B Social Mobility, 1300 - 1900: Supplementary Analyses

B.1 Income Definition

This section examines the role of various aspects of my definition of income groups for the results. First, one difference compared to the pioneering work by Ho (1967) is that my sample includes a substantial number of men from lower income groups. To see how important the influence of top-status levels are for the results, I re-estimate equation (11) with the central 95% of the father distribution. As shown in Table A.4, this does not strongly affect the results (column (2)).

Further, one of the income groups in the data refers to those who purchased a degree as a way to get a minor official position. While as a rule official positions in China were open to men who obtained a degree by passing the official government entrance exam, people who were not able to pass the exam, but were sufficiently wealthy could sometimes buy their way in. If one assumes that a purchased degree would give the holder the same permanent income as men who passed the exam, this does not change the estimated mobility substantially (see Table A.4, column (3)). One reason for this is that purchased degrees and positions account for only 1.5% of the sample and holders of purchased degrees were generally not appointed to high offices.

Table A.4: Income Distribution: Definition and Alternative Samples

	(1) Baseline	(2) Central 95%	(3) Purchased & Exam Same	(4) Single Men	(5) Multiple Wives	(6) Family Wealth first degree
β	0.528	0.506	0.528	0.377	0.625	0.306
	(0.009)	(0.010)	(0.009)	(0.007)	(0.008)	(0.012)
$E\left[R_S(i)\,\vert R_F(i) = 0.25\right]$	0.418	0.417	0.419	0.459	0.418	0.467
$E\left[R_S(i)\,\vert R_F(i) = 0.80\right]$	0.620	0.610	0.619	0.595	0.622	0.550

Notes: Estimation of β (equation 11) by OLS; $E\left[R_S(i)\,|R_F(i) = 0.25\right]$ is the expectation of percentile rank of sons with fathers at percentile rank < 0.5. $E\left[R_S(i)\,|R_F(i) = 0.80\right]$ is the expectation of percentile rank of sons with fathers at percentile rank > 0.6. N = 8,893 in (1), (3), (5), and (6); N = 8,760 in (2) and N = 14,905 in column (5). Robust s.e. in parentheses.

In a society such as this one where illegitimacy rates were very low the married population is the most relevant for studying mobility from parents to children. I now expand the sample to include individuals who appear in the genealogy as children but not as married husbands. While marriage was nearly universal among women, the fraction of single men is generally estimated to have been substantially higher.[64] There are at least two reasons for this. First, female mortality rates were higher than those for males, and second, men tended to remarry more frequently than women did upon the death of a spouse. In a stratified society, this implied it was not uncommon for higher-income men to marry multiple times while low-income men stayed single. For the analysis of single men, I focus on men who reached at least fifteen years of age and for whom I know the date of death.[65] Because there is no information on the permanent income of unmarried men it has to be estimated. I assign these individuals to the lowest income group, which is a plausible assumption. At the same time, this approach (weakly) underestimates the income of single men and should therefore be seen as a bound on how social mobility would change if single men were included.

The relative mobility estimate β declines from 0.53 in the baseline to 0.38 (column (4)). Further, in terms of upward mobility sons of fathers in the lower half of the income distribution can expect to rise to rank 0.46, compared to 0.42 in the baseline (row 2). Also, a son of a father at the 80th percentile can expect to drop to the 60th, in contrast to 62nd percentile rank in the baseline (row 3). In part, mobility is higher including single men because father-son pairs are added to the sample in a way that tends to increase downward mobility. In fact, some single sons to whom I assign income group level zero are likely to have had a level higher than that. Further, the extent to which this was the case may have be correlated with income of the father.[66] To understand how important this might be for estimating mobility, instead of level zero I assign level 1 to a randomly selected subset of the single men, plus a boost in income proportional to the income of their fathers.[67] With this approach intergenerational mobility would be estimated at β = 0.48, not too different from the baseline estimate of 0.53. Thus, under plausible assumptions the inclusion of single men would yield a moderately higher intergenerational mobility.

[64] The proportion of unmarried men in Tongcheng during the Ming was above 20% (Telford 1986).

[65] The presence of a death date record increases the likelihood that the individual did not out-migrate. In any case, out-migration over longer distances was rare; 1.9% is the figure reported by Telford 1986).

[66] Consistent with this there is evidence that the fraction of sons that would not marry is declining with status during the Ming (Telford 1986, Table 5.1).

[67] I assume that sons born in January have income group one plus a 10% boost depending on their fathers' income group in column 2 of Table A.1.

I also explore the influence of particular elements of a person's biography on the definition of income group. One assumption used in the definition of incomes is that men who are sequentially married more than once would have attained a certain level of income.[68] While this is supported by historical accounts that a man who could afford to marry more than once would have likely had been somewhat better off, an alternative is to drop this category. If we were to instead code these men no differently from someone in the lowest income group (level 0), the relative mobility estimate β changes to 0.63, while upward and downward mobility remain largely unchanged (column (5)). This suggests that the way in which this particular group was coded has only a limited influence on the findings.

Another item that is among the biographical information is whether one of a man's immediate family member had some wealth, including from a lower level official position or a lower academic degree (*sheng-yuan*). If instead I treat these men as commoners (income group level 0), β is estimated to be lower than before, at 0.306 (more mobility), and there is also evidence for more upward and downward mobility than in the baseline (compare columns (6) and (1), respectively). However, in China during the sample period these men would have higher incomes than commoners even though this is based on family, not individual characteristics. Consistent with that, in the sample such men typically live three years longer than commoners.[69] I also show below that my findings on changes of mobility over time are robust to how I treat these men with family resources, see Table 8.

B.2 Measurement Error in Income

The preceding analysis has shown results for a number of alternative income definitions. In the following, I treat true permanent income as a variable that is observed with error. Depending on the size of the error, a person's income may be very similar, or it could also be quite different from what is given in the Tongcheng genealogies; in addition, it could be higher *or* lower. While this approach does not 'undo' whatever measurement error is present in the data, it allows one to see, both qualitatively and quantitatively, the influence of measurement error for mobility.

Table A.5: Measurement Error in Permanent Income

	(1) Small Error Father	(2) Large Error Father	(3) Same Error Father and Son	(4) Error Father and Son orthogonal	(5) Error Father and Son opposite
β	0.528	0.421	0.553	0.506	0.486
$E[R_S(i)\,\vert R_F(i) = 0.25]$	0.418	0.430	0.412	0.421	0.425
$E[R_S(i)\,\vert R_F(i) = 0.80]$	0.618	0.605	0.629	0.617	0.610

Notes: Estimation of β (equation (1)) by OLS; $E[R_S(i)\,|R_F(i) = 0.25]$ is the expectation of percentile rank of sons with fathers at percentile rank < 0.5. $E[R_S(i)\,|R_F(i) = 0.80]$ is the expectation of percentile rank of sons with fathers at percentile rank > 0.6. N = 8,893.

My starting point is a mean-zero error in father permanent income with a standard deviation of 0.01.

[68] Level 1 in Table A.1, column 1.

[69] Furthermore, the hypothesis that family relations can affect permanent income finds support in the fact that the official *jin-shi* lists of the Chinese state mention the name of the immediate kin of the *jin-shi* (including father, mother, wives, brothers, and sons).

Results are shown in Table A.5. As one would expect, the effect of this relatively small error on the mobility estimate is limited (Table A.5, column (1)). When the size of the error is drastically increased to a mean-zero error with standard deviation of 0.5, the relative mobility estimate falls from 0.53 to 0.42 (column (2)). Note that father income is the independent variable in the regression. Therefore, the change in β towards zero is in line with classical measurement error that biases the regression coefficient towards zero. In comparison, changes in my nonparametric upward and downward mobility estimates are relatively small.

The following three specifications add errors to both the father and son data. I begin with the same measurement error for father and son (column (3)), which can be thought of as perfect correlation in terms of income mismeasurement. Generally, this leads to more persistence than before. In particular, the coefficient in the rank regression is now 0.55, up from 0.53 in the baseline. Persistence is higher because a term is introduced that is the same for father and son in a given pair.

If, in contrast, the error in father income is orthogonal to the error in son income, mobility is estimated slightly higher than in the baseline analysis (compare columns (4) and (1), Table A.5). Finally, column (5) shows results when the error in father and son status is negatively related with a correlation of -1; as one might expect, the result is a higher level of mobility. For example, β in the rank regression falls from 0.53 to 0.49. Generally, positive correlation in measurement error is more plausible than negative correlation for a given father-son pair. As seen above, positively correlated measurement error will tend to yield lower mobility (column (3)). To the extent that in the data there is positively correlated measurement error in a given father-son pair, the β estimate would be biased towards persistence, as in column (3), and the true β would be lower than the estimate of 0.53 in column (1). At the same time, comparing columns (1) and (3) of Table A.5 suggests that, quantitatively, the difference is likely to be small.

B.3 Intergenerational Mobility with More Than Two Generations

While most of the intergenerational mobility research is based on two generations (typically father and son), mobility may depend on more than two generations (so-called multigenerational mobility).[70] The influence of previous generations could be due to personal contact, as may be the case between grandfather and son, or due to indirect effects, such as from accumulated resources and values in earlier generations. In this section, I provide evidence on the roles of the paternal grandfather, great-grandfather, and great-great-grandfather income for mobility. Instead of the father-son pair, the unit of observation now is the five-generation quintuplet, from great-great-grandfather to son. The following are simple extensions of the rank regression, equation (11):

$$R_S(i) = \alpha + \beta_1 R_F(i) + \beta_g R_g(i) + \epsilon(i),$$

where g, $g = \{GF, GGF, GGGF\}$ indicates grandfather, great-grandfather, and great-great-grandfather, respectively, and $R_{GF}(i)$, for example, is the percentile rank of the grandfather in quintuplet i. With

[70] See, e.g., Long and Ferrie (2018).

roughly thirty years between any two generations, a five-generation linked quintuplet covers about 150 years of calendar time. Similar to how lagging in a time series regression reduces the number of observations, the linking of five generations results in a fall in the number of observations.

Table A.6: Intergenerational Mobility with Five Generations

	(1)	(2)	(3)	(4)	(5)
Rank Father	0.528^{**}	0.533^{**}	0.484^{**}	0.488^{**}	0.498^{**}
	(0.009)	(0.009)	(0.013)	(0.011)	(0.010)
Rank Grandfather			0.072^{**}		
			(0.013)		
Rank Great-Grandfather				0.087^{**}	
				(0.011)	
Rank Great-Great-Grandfather					0.078^{**}
					(0.009)
R^2	0.311	0.322	0.326	0.329	0.329
N	8,893	7,328	7,328	7,328	7,328

Notes: Dependent variable is percentile income rank of son. Estimation by OLS. Robust standard errors in parentheses. ** means significant at a 1% level.

I begin by comparing the regression results for the father-son sample and the five-generation sample. Results in Table A.6 shows there is little difference in the mobility estimates, see columns (1) and (2). This indicates that the results are not strongly affected by whether there are five generations or not. The first set of results for multigenerational mobility are for Rank Grandfather, shown in column (3). The coefficient for Rank Father falls somewhat while the coefficient on Rank Grandfather is about $\beta_{GF} = 0.07$. One interpretation of this result is that the advantage that sons from higher income families have is not exclusively due to his father, but rather to both father and grandfather. Also, the ratio of grandfather to father coefficient is 0.15; this is somewhat lower but comparable to the relative influence of grandfathers in occupational mobility regressions for the US and UK starting in the 19th century (Long and Ferrie 2018).

Moving to additional generations, the coefficient of great-grandfather rank is positive at 9% (column (4)), while the great-great-grandfather variable enters with a coefficient of 8% (column (5)). Notice that the size of the coefficient for the earlier generation does not monotonically fall as the generational distance to the son increases. This suggests that the results do not capture the direct contact between persons.

Quantitatively, the results of column (5) imply that a 10-percentage point advantage in father income translates on average into an advantage of the son between 5.1 percentage points and 5.8 percentage points; the former is the case when the great-great-grandfather has a low income level, the latter when he has top income. This means that having a high-income great-great-grandfather increases the advantage of the privileged son by a sizable 14%. The sum of the coefficients of father plus great-grandfather, or of father plus great-great-grandfather (columns (4) and (5)) is somewhat above the estimate of β_1 when only father income is included (column (2)). This indicates that if a family has high income for multiple generations, the advantage to the son is higher than if only the father had the high income.

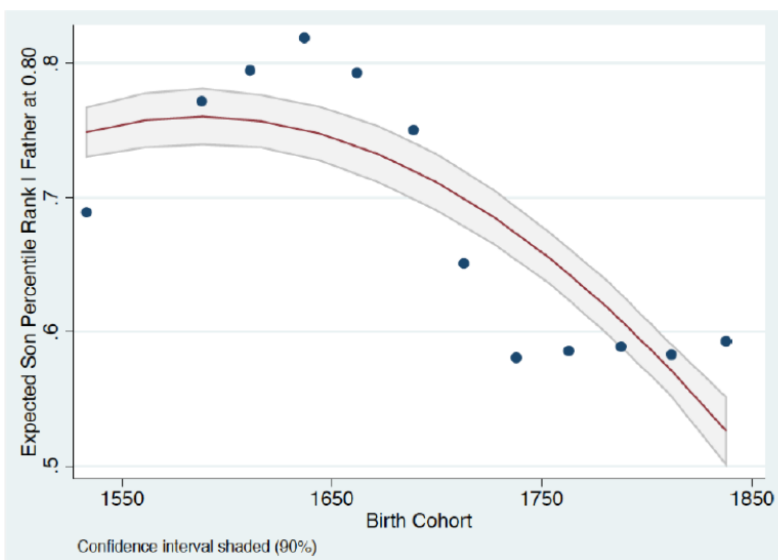
The results indicate that while generations before father enter significantly into the regression the additional mobility-relevant information coming from these earlier generations appear to be largely shared. In section 6.5 I show that entering the income of four earlier generations (father to great-great-grandfather) simultaneously into the regression yields a sum of coefficients of about 0.57, somewhat higher than when only father is included (giving a value of 0.528, see Table A.6, column (1)). This confirms other recent results on multigenerational mobility (e.g. Long and Ferrie 2018, Braun and Stuhler 2018). At the same time, the difference between two-generational and multigenerational mobility estimate does not change much over time, see Figure 11.

C Intergenerational Mobility over Time: Supplemental Analyses

The following shows complementary results on intergenerational mobility over time (section 5.2) in the text.

Downward Mobility with Fixed-Length Birth Cohorts Figure A.5 shows results on downward mobility for birth cohorts with fixed size in terms of years (not observations). Around 1600, a son whose father was at the 80th percentile could expect to be at the 75th percentile himself, whereas by the early 19th century the son could expect to be closer to the 55th percentile, a 27 percent drop in expected rank for these sons from high-status families.

Figure A.5: Downward Mobility over Time - Cohorts of Fixed Length in Years

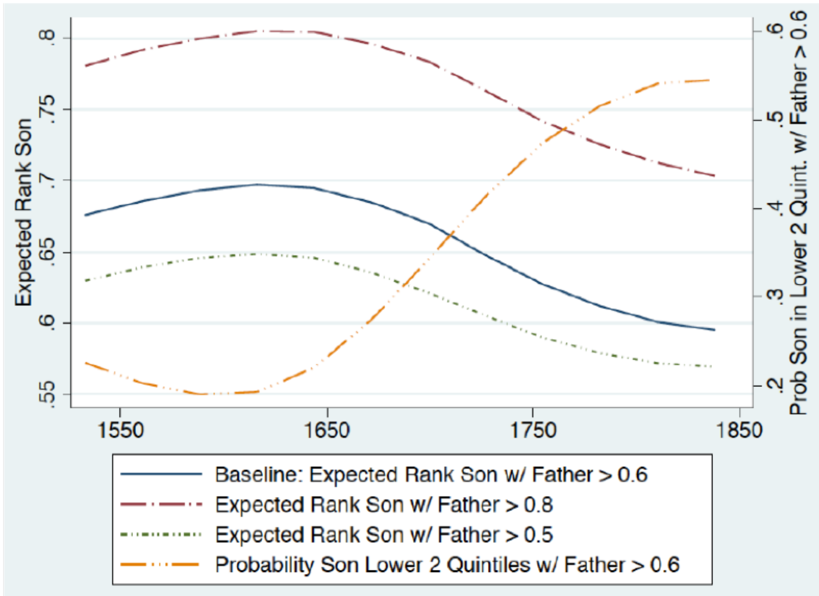


Notes: Shown is the expected percentile rank of sons with father percentile income rank > 0.6 for twelve birth cohorts. Horizontal axis gives median birth year in cohort; earliest birth year is 1330.

These results are similar to those for forty equal-sized birth cohorts, see Figure 4.

Downward Mobility with Alternative Measures Figure A.6 shows results for downward mobility for alternative measures.

Figure A.6: Downward Mobility over Time - Alternative Measures

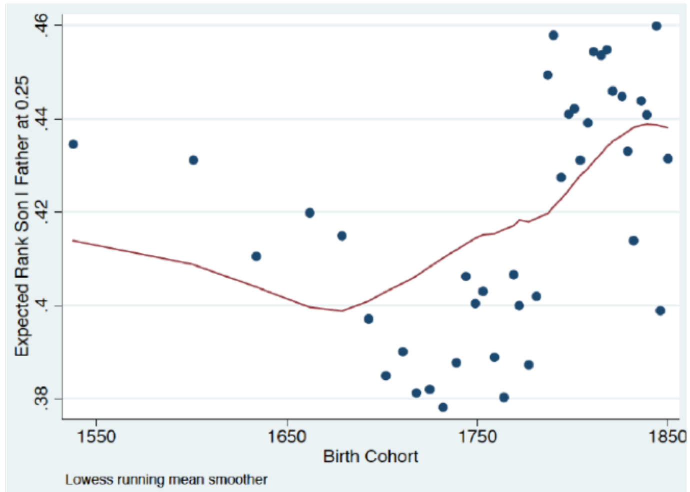


Notes: Shown is expected percentile rank of son with fathers of percentile rank greater than 0.6, greater than 0.8, and greater than 0.5, as well as the probability that a son is the lower two quintiles of the distribution given the father has percentile rank above 0.6. Locally smoothed series across birth cohorts. Horizontal axis gives median birth year in cohort; earliest birth year is 1330.

The baseline results for downward mobility are shown in Figure A.6 with a solid line. Results are qualitatively similar for sons of fathers with rank above 0.8 (dash-dot) and above 0.5 (dash-dot-dot). Figure A.6 presents also evidence on downward mobility between sets of quintiles. In particular, the probability that a son from a family in the top two quintiles (above 0.6) ends up in the lower two quintiles (below 0.4) is about 20% in 1600 but more than 50% by the beginning of the 19th century. The above results indicate that the trend towards higher downward mobility noted in the text is robust to alternative measures.

Upward Mobility Over Time with Fixed-Sized Birth Cohorts Figure A.7 presents evidence for a higher level of social mobility towards the end of the sample period in the case of birth cohorts of equal size, similar to what was shown for the case of birth cohorts with an equal number of years (see Figure 5 in the text).

Figure A.7: Upward Mobility over Time - Cohorts of Fixed Size

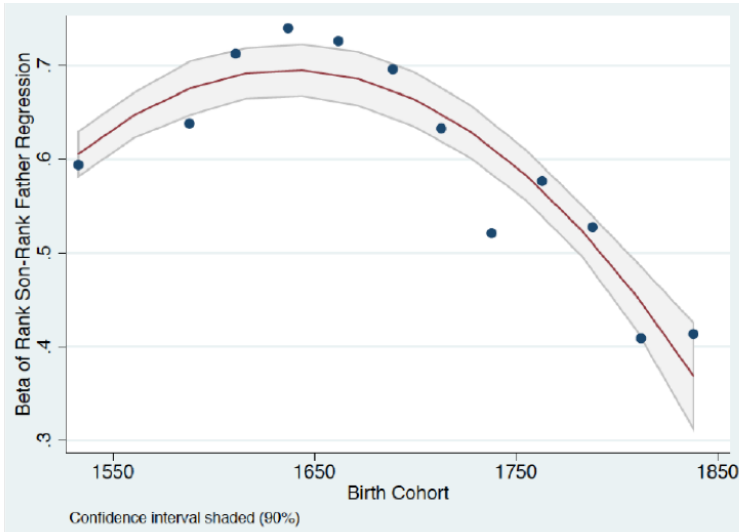


Notes: Shown is expected percentile rank of son with fathers in the lower half of rank distribution for forty birth cohorts. Horizontal axis gives median birth year in cohort; earliest birth year is 1330.

I have also confirmed that my findings on upward mobility are robust to changes in the particular measures, analogous to the results for downward mobility shown in Figure A.6.

Intergenerational Mobility with Constant Income Distribution Instead of income distributions that are specific to each cohort, I have estimated changes in mobility when each son is ranked in the distribution relative to sons over the entire sample period of six hundred years. The result, shown in Figure A.8, is quite similar to the findings with birth-cohort specific income ranks: β is estimated around 0.6 or higher until about the year 1700, before the parameter estimate β drops to around 0.4 by the early 19th century.

Figure A.8: Mobility over Time with Constant Income Distribution

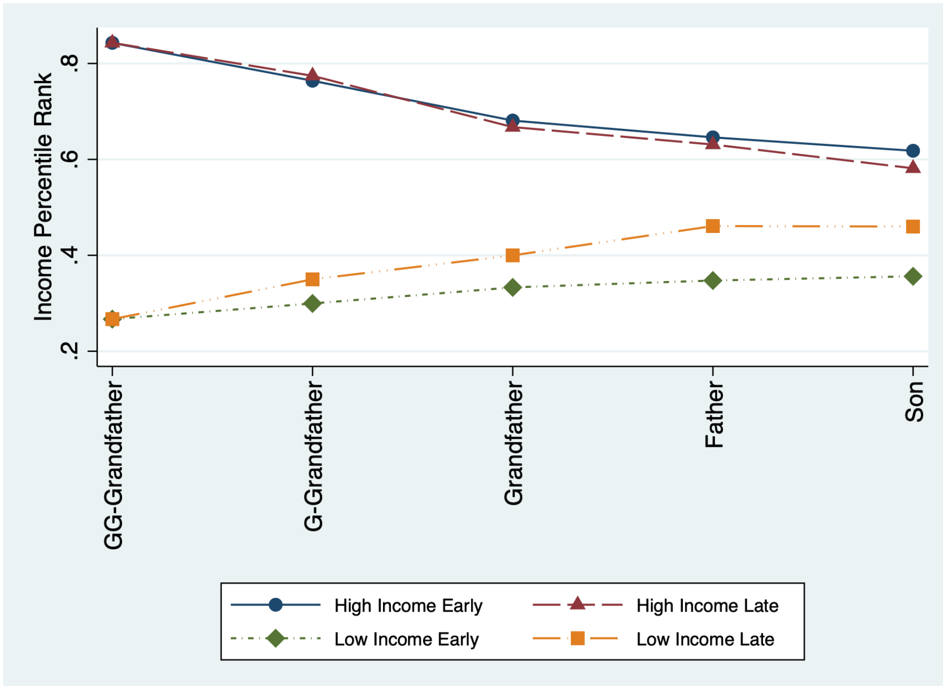


Notes: Shown is β (equation (11)) for twelve birth cohorts based on distributions of father and son income held constant for the entire sample period; the first birth year is 1330.

The results in Figure A.8 indicate that the increase in mobility over the sample period is not driven by the way in which the father and son were ranked relative to others in their specific birth cohort.

Multigenerational Mobility over Time with Five Linked Generations In Appendix section B I have analyzed multigenerational mobility for the entire sample period 1300 to 1900. Below I present evidence on changes in multigenerational mobility with five linked generations over time.

Figure A.9: Multigenerational Mobility Over Time



Notes: Figure shows average percentile income rank for consecutive generations in the early versus late period, defined as son's birth year before versus including and after 1786. For each period, figure shows income of four subsequent generations when great-great-grandfather has high income (upper two lines) and when great-great-grandfather has low income (lower two lines). GG-Grandfather: great-great-grandfather, G-Grandfather: great-grandfather.

Figure A.9 shows results for the first versus the second half of sample. The two lines on top describe downward mobility conditional on the great-great-grandfather being in the top 30% of the income distribution. In the first half of the sample, a son with such a high-income great-great-grandfather could expect to be at percentile 62 of the distribution, whereas in the second half the expected rank of the son would be 58. Thus, there is a higher level of downward mobility in the second half of the sample (based on point estimates). Moreover, with five generations it becomes clear that the increase in downward mobility—the difference between the two downward trending lines--does not materialize in the first or second generation; instead, it is more pronounced in subsequent generations.

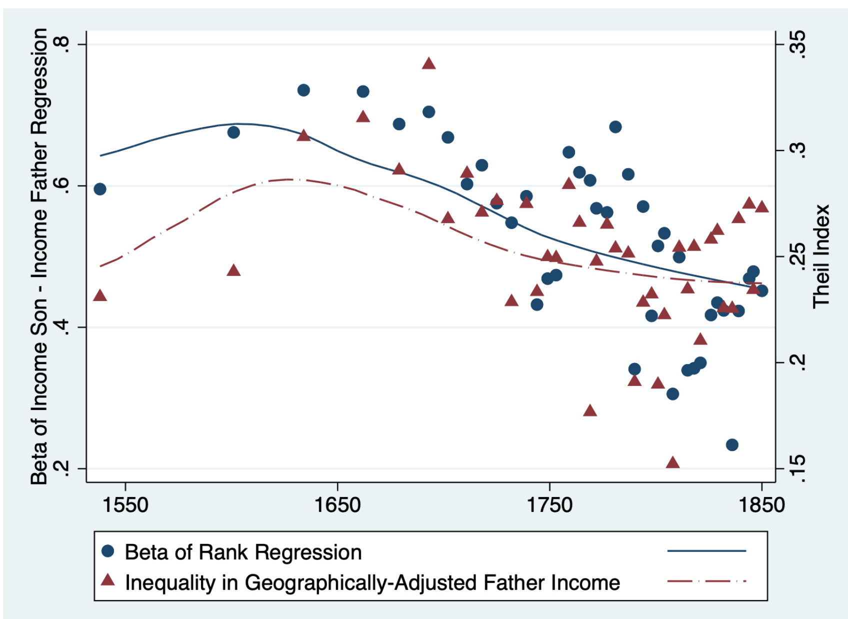
The two lines at the bottom of Figure A.9 show evidence of changes in upward mobility. While in the first half of the sample the son of a low-income great-great-grandfather could expect to be at rank 36 of the percentile distribution, in the second half of the sample period the son could expect to be at rank 46, a full ten percentile ranks higher. Furthermore, the figure indicates that the increase in upward mobility is strongest for the second to fourth generation; for the fifth generation, in contrast, there is little difference in upward mobility in the early versus the late part of the sample. Thus, there is evidence for an increase in mobility over time from five-generation-linked observations.

D Explaining Temporal Variation in Mobility: Supplemental Analyses

D.1 The Role of Geographic Location for Mobility

The results indicate that times of high cross-sectional income inequality are times of low intergenerational mobility. According to the model of section 2, this is because a changing earnings elasticity provides different incentives to parents to invest into the human capital of their child. Geographic location may matter for the efficacy of these investments in several ways. If the family lives close to an urban area, a given investment may be comparatively more effective, perhaps because personal interaction with high-income individuals (outside of the family) is more likely. While residential location is an endogenous choice of the family, it is useful to examine whether location matters for intergenerational mobility. To do so I employ information on the geographic distance of each individual's location to the most important city in the region, the capital of Tongcheng (see also the summaries in Table 2).

Figure A.10: Mobility and Inequality in Location-Adjusted Income

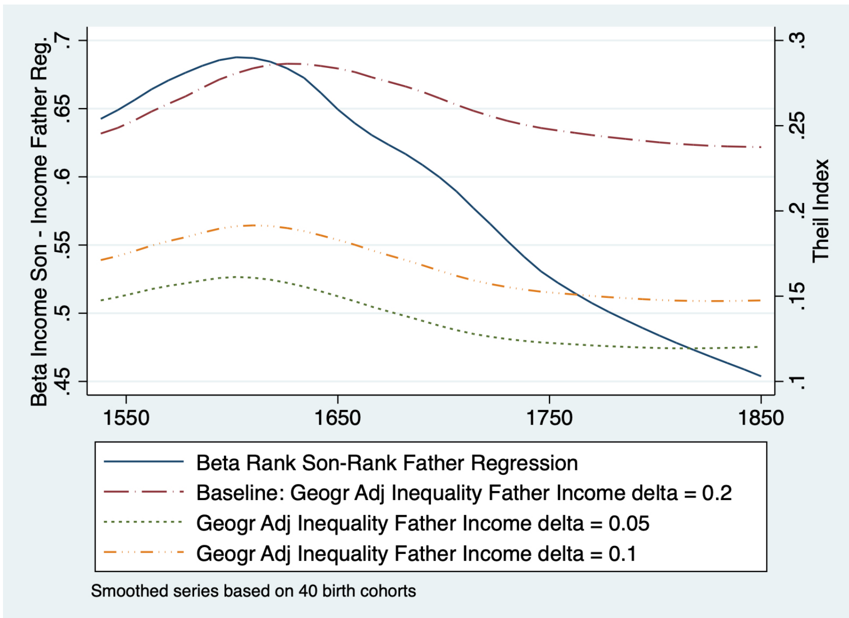


Notes: Shown is β_c and location-adjusted status inequality in the father generation. Horizontal axis shows median birth year in birth cohort; first birth year is 1330.

Using GIS methods to obtain for each individual the distance, as the crow flies, from his location to Tongcheng City, I define geographically-adjusted father income as the percentile rank divided by distance to Tongcheng City, raised to some power δ: $\frac{R_F}{dist^\delta}$, where I choose δ = 0.2 based on empirical fit (results for alternative values of δ are shown below). Figure A.10 shows that inequality in geographically-adjusted income is also positively correlated with the rank regression coefficient. While the adjusted income inequality is not higher in 1550 versus 1850, in contrast to the unadjusted income inequality measure, the correlation between inequality and mobility is now stronger (0.58, compared to 0.43). This is consistent with geographic location playing a role for mobility.

The following Figure A.11 shows that I obtain broadly similar results for alternative decay parameters δ. It presents results for two other values of δ, namely $\delta = 0.1$ and $\delta = 0.05$.

Figure A.11: Location and Cross-Sectional Inequality - Alternative Assumptions

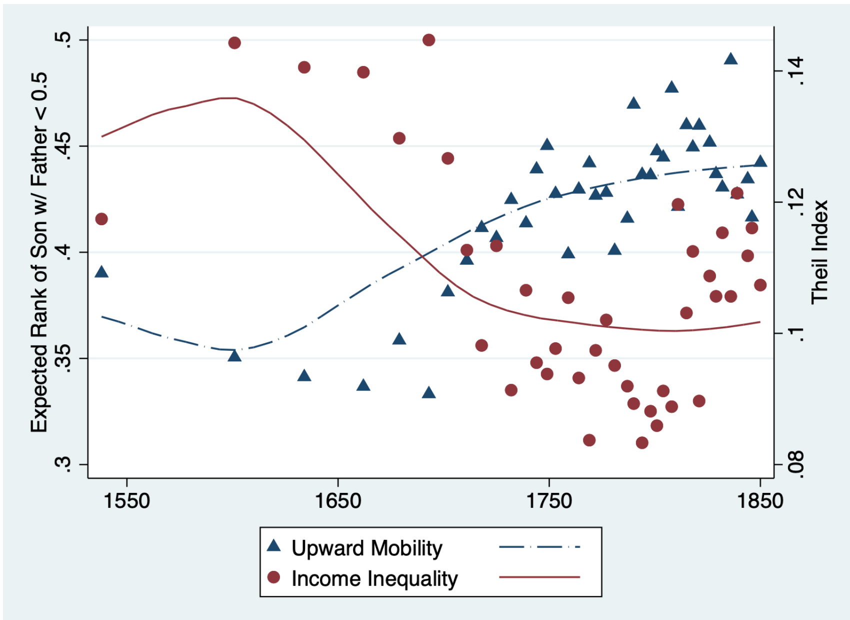


Notes: Shown is locally smoothed β (equation 11) and alternative measures of geographically-adjusted income inequality for forty birth cohorts; the first son birth year is 1330.

D.2 Income Inequality and Upward versus Downward Mobility

The following Figure A.12 shows the relationship between a measure of upward mobility and cross-sectional income inequality. Times of high inequality are times of low upward mobility (correlation is -0.80).

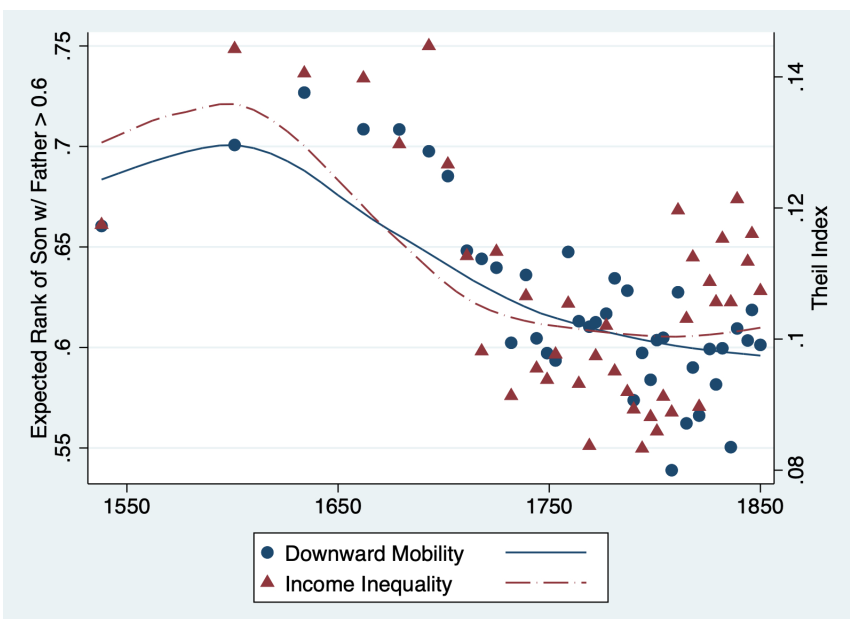
Figure A.12: Income Inequality and Upward Mobility



Notes: Shown is the Theil index of income inequality and the expected rank of sons with fathers at percentile ranks < 0.5 across birth cohorts. Horizontal axis shows median birth year in birth cohort; earliest birth year of son is 1330.

The following Figure A.13 shows the relationship between a measure of downward mobility and income inequality. Also here the high correlation, which is 0.75, confirms the results in the text.

Figure A.13: Income Inequality and Downward Mobility

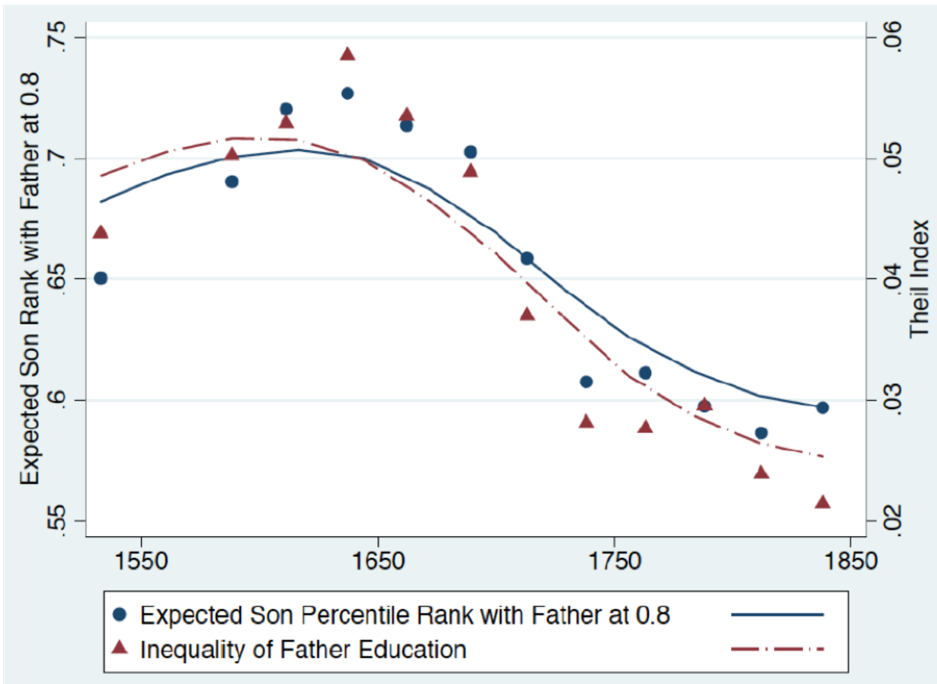


Notes: Shown is the Theil index of income inequality and the expected rank of sons with fathers at percentile ranks > 0.5 across birth cohorts. Horizontal axis shows median birth year in birth cohort; earliest birth year of son is 1330.

D.3 Educational Inequality and Mobility Changes

The following presents the relationship between a measure of downward mobility and educational inequality (Figure A.14).

Figure A.14: Educational Inequality and Downward Mobility over Time

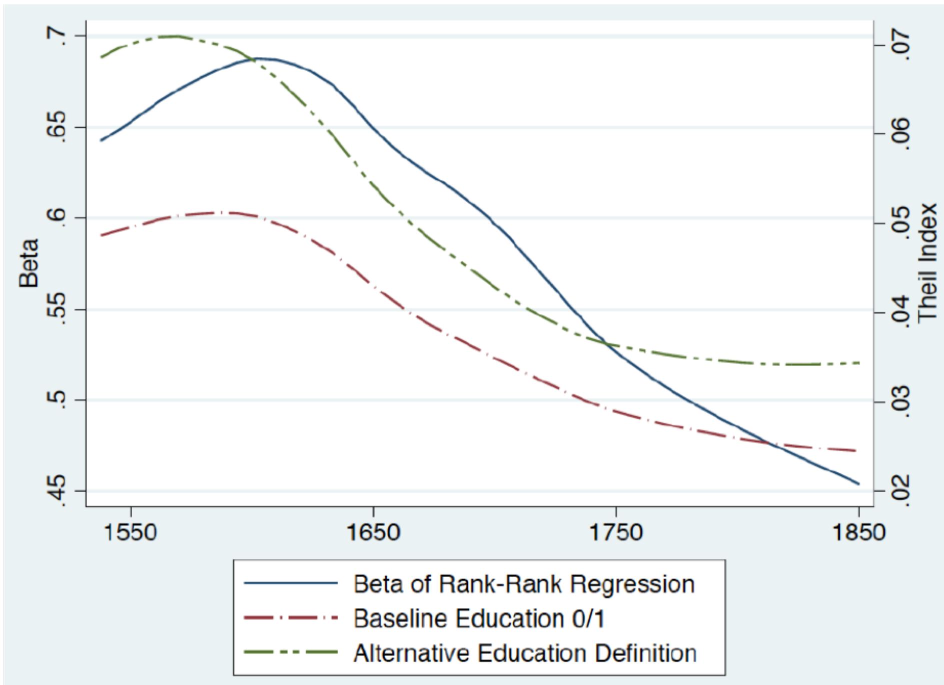


Notes: Shown is the expected rank of sons with fathers at percentile rank income > 0.6 and the Theil index of father education for twelve birth cohorts. Horizontal axis shows median birth year in cohort; birth year of first son is 1330.

Figure A.14 indicates that there is a strong relationship between downward mobility and educational inequality (correlation of 0.91). Furthermore, the correlation between upward mobility and educational mobility is also very strong (-0.91; not shown). As in the case of β_c discussed in the text in section 6.3, upward and downward mobility are more strongly correlated with educational inequality than with income inequality. This is consistent with the model in which parental human capital is complementary to nonhuman investments.

The analysis in the text provides evidence that educational inequality in the father generation is strongly and negatively correlated with social mobility. The measure of education in this baseline is an indicator variable of whether a person has prepared to take the civil entrance exam at any level. Distinguishing those that passed at a high levels from those that only passed at a lower level (or prepared but did not pass) does not change the strong relationship between educational inequality and mobility, as Figure A.15 shows.

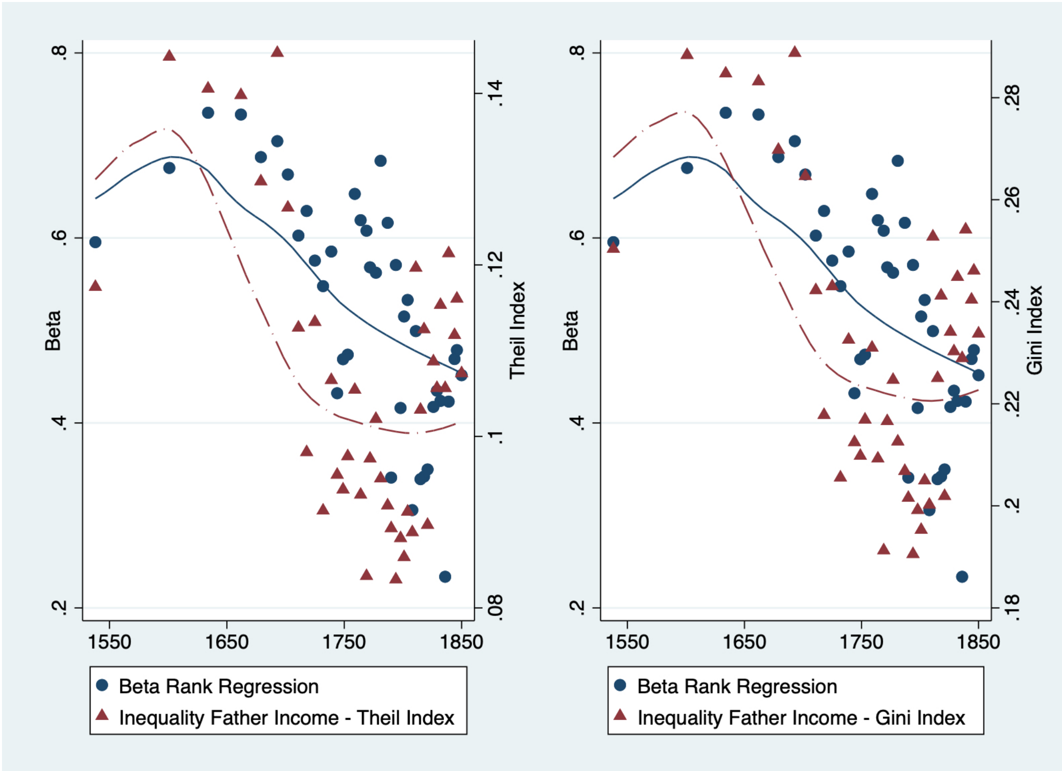
Figure A.15: Mobility and Alternative Measures of Educational Inequality



Notes: Shown are locally smoothed series β_c and Theil indices for two alternative definitions of education for forty birth cohorts. Horizontal axis shows median birth year in cohort, birth year of first son is 1330. Baseline education definition is indicator variable (Table A.1, last column). Alternative definition has education variable taking the value of 1 for preparation and passing of lower exams (income levels 6, 9, 12, and 13) and a value of 2 for passed examinations at a higher level (income levels 15, 16, and above 17), see Table A.1.

To examine the robustness of the results to different definitions of inequality, I compare the results using the Theil index with findings based on the Gini index.

Figure A.16: Alternative Measures of Inequality



Notes: Shown is β_c and inequality in father percentile rank income for forty birth cohorts; Theil index on the left, Gini index on the right. Horizontal axis shows median birth year in cohort; birth year of the first son is 1330.

As Figure A.16 indicates, the correlation between β_c and inequality is similar employing the Theil and the Gini index. Furthermore, results for the Gini and the Theil indices are similar in the case of upward and downward mobility over time as well (not shown). I have also explored other well-known measures of inequality, including the standard deviation of logs, the coefficient of variation, and the relative deviation from the mean, with similar results. I conclude that the finding that inequality is negatively correlated with intergenerational mobility is robust to alternative measures of inequality.

D.4 The Role of Clans

D.4.1 Clan Rules on Resource Sharing and Transfers

The clan as an organization form of local society was present over the entire period of the analysis and played a key role in Chinese society.[71] Family or clan laws were written documents in genealogies and were generally guided by state laws.[72] The rituals of ancestral worship helped to unify the common ties and

[71] Dynasties and genealogies were prevalent in the 20th century until the communist party under Mao Zedong suppressed paternalistic sources of authority (Yang 1959a, 1959b).

[72] One clan rule reads (translation from Liu 1959): "Clan rules rely on the law of the state as their guide. The law, in turn, depends on clan rules to supplement it. In comparison, clan rules are more lenient than the law. But whoever violates clan rules will eventually find himself in violation of the law."

bonds of the group, but more than that, the written rules of behavior had the dual effect of increasing trust and social order within the group on the one hand, while on the other hand reducing costs of enforcement when interventions by the clan elders became necessary. While the emphasis was on moral persuasion, punishments for misbehavior— such as by flogging — were set out in these family laws as well. The rules were set out by custom and tend to share broad similarities across clans, although many differences in wording and specifics can also be identified.

Importantly, these family laws were not just empty ritual but an essential part of what made the clan organization a means of insurance and resource sharing. These laws and regulations applied to a wide set of concerns with respect to common property, awards, theft, poor relief, and social relationship within the clan as well as towards non-clan individuals (Liu 1959). Similar to how progressive government policies redistribute resources, family laws had the effect of creating a range of progressive policies that applied to members of the clan.

Given the lack of publically funded education at the time, the bulk of education was undertaken at the clan level, through clan schools or hired tutors. Clans that had more resources, in particular, could set aside common lands or clan funds for the education of the children of poorer kin as well as to pay the fees and expenses associated with taking the civil service examinations. Clans also customarily used their common funds and school lands for rewards and support of its students. According to Liu (1959):

> "The greatest emphasis is placed, however, upon the honoring of successful scholars and promising aspirants as well as assisting them in the consecutive stages of their career, as the public honors and prestige they gain are shared by the clan group".

The amounts varied according to clan wealth, but the range given to those who successfully earned degrees ranged from 2-10 *taels* for the first degree to 80 *taels* for the highest degree, which was a substantial sum.

Most rules supported the financial investment in education of its students while they studied. The genealogy of the Yeh clan states: "Our clan has long remained undistinguished, since few members ever become officials or degree holders. We should therefore give in order to help promising members study." One clan held semi-monthly essay contests at its ancestral hall to encourage study, while another set up rules to support examination fees; yet another clan decided to give priority of free schooling to the orphans and poorest members in the clan (Liu 1959).

Clans with school lands used rents to support students. One such clan stipulated that half of the rent would be used for teacher's salary and the other half to scholarships (split evenly between poor members and the descendants of the original donor of the land).[73] These strategies of upward mobility could potentially pay off if a member succeeded in the civil service examinations. Wealth would have meant better access to tutors, but kinship networks and connections among the dynasty in high status positions would have helped to solidify personal advantages for other clan members.

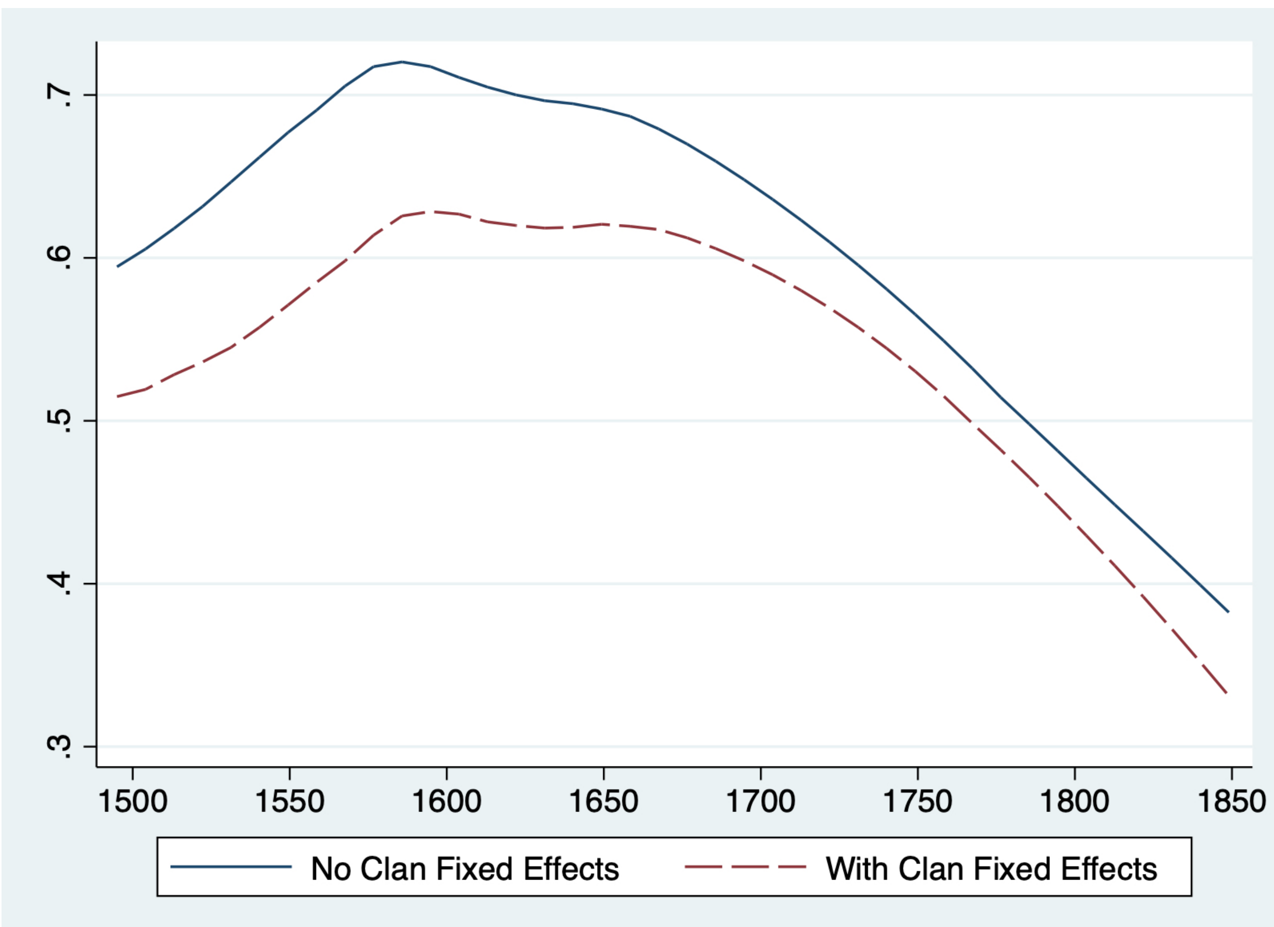
[73]According to a survey of the awards of the School and Ritual Land in genealogies by Liu (1959).

Anecdotal evidence paints a picture of how the rewards of high income was beneficial for the larger family entity. Chang (1962) documents officeholders giving up their share of the family inheritance to their brothers, or, in other cases, paying off debts. Although it is difficult to give exact magnitudes, the prestige of high income carried with it an expectation of charitable philanthropy to a larger community through financial contributions to schools, organizations, the welfare of the poor and widows. Indeed, the genealogy itself is seen by some historians and anthropologists as a document demarcating the members in the corporate group who are entitled to have access to clan lands, schools, and other resources (Watson 1982).

D.4.2 Regression Results with Clan Fixed Effects

The following result shows that mobility estimated with and without clan fixed effects evolves similarly over time.

Figure A.17: Mobility over Time and Clan Fixed Effects



Notes: Shown is β_c (equation 9) and the Theil index of father education for forty birth cohorts (smoothed series). Horizontal axis shows median birth year in cohort; birth year of first son is 1330.

E Extension: Inequality and Intergenerational Mobility: Multi-variate Results

The following extends the analysis of the relationship between inequality and mobility to a multi-variate regression framework. Table A.7 shows OLS results for relative mobility,

$$\beta_c = \pi_0 + \pi_1 X_c + \varepsilon_c, \qquad (4),$$

where X_c is a vector of correlates of the estimated birth-specific β_c. Reported are standardized coefficients once variables are transformed to have mean zero and standard deviation of one (also called beta coefficients), with robust t-statistics in parentheses.

Table A.7: Cross-Sectional Inequality and Intergenerational Mobility - Multi-Variate Analysis

	(1)	(2)	(3)	(4)	(5)
Father Income Inequality	0.429 (3.449)		0.065 (0.346)		
Geogr. Adj. Father Income Inequality		0.575 (4.432)	0.530 (2.853)		0.157 (1.232)
Educational Inequality				0827 (10.924)	0.739 (7.771)
R^2	0.184	0.331	0.333	0.684	0.701

Notes: Dependent variable is β_c for N = 40 birth cohorts. Estimation by OLS; standardized coefficients reported, with robust t-statistics in parentheses. Inequality is Theil index.

The first variable introduced on the right hand side is father income inequality, T_c. The result in column (1) indicates that cross-sectional income inequality is significantly associated with persistence, confirming the patterns in Figure 8. With a t-statistic of close to 3.5, the correlation is significant at standard levels. Results for inequality in location-adjusted father income are shown in column (2). Geographic location increases the correlation between income inequality and mobility, with the R^2 almost twice of what is obtained in column (1). When both location-adjusted and unadjusted income inequality variables are included, the former is more powerful in accounting for variation in relative mobility (column (3)).

Next, Table A.7 confirms the strong relationship between educational inequality, T_c^e, and mobility (column (4)). Variation in fathers' human capital, a single variable, account for more than two thirds of all the variation in mobility. The analysis also shows that educational inequality is a more powerful predictor of intergenerational mobility than income inequality, even if location-adjusted (column (5)).

The following set of results examines the role of average income, as well as average education levels and average longevity (age at death), for the results. Recall that father income is in terms of percentile rank for a given birth cohort, so that the average is 0.5 by construction. However, if one calculates the income average in a given birth cohort using the six income levels, zero to five (see Table 1), changes in average income might account for some of the variation in mobility. In addition, other factors such as average grandfather status or family size might also help to account for mobility differences. These questions are addressed in the following Table A.8.

Table A.8: Cross-Sectional Inequality and Intergenerational Mobility - Other Influences

	(1)	(2)	(3)	(4)	(5)	(6)
Income	−0.13					
Average	(−0.83)					
Education		−0.73				
Average		(−2.34)				
Longevity			0.16			
Average			(2.95)			
Grandfather				0.07		
Inequality				(0.48)		
Geogr.					0.15	
Dispersion					(1.29)	
Family						−0.03
Size						(−0.39)
Educational	0.94	1.34	0.77	0.77	0.92	0.81
Inequality	(5.55)	(4.42)	(11.26)	(5.28)	(9.57)	(9.91)
R^2	0.69	0.71	0.71	0.69	0.70	0.69

Notes: Dependent variable is β_c; estimation by OLS. Family size is number of brothers including the son himself. Standardized coefficients shown, robust t-statistics in parentheses. N = 40.

The first specification in A.8 introduces the average of income in a birth cohort (values from 0 to 5, column 1 in Table 1) together with cross-sectional inequality in education. Recall that in a bivariate analysis educational inequality enters with a positive coefficient of 0.83 (see Figure 8). Adding educational inequality together with average income turns the latter coefficient to close to zero (column (1)). The regression with average education and inequality in education has a negative coefficient on average education and a larger positive on educational inequality (column (2)). We see in column (3) that average longevity enters significantly positive with educational inequality. This is consistent with health having an independent influence on intergenerational mobility, although quantitatively, the role of educational inequality is about five time as large (the coefficients are standardized). In contrast, there is no significant independent correlation of grandfather inequality, geographic dispersion, and family size once inequality in education is included (columns (4), (5), and (6)).

Summarizing, the results indicate that educational inequality is a robust and important correlate of intergenerational mobility. This provides additional support both for the role of parents' investments into their children and the complementary role of parental human capital in making them.