California Center for Population Research
University of California - Los Angeles

Social, Economic, and Demographic Determinants of Descent Line Growth and Extinction Over the Long Term in Historical China

Cameron Campbell
James Lee

PWP-CCPR-2010-005

March 2010
Last Revised: March 2010

California Center for Population Research
On-Line Working Paper Series

Social, Economic, and Demographic Determinants of Descent Line Growth and Extinction Over the Long Term in Historical China

Cameron Campbell
University of California, Los Angeles
camcam@ucla.edu

James Lee
Hong Kong University of Science and Technology
jqljzl@ust.hk

This paper was originally prepared for the workshop "The Persistence of the Past," May 25 – 27, 2009, Max Planck Institute for Demographic Research, Rostock, Germany. Revised versions were presented at the Institute of Behavioral Science Population Program and Center for Asian Studies, University of Colorado, September 17, 2009, Workshop in Economic History, Northwestern University, April 22, 2009, the Center for Population Economics, University of Chicago, April 23, 2009. We are grateful to the discussants and audience members for their comments and suggestions.

This research was supported by NICHD 1R01HD045695-A2 (Lee PI), "Demographic Responses to Community and Family Context."

This is a work in progress, and subject to substantial revision. Please check with the authors to obtain the latest version before citing or circulating.

ABSTRACT

We assess the implications of demographic differentials associated with socioeconomic status for long-term changes in population composition through analysis of a household register data from northeast China during the eighteenth and nineteenth century. Specifically, we investigate whether the descendants of socioeconomically advantaged men accounted for a disproportionate share of the population in later generations. We take advantage of a multigenerational dataset of historical population registers that allows for direct measurement of numbers of surviving descendants in subsequent generations because it spans seven generations over 150 years, covers a population that is largely closed to in- and out-migration, and supports reconstruction of male pedigrees via automated record linkage. We use these data to relate the numbers of descendants in subsequent generations to male socioeconomic status and other characteristics. The study contributes to the growing literature on interactions between social stratification and socioeconomic differences in demographic behavior. In particular, we assess whether processes like ones that Clark (2007) suggested operated in preindustrial England were also present in historical China. Preliminary results indicate that descendants of socioeconomically privileged males accounted for larger shares of the population two and four generations later, suggesting that processes like the ones Clark suggested for England may have operated in China as well.

Introduction

We explore how differences between the social, economic, and demographic performance of lineages shaped patterns of inequality and social organization over the very long term. The key question is whether differences in marriage, reproduction, and mortality by social and economic status worked together to drive changes in population composition across generations, or were cancelled out by each other or other factors. If the net result of differences in marriage, reproduction, and mortality by socioeconomic status was that socially and economically advantaged lineages experienced more rapid demographic growth over the long term, members of such lineages would account for a steadily rising share of the population. To the extent that knowledge, skills, and attitudes conducive to socioeconomic success were transmitted within lineages, this would also be a mechanism for the diffusion of such human capital through the population, leading over the long term to social and perhaps economic change.

We test directly whether socioeconomic differentials in demographic behavior combined with the transmission from parents to children of traits conducive to socioeconomic success contributed to long-term population change in China in the same way that Clark (2007) proposed was the case in preindustrial England. Clark suggested that preindustrial England was distinguished by a link between economic advantage and reproductive success that when combined with intergenerational transmission of attitudes, knowledge, and skills led to the dissemination through the population of traits conducive to economic growth. Clark suggested that this interaction between reproductive advantage for the most privileged and the transmission of traits conducive to socioeconomic success was unique to preindustrial England, and that by driving the dissemination of attitudes, skills, and knowledge that were advantageous in a market economy, it helps account for why England experienced the Industrial Revolution before other countries. If we observe processes in northeast China like the ones Clark argued characterized preindustrial England, the implication would be that such processes were not unique to preindustrial England. Such processes may very well have played an important role in social and economic change, as Clark argued they could, but they may have been more widespread than he suggested.

The paper contributes to the rapidly growing literature in demography and stratification on interactions between population composition, social stratification, and socioeconomic differentials, in demographic behavior. The distinguishing feature of the analysis is that the implications of socioeconomic differences in demographic behavior for trends in population composition are measured directly from a multigenerational database in which a population was recorded prospectively. Almost all previous studies in demography and stratification apply population renewal models and other techniques of demographic projection to infer implications for population composition and patterns of inequality of demographic differentials and intergenerational transmission of status (Preston and Campbell 1993; Mare 1996, 1997; Mare and Maralani 2006; Musick and Mare 2004). Such simulation-based studies rely heavily on assumptions about the relative strengths of specific relationships among demographic rates and patterns of status transmission. By contrast, this study measures changes in population composition empirically and accordingly makes no assumptions about the structure of correlations in demographic and social outcomes.

The paper also contributes to the literature on kinship in Chinese society by examining the dynamics of lineage growth, decline, and extinction prospectively. Lineages have long been a preoccupation in the study of Chinese kinship. There are countless anthropological, historical, and sociological studies of particular lineages, especially in south China. Such studies describe in detail the internal dynamics of lineage organizations, the histories of specific lineages, and the roles that lineages played in the community, region, or nation. Their most important limitation is that because they are retrospective, they overemphasize the experience of the successful lineages that not only survived but prospered, and neglect the lineages that became extinct, or survived but were too small or undistinguished to come to the attention of a researcher. The ideal study of the determinants of lineage growth or decline should be prospective, starting with a collection of lineage at some point in time and following them forward to compare the winners and losers.

We assess the implications of socioeconomic differentials in demographic behavior for inequality over the long term by measuring whether the descendants of socioeconomically advantaged men accounted for a disproportionate share of the population in later generations. We take advantage of the multigenerational nature of our historical population registers to measure directly the numbers of surviving descendants in subsequent generations. The Liaoning population data are suited to this task because they span seven generations, covers a population that is largely closed to in- and out-migration, and supports reconstruction of male pedigrees via automated record linkage. Through regression analysis, we relate the numbers of descendants in subsequent generations to male socioeconomic status and other characteristics.

Background

Socioeconomic status and fertility were positively associated in historical China (Harrell 1986; Lee and Campbell 1997; Wang, Lee, and Campbell 1995; Wolf 1980, 1995), but such an association doesn't guarantee that descendants of the privileged were overrepresented in later generations. Trends in population composition depended on interactions between differentials in marriage, fertility and survival as well as processes of intergenerational transmission of status. Especially in historical China, socioeconomic gradients in fertility and survival may have counteracted each other, with socioeconomic advantage conferring higher fertility rates but lower survival chances (Lee and Campbell 1997). Alternatively, fertility differentials by socioeconomic status may have had only transitory effects on population composition, because children growing up in larger families may have been more likely to experience downward mobility as a result of resource dilution and experience higher mortality or a reproductive disadvantage, yielding no net advantage in numbers of grandchildren.

Simulation has been used to assess the implications of demographical differentials and intergenerational transmission for population composition, but is not suited to the specific needs of this analysis. Such simulations typically apply some extension to population projection that incorporates assumptions about fertility and mortality differentials by socioeconomic status, and intergenerational mobility between statuses across generations (Preston and Campbell 1993; Mare 1996, 1997; Mare and Maralani 2006; Musick and Mare 2004). In this approach, the population is divided into a fixed number of groups defined by socioeconomic status or some other variable, and the size and composition of the groups is projected by aging the population. The limitation of such macro-level approaches is that they do not account for micro-level feedbacks and correlations that may influence macro-level population composition. For example, it is difficult for such models to incorporate complex feedbacks by which children born to larger families also experience higher mortality or greater downward mobility. It is also difficult for such approaches to incorporate assumptions about correlations within families, lineages, or communities in unobserved characteristics that influence demographic behavior and socioeconomic attainment. In principle, a very detailed micro-simulation could address these limitations by incorporating appropriate assumptions about relevant processes, but given that the number of possible assumptions to choose from to incorporate is unbounded, strong priors would be needed to select among them.

Another approach to studying the implications of demographic differentials and intergenerational transmission for population composition is direct measurement. The problem is that appropriate data that follow populations prospectively over generations are not widely available, and the data that are available have not been exploited to study the determinants of lineage growth. One prominent exception is the study of preindustrial England by Clark (2007). Since prospective data were not available that followed the English population prospectively across multiple generations, Clark pieced together indirect evidence. Most relevantly, he proxied the growth of lineages by examining the prevalence of rare surnames for which a founder could be plausibly identified and their socioeconomic status reliably measured. From these and other pieces of indirect evidence he argued that the descendants of socioeconomically privileged males were overrepresented in the population generations later, and that if they transmitted traits conducive to success to their offspring, this would diffuse knowledge, skills, and attitudes conducive to success in a market economy through the population.

We address the same question in northeast China, but improve on the analysis in Clark (2007) by measuring the growth of lineages over generations directly from a database of population register that follows patrilines prospectively for 150. We describe features of the Liaoning population database relevant to the measurement of lineage growth in detail in the following section. Here we measure net reproductive success in terms of numbers of surviving descendants 2 and 4 generation later. Any differences in population composition apparent 2 and 4 generations later are net of fertility and mortality differentials. If they exist, they would challenge Clark's claim that there was no combination of reproductive differentials and intergenerational transmission of traits, attitudes, and knowledge in China comparable to the combination he claims existed in England.

Our approach is straightforward. Through record linkage, we construct counts of the numbers of surviving male descendants for men 48 and 96 years later. We use these counts as an outcome measure in a regression analysis in which the explanatory variables are measures of men's current socioeconomic status. The results assess directly whether the descendants of economically advantaged men are overrepresented in subsequent generations. We begin by introducing our data, briefly discuss our results, present descriptive results, and then present our multivariate results. We show that even after the introduction of controls for current fertility, descendants of socioeconomically advantaged males are overrepresented in the population 2 and 4 generations later. Processes of descent line extinction, survival, and growth were not random, but in fact interacted with socioeconomic characteristics.

Data

The data we analyze consists of triennial household register data for 1749 to 1909 for more than 600 villages in Liaoning province in northeast China. The database comprises 1.4 million observations of one-quarter million people who lived in 28 administrative populations between 1749 and 1909. Figure 1 summarizes the distribution of observations by year. We have been able to produce such historical data because of the internal consistency of the core household register data, their availability through the Genealogical Society of Utah and the Liaoning Provincial Archives, and the sustained efforts of teams of colleagues and data entry personnel in the People's Republic of China. We have already described the origins of the registers as well as our procedures for data entry, cleaning and linkage in Lee and Campbell (1997, 223-237).

Figure 1 here

The features of the registers relevant for studying specific demographic and social outcomes have been described in published investigations of the determinants of individual survivorship (Campbell and Lee 1996, 2000, 2002, 2004), transmission of household headship (Lee and Campbell 1998), migration (Campbell and Lee 2001), ethnic identity (Campbell, Lee, and Elliott 2002), social mobility (Campbell and Lee 2003b), the influence of secular economic change on demographic behavior (Lee and Campbell 2005), and kinship organization (Campbell and Lee 2006). The description of the data here is based in large part on the discussions of the data in these publications, especially Campbell and Lee (2006).

The geographic and economic contexts of these populations varied. As Figure 2 shows, the more than 600 Liaoning villages are arranged in four distinct regions spread over an area of 40,000 square kilometers, larger than the province of Taiwan. These regions include a commercialized coastal area around Gaizhou that we identify as Liaoning South in the analysis, a farming region around Haizhou and Liaoyang that we identify as Liaoning South Central, an administrative center on the Liaodong Plain around the city of Shenyang that we refer to as Liaoning Central, and a remote agricultural area in the hills and mountain ranges in the northeast that we refer to as Liaoning Northeast. The institutions, regions and communities covered in the data are diverse enough that even if the population is not representative of China or even Liaoning in a formal statistical sense, results are likely to be relevant for understanding family and social organization in other parts of China.

Figure 2 here

The Liaoning household registers provide far more comprehensive and accurate demographic and sociological data than other household registers and lineage genealogies available for China before the twentieth century (Harrell 1987, Jiang 1993, Skinner 1987, Telford 1990). This is because the Northeast, which was the Qing homeland, was under special state jurisdiction, distinct from the provincial administration elsewhere. Regimentation of the population actually began as early as 1625, when the Manchus made Shenyang their capital and incorporated the surrounding communities into the Eight Banners (Ding 1992, Elliott 2001). By 1752, with the establishment of the General Office of the Three Banner Commandry, the population was also registered in remarkable precision and detail, and migration was strictly controlled, not just between Northeast China and China Proper, but between communities within Northeast China as well. Government control over the population was tighter than in almost any other part of China (Tong and Guan 1994, 1999). Movement within the region was annotated in the registers, and individuals who departed the area without permission were actually identified in the registers as 'escapees' (taoding).

The Qing state implemented a system of internal cross-checks to ensure the consistency and accuracy of the registers. First, they assigned every person in the banner population to a residential household (*linghu*) and registered him or her on a household certificate (*menpai*). Then they organized households into groups (*zu*), and compiled annually updated genealogies (*zupu*). Finally, every three years they compared these genealogies and household certificates with the previous household register to compile a new register. They deleted and added people who had exited or entered in the previous three years and updated the ages, relationships, and official positions of those people who remained as well as any changes in their given names. Each register, in other words, completely superseded its predecessor.

The result was a source that closely resembled a triennial census in terms of format and organization. Entries in each register were grouped first by village, then by household group (*zu*) and then by household. Individuals in a household were listed one to a column in order of their relationship to the head, with his children and grandchildren listed first, followed by siblings and their descendants, and uncles, aunts, and cousins. Wives are always listed immediately after their husbands, unless a widowed mother-in-law supersedes them. For each person in a household, the registers recorded relationship to household head; name(s) and name changes; adult occupation, if any; age; animal birth year; lunar birth month, birth day, and birth hour; marriage, death, or emigration, if any during the intercensal period; physical disabilities, if any and if the person is an adult male; name of their household group head; banner affiliation; and village of residence.

The registers also record official positions held by adult males. We have identified four broad categories of official position: banner, civil service, examination, and honorary. These constituted the local elite. The first three categories were formal governmental offices and included a salary and other perquisites. They predominantly comprise lower-level occupations such as soldier, scribe, or artisan. Positions also included some high administrative offices that entailed not only a salary, but substantial power as well. Positions and titles in the fourth category, honorary, were typically purchased, and indicate substantial personal resources or access to such resources through the family. For the purposes of this analysis, we divide the official positions into three categories according to their salaries. The salaries in the lowest category were equivalent to the annual consumption needs of 10-15 individuals, and the salaries in the highest category were equivalent were equivalent to the consumption needs of 50-100 individuals.

We also treat disability status as an indicator of socioeconomic advantage because in a preindustrial rural setting, disability was likely to be associated with impaired capacity for agricultural labor. The registers distinguish disabled adult males, classified as *feiding* or *chenfei*, from other adult males (Lee and Campbell 1997). Campbell and Lee (2003) assess the quality of the disability data and examine causes and consequences of being recorded as detailed in detail. Classification as disabled could occur for any one of a number of reasons and until 1786, the registers generally specified a specific disease or condition for each disabled male. Reflecting the prevalence of tuberculosis in the eighteenth and nineteenth centuries, respiratory diseases, especially consumption (*laozheng*), were by far the most common, affecting more than 5 percent of adult males and 25 percent of all the disabled. Eye diseases were second most common, affecting more than 3 percent of adult males and 15 percent of the disabled, followed by such neurological disorders as retardation, insanity, dementia, and epilepsy which affected and afflicted 2 percent of adult males and 10 percent of the disabled.

Other proxies for socioeconomic status available in the registers include being an eldest son, service as a *zuzhang*, and being recorded with a diminutive name. Eldest sons were supposed to be privileged because of their role as family head, or heir to the family headship. They married earlier and were otherwise advantaged (Lee and Campbell 1997). *Zuzhang* were heads of groups anywhere between 1 and 4 households. This was the lowest position in the administrative hierarchy and was not salaried, thus in itself was not a source of privilege. However *zuzhang* were selected by the households themselves, ostensibly on the basis of ability, thus selection to serve as a *zuzhang* may have reflected advantage in the form of talent or possession of useful skills. Finally, recording with a diminutive name in adulthood appears from other analysis to have been associated with disadvantages on a number of outcomes such as marriage and attainment of official position. It seems to have been a marker for lack of aspirations or a acknowledgment of a lower socioeconomic status.

In contrast with most historical censuses, the triennial registers allow for linkage of the records of an individual in successive registers. Households and their members appeared in almost the same order in each register, even if they moved to another village. Linkage from one register to the next is straightforward: as our coders transcribe each new register, for each individual they list the record number of his or her entry in the previous register. Since the coders transcribe each new register by copying over the file for the preceding register and then editing it, this is straightforward. From the linked records for each individual, we reconstruct life histories. By comparing observations for the same individual in successive registers, we can construct outcome measures indicating whether particular events or transitions took place in the time between two successive registers.

These registers have a number of features that distinguish them as a source for historical demography. The population is closed, in the sense that the registers followed families that moved from one village to another within the region. Entries into and exits from the region were rare, and when they did occur, their timing was recorded or can be inferred (Lee and Campbell 1997, 223-237; Lee and Wang 1999, 149-153). In contrast with historical Chinese demographic sources such as genealogies that only record adult males, the Liaoning registers record most boys and some girls from childhood, as well as all women from the time of their marriage. Unlike genealogies, they also provide detail on village and household residence. In contrast with parish registers, an important source for European historical demography, they allow for precise measurement of the population at risk of experiencing most demographic events and social outcomes.

As a result, the registers are uniquely suited to prospective reconstruction of descent lines through intergenerational record linkage, and prospective study of the predictors of descent line growth, decline, and extinction. Figure 3 summarizes the proportions of male children whose fathers, grandfathers, and earlier ancestors have been identified through automated record linkage. Approximately eighty percent of boys in the registers in the first decade of the twentieth-century may be linked back to a great-great-great-grandfather, and more than ninety percent can be linked back to a great-great-grandfather. The remaining men are mainly descendants of men who entered the registration system in the nineteenth century because they migrated in. Based on the results of this linkage, we generate the counts for males of the numbers of surviving male descendants recorded 24, 48, 72, and 96 years later that are the basis of our analysis.

Figure 3 here

A key difference from Chinese lineage genealogies that is crucial to this analysis is that recording of adult males in the household registers was not selective in terms of socioeconomic status, marital status, or number of offspring. Because Chinese lineage genealogies are almost always compiled retrospectively by the descendants, men who had few or no offspring were more likely to be omitted (Campbell and Lee 2002), biasing counts of numbers of offspring upward, and underestimating the frequency of extinction. A related problem is that in Chinese lineage genealogies, men of higher status were more likely to be included and men of lower status were more likely to be omitted, because one of the purposes of a genealogy was to demonstrate the overall prestige of the lineage and document an individual's connection to a prestigious ancestor or other relative. To the extent that the descendants of such men were more assiduous about maintaining their genealogy, an analysis might overstate the relationship between socioeconomic status and numbers of descendants generations later.

The registers omit many boys who died in infancy or early childhood, and omit most daughters, but these are not relevant to the analysis here because of the focus on the numbers of surviving male descendants. These omissions mainly affect the estimation of infant and child mortality, and fertility. Thus the registers do not allow for estimation of infant or early childhood mortality for either sex. While they yield reliable estimates of mortality in later childhood and early adolescence for boys, the estimates for girls may be based on the potentially selective subset of families who bothered to register daughters. Analysis of fertility, meanwhile, is based on boys who survived long enough to be included in the registers. Apparent fertility differentials may also reflect differentials in infant and child survival.

Methods

We estimate three different types of regression to assess how the current status of males affected the numbers of descendants they would have 48 and 96 years later. We consider three different indicators of descent line survival and growth: how many surviving male descendants there were 48 and 96 years later, whether or not a male was in the top 15 percent of men in terms of the numbers of surviving male descendants 48 and 96 years later, and whether or not a male had any male descendants at all 48 and 96 years later. For the regression of total number of surviving descendants, we estimate a negative binomial regression, which like Poisson regression is typically used for count variables, but differs in that it allows for overdispersion in the distribution. For the regression of the chances of being in the top 15 percent in terms of number of surviving descendants, and the regression of whether or not there were any descendants at all, we use logistic regression. We choose 48 and 96 years later as times to measure outcomes because they correspond roughly to 2 and 4 generations. The actual mean generation length may have differed but because registers are 3 years are apart, any length of time we choose for the follow-up must be a multiple of 3.

We consider these different measures of the survival and growth of descent lines to ensure that our conclusions are robust to alternative specifications of the outcome. The first specification, numbers of surviving male descendants, is an overall general measure of reproductive success and survival. The second specification, being in the top 15 percent of men in terms of male descendants, allows for the possibility that the determinants of being especially successful differed from the determinants of average numbers of descendants. The final specification, having any descendants at all, allows for the possibility that the determinants of extinction differ from the determinants of average number of children.

We consider the effects of multiple measures of socioeconomic advantage described earlier in the section on data. We consider three categories of income for official position, recording as disabled, recording with a diminutive name, *zuzhang* status, and being an eldest son. Positive coefficients for any of these variables indicate that descendants of such men are overrepresented in the population 48 or 96 years later. Negative coefficient indicate that descendants of such men are underrepresented in subsequent generations.

To account for the possibility that observed differences 48 and 96 years later are only artifact of transitory reproductive success of advantaged males in the current generation, we also estimate models that include counts of numbers of boys born. Descendants of men who had more children in the current generation should be overrepresented among the population in future generations if only because those children were all at risk of marrying and having children themselves. Effects of socioeconomic advantage that persist in models that include controls for current reproduction indicate that overrepresentation in future generations is not simply path dependence or a random walk, but reflects reproductive advantages that persist in subsequent generations, most likely because of parallel transmission of socioeconomic status that is the basis for such advantage. If coefficients for the measures of socioeconomic advantage all become 0 after the introduction of a control for current number of boys, the implication is that effects of reproductive advantage associated with current socioeconomic advantage are transitory, and do not lead in the long run to overrepresentation of the descendants of such men in the population.

Results

Descriptive

Extinction of descent lines was a common phenomenon in Liaoning, as elsewhere. Table 1 presents distributions of men according to the numbers of surviving male descendants they had 24, 48, 72, and 96 years later, as counted from the household registers. The proportions of men currently living who had no surviving descendants was grew larger moving forward in time, so that whereas only one-sixth of men had no living male descendants 24 years later, 45 percent had no living male descendants 96 years later. At the same time, the share of men who had large numbers of descendants also grew as the time horizon considered move forward. In other words, as the time considered moved forward, men were increasingly divided into those who would have no living descendants, and those who had many.

Table 1 here

As a result, descendants of a steadily smaller share of men accounted for a steadily larger proportion of males as time went forward. Figure 4 presents Lorentz curves that summarizes the shares of males accounted for 24, 48, 72, and 96 years later by shares of men sorted in order of their reproductive success. 48 years later, more than one-third of males were accounted for by the descendants of the top 10 percent of males in terms of reproductive success, and more than one-half were accounted for by the top 20 percent. 96 years later, more than one-half of males 96 years later was accounted for by the descendants of the top 10 percent of males, and more than three-quarters of males were accounted for by descendants of top 20 percent of males.

Figure 4 here

Processes of descent line growth, decline and extinction were anything but random because socioeconomic and demographic outcomes varied systematically and substantially between descent groups in the same village, and between households within the same descent group. Table 2 presents results on the standard deviations of intercepts from models that include random effects of village, descent group, and household. Comparisons of the standard deviations for the intercepts of the different levels illuminate the relative importance of variation at different levels in accounting for individual outcomes. For example, a larger standard deviation for the intercepts for village and a small standard deviation for the intercepts for descent group would indicate that differences in an outcome between village were more pronounced than differences in that outcome between descent groups in the same village. According to the results, not only was there systematic variation in key demographic and social outcomes according to village, descent group, and household, but patterns varied according to the outcome. Broadly speaking, demographic outcomes such as mortality and fertility varied more between villages than between descent groups in the same village, or households in the same descent group. Social outcomes such as marriage, attainment of official position, and use of diminutive names varied more within villages than between villages.

Table 2 here

Arithmetic tabulations confirm that the descendants of higher status males accounted for a disproportionate share of the male population 96 years later. Table 3 presents tabulations of adult males by various categories of socioeconomic status, and tabulations of males 96 later according to whether they were descended from males in each of these categories. The share of the population 96 years later accounted for males descended from men in the two highest categories of income was larger than the share of the present population accounted for by men in those categories. Similarly, the share of the population accounted for by descendants of privileged males such as *zuzhang* and eldest brothers was larger than the share of the current population accounted for by such men. Conversely, the share of population 96 years later accounted for by the descendants of men with diminutive names was smaller than the share of men currently who had such names. The only trend in population that appears counter-intuitive is that the share of men who are descendants of disabled men appears higher than the share of men who are disabled.

Table 3 here

Multivariate

Socioeconomic advantage was associated with larger numbers of surviving descendants. Table 4 presents results from negative binomial regressions of the numbers of surviving male descendants 48 and 96 years in the future. According to a model with no controls for the number of boys already born, men with official positions that paid salaries in the 24 to 36 tael range had 1.42 (=$e^{0.352}$) as many surviving descendants 48 years later as men without an official position, and 1.29 times as many surviving descendants 96 years later. Similarly, men whose official positions were in the highest income category had 1.74 times as many surviving descendants 48 years later, and 2.08 times as many surviving descendants 96 years later. Conversely, men who had diminutive names had 0.86 times as many descendants 48 years later as men without such names, and 0.90 times as many descendants 96 years later as other. Effects for *zuzhang* and disability status were statistically significant and in the expected direction, while effects for being an eldest brother were not statistically significant.

Table 4 here

These differences in success reflected reproductive advantage not only of the founder male, but persistent reproductive or survival advantage of his descendants in later generations as well. Statistically significant differences by income, *zuzhang* status, and diminutive name all persist after the introduction of controls for the number of boys already born to a male. This rules out the possibility that advantages apparent generations later are simply an outcome of a path dependent process where a transitory increase in reproduction associated with an advantaged current status leads to larger numbers of descendants generations later because there are more offspring in the next generation who experience average levels of reproduction. The persistence of effects after controls for current reproduction suggest that the descendants of privileged males themselves enjoyed a reproductive or survival advantage, presumably socioeconomic status that conferred such advantage was transmitted from father to son across generations.

Patterns of effects are similar when we consider an alternative outcome, an indicator of whether or not a male is among the top 15 percent in terms of numbers of descendants 48 or 96 years later. Effects of being in the top two income categories, and being *zuzhang*, were strongest and most robust. Even after controlling for number of boys born, men in the highest income category were 2.32 times more likely ($e^{0.842}$) to be among the top 15 percent of men ranked by number of descendants 96 years later. Men who were *zuzhang* were 25.4 percent more likely to be in the top 15 percent. These results confirm that the ones in Table 4 are robust to alternative specifications of measures of reproductive success, in particular a specification that makes no assumptions about the distribution of the outcome measure.

Table 5 here

Similarly, socioeconomic advantaged substantially reduced the chances of descent line extinction over the long term. Table 6 presents results from a logistic regression in which the outcomes considered are whether males had any surviving descendants at all 48 and 96 years later, with and without controls for their current number of sons born. Effects for income, *zuzhang* status, diminutive name, and disability status were all large, statistically significant and in the expected direction, and all except the effects of disability persisted after the introduction of controls for current reproduction. Controlling for current number of sons, men in the highest income category were 2.86 times ($e^{1.052}$) times more likely to have at least one surviving male descendant 96 years later. Effects of being an eldest brother were not statistically significant.

Table 6 here

Conclusion

In northeast China in the past, socioeconomic gradients in reproduction had implications for trends in population composition. Descendants of advantaged men were overrepresented in later generations, in the sense that they accounted for a larger share of the population 48 and 96 years later than their forefathers. While the magnitudes of the effects here may appear small, they are observed only over four generations. To the extent that some of the effects four generations later represented processes that worked over even more generations, the cumulative effects over centuries might have been much larger. Given the limited time depth of the available data, exploring the possible implications of the processes reported here for trends in population composition over centuries or millenia will require demographic projection or simulation that can take the results here as inputs. In the meantime, we can speculate that if these processes were general to all of China, and persisted over centuries or millenia, there is a reasonable probability that the ancestors of the contemporary population were not a random draw from the population of their day, but were concentrated among the elite.

References

Campbell, Cameron and James Lee. 2002 (publ. 2006). "State views and local views of population: Linking and comparing genealogies and household registers in Liaoning, 1749-1909." *History and Computing.* 14(1+2):9-29.

Clark, Greg. 2007. *A Farewell to Alms: A Brief Economic History of the World.* Princeton: Princeton University Press.

Harrell, Stevan. 1985. "The Rich Get Children: Segmentation, Stratification, and Population in Three Zhejiang Lineages." In Susan B. Hanley and Arthur P. Wolf, eds., *Family and Population in East Asian History*. Stanford: Stanford University Press, 81-109.

Lee, James and Cameron Campbell. 1997. *Fate and Fortune in Rural China: Social Organization and Population Behavior in Liaoning, 1774-1873.* Cambridge University Press.

Mare, Robert D. 1996. "Demography and the evolution of educational inequality." In J. Baron, David Grusky and Donald Treiman, eds. *Social Differentiation and Social Inequality: Theoretical and Empirical Inquiries. Essays in Honor of John C. Pock.* Boulder, CO: Westview Press.

Mare, Robert D. 1997. "Differential Fertility, Intergenerational Educational Mobility, and Racial Inequality." *Social Science Research*. 26: 263-91.

Mare, Robert D., and Vida Maralani. 2006. "The Intergenerational Effects of Changes in Women's Educational Attainments." *American Sociological Review*. 71: 542-64.

Musick, Kelly, and Robert D. Mare. 2004. "Family Structure, Intergenerational Mobility, and the Reproduction of Poverty: Evidence for Increasing Polarization?" *Demography*. 41: 629-48.

Wang Feng, James Lee and Cameron Campbell. 1995. "Marital fertility control among the Qing nobility: implications for two types of preventive check." *Population Studies.* 49(3), 383-400.

Wolf, Arthur P. 1995. *Sexual Attraction and Childhood Association: A Chinese Brief for Edward Westermarck.* Stanford: Stanford University Press.

Wolf, Arthur P. and Chieh-shang Huang. 1980. *Marriage and Adoption in China, 1845-1945.* Stanford: Stanford University Press.

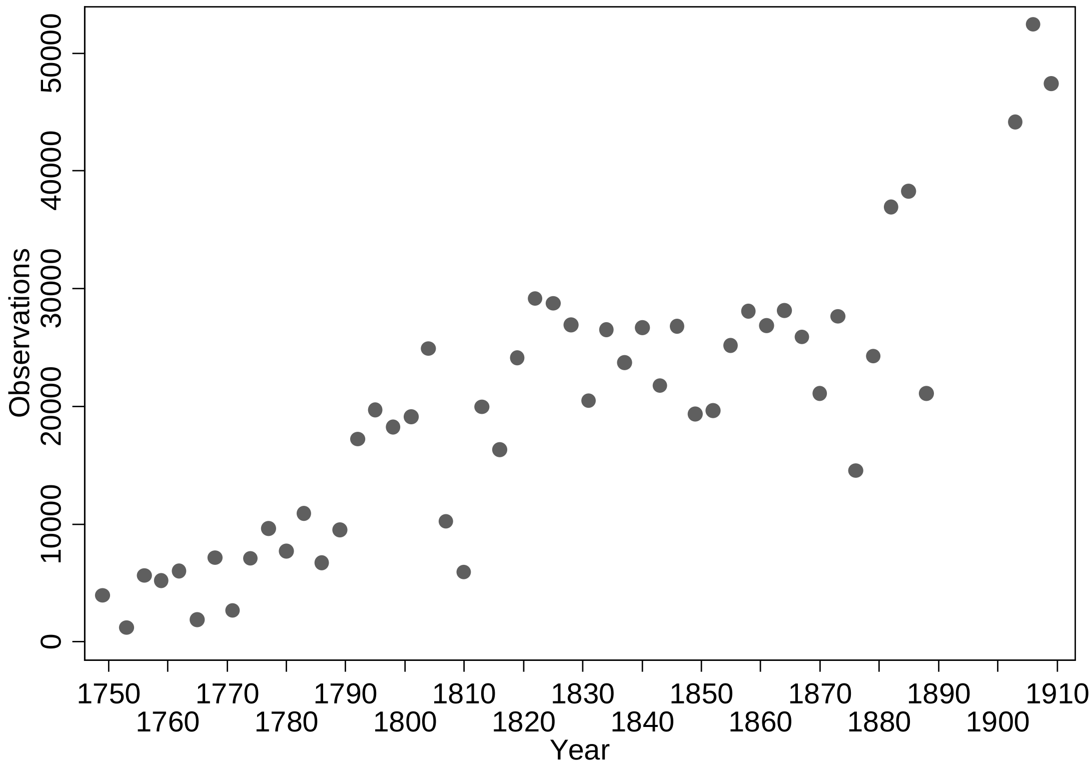


Figure 1 Number of Observations by Year, Liaoning Household Registers, 1749-1909

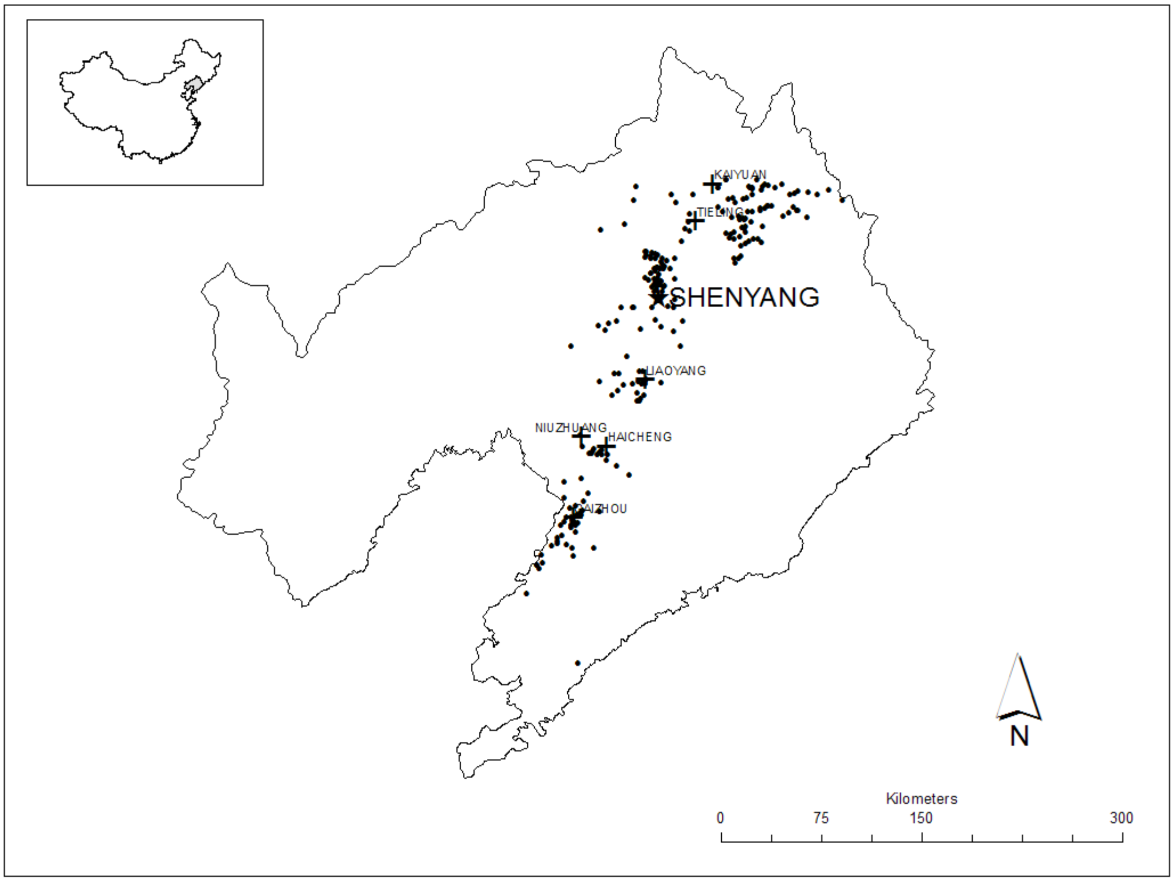
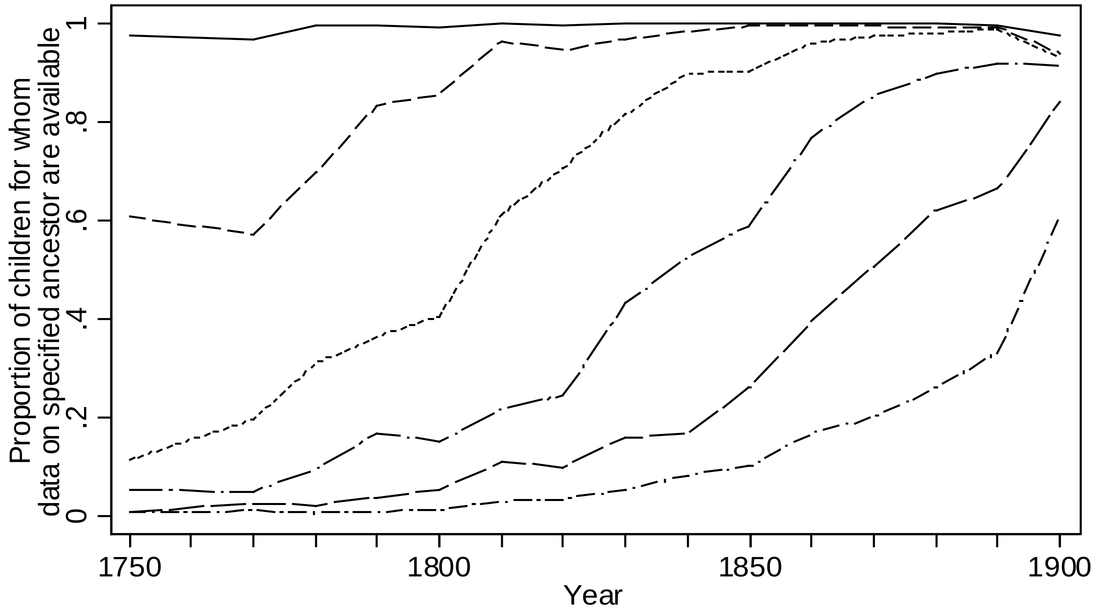


Figure 2 Communities Covered in the Liaoning Household Registers, 1749-1909



[redacted]

Figure 3 Proportions of Boys For Whom Specified Ancestors Can Be Traced, 1749-1909

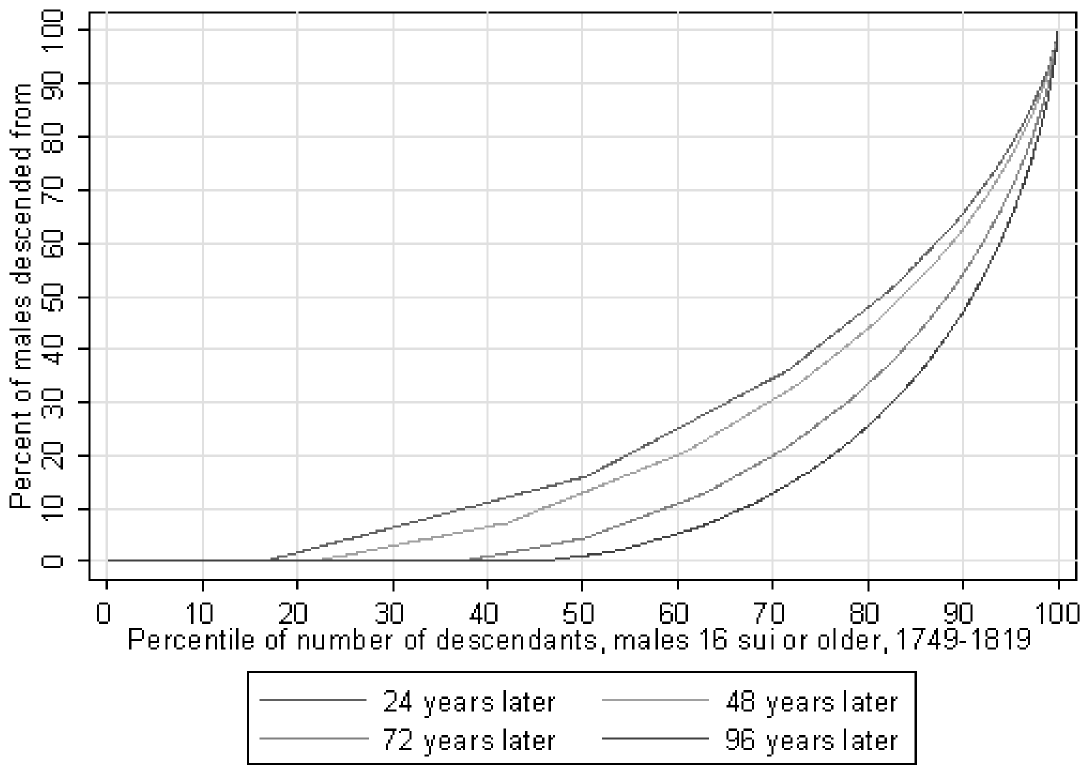


Figure 4 Lorentz Curves for Shares of Males 24, 48, 72, 96 Accounted for by Living Males

Table 1. Men by Numbers of Descendants 24, 48, 72 and 96 Years Later, Liaoning, 1749-1909

Number of Descendants	24 years	48 years	72 years	96 years
0	16.37	21.66	37.07	45.81
1	34	20.35	12.89	8.34
2	21.18	18.79	12.72	8.4
3	11.33	11.76	8.56	6.16
4	6.52	8.39	6.69	5.31
5	3.79	5.57	4.73	4.14
6	2.08	3.43	3.37	2.98
7	1.46	2.49	2.78	2.69
8	1.04	1.97	2.05	2.16
9	0.68	1.32	1.66	1.94
10	1.17	2.92	4.46	5.78
15	0.34	1.18	2.32	4.13
25	0.02	0.17	0.66	1.89
50	0	0	0.04	0.25
N	53588	59737	66250	45742

Table 2. Standard Deviations of Intercepts for Villages, Descent Groups, and Households from Random Effects Models, Liaoning, 1789-1909

	Village		Descent Group		Household			
	S.D. of Intercept	S.E.	S.D. of Intercept	S.E.	S.D. of Intercept	S.E.	N	
Age-specific marital fertility								
Male	0.17	0.02	0.13	0.01	0.09	0.02	71308	
Female	0.18	0.02	0.15	0.01	0.09	0.02	76223	
Male completed fertility								
By ages 41-45 sui	0.19	0.02	0.14	0.02	0.32	0.01	22415	
By ages 51-55 sui	0.17	0.02	0.13	0.02	0.29	0.01	14875	
Mortality								
Males								
Age 1-15	0.43	0.04	0.27	0.04	0.51	0.04	76405	
Age 16-55	0.25	0.03	0.23	0.02	0.36	0.03	102257	
Age 56-80	0.20	0.02	0.17	0.03	0.00	0.04	28776	
All ages	0.22	0.02	0.23	0.02	0.36	0.01	151677	*
Females								
Age 16-55	0.29	0.03	0.21	0.02	0.33	0.03	89022	*
Age 56-80	0.21	0.03	0.17	0.03	0.00	0.03	28114	*
All ages	0.23	0.02	0.21	0.02	0.24	0.02	99854	*
Socioeconomic Outcomes								
Male Marriage								
11-15 sui	0.52	0.06	0.36	0.06	0.88	0.05	58673	
16-20 sui	0.22	0.03	0.32	0.03	0.76	0.02	54396	
21-25 sui	0.27	0.03	0.44	0.03	0.91	0.02	51548	
26-30 sui	0.29	0.04	0.43	0.03	0.98	0.03	48651	
31-40 sui	0.32	0.04	0.51	0.03	1.06	0.03	56985	
Official Position	0.19	0.03	0.52	0.10	8.01	0.24	74554	*
Diminutive Name								
1-10 sui	0.37	0.03	0.37	0.02	0.65	0.01	81962.00	
31-50 sui	0.39	0.08	0.87	0.05	0.99	0.05	40645	

* = Provisional result.

Table 3. Distributions of Adult Males Now And Their Descendants 96 Years Later by Current Status, Liaoning, 1749-1909

	Now	96 Years Later
Income from Official Position		
No Position	95.83	95.08
1-12 taels	0.74	0.68
24-35 taels	3.21	3.82
36+ taels	0.22	0.42
Zuzhang	13.03	16.97
Disabled adult male	25.51	26.69
Diminutive name	16.36	14.26
Eldest brother	56.34	58.08
N	20988	85985

Table 4. Negative Binomial Regression of Number of Male Descendants 48 and 96 years Later, Liaoning, 1749-1909

					w/ Control for Current Number of Boys			
	48 years		96 years		48 years		96 years	
	Coefficient	p	Coefficient	p	Coefficient	p	Coefficient	p
Region (*Ref: North*)								
Central	0.139	0.00	-0.124	0.00	0.088	0.00	-0.202	0.00
South Central	0.204	0.00	0.244	0.00	0.205	0.00	0.252	0.00
South Central	0.187	0.00	0.586	0.00	0.237	0.00	0.634	0.00
Age	0.013	0.00	0.016	0.00	0.002	0.02	0.005	0.00
Year of Birth	0.002	0.00	0.009	0.00	0.003	0.00	0.011	0.00
Income from Official Position (*Ref: None*)								
1-12 taels	0.137	0.12	-0.002	0.99	0.028	0.73	-0.006	0.97
24-35 taels	0.352	0.00	0.255	0.00	0.165	0.00	0.110	0.13
36+ taels	0.553	0.00	0.733	0.01	0.267	0.05	0.612	0.02
Zuzhang	0.093	0.00	0.264	0.00	-0.005	0.83	0.180	0.00
Disabled	-0.061	0.00	0.014	0.66	-0.111	0.00	-0.031	0.29
Diminutive Name	-0.156	0.00	-0.124	0.00	-0.110	0.00	-0.080	0.02
Eldest brother	0.007	0.66	0.027	0.31	-0.020	0.17	0.008	0.75
Number of Boys					0.343	0.00	0.309	0.00
Constant	-1.490	0.00	-6.588	0.00	-2.189	0.00	-7.458	0.00
Ln(alpha)	-0.076		1.191		-0.352		1.127	
N	22262		20988		22262		20988	

Table 5. Logistic Regression of Being in the 85th Percentile or Higher of Number of Surviving Male Descendants 48 and 96 Years Later, 1749-1909

					w/ Control for Current Number of Boys			
	48 years		96 years		48 years		96 years	
	Coefficient	p	Coefficient	p	Coefficient	p	Coefficient	p
Region (*Ref: North*)								
Central	0.285	0.00	-0.018	0.74	0.188	0.00	-0.096	0.08
South Central	0.341	0.00	0.389	0.00	0.394	0.00	0.393	0.00
South Central	0.254	0.00	0.799	0.00	0.417	0.00	0.903	0.00
Age	0.024	0.00	0.021	0.00	0.004	0.06	0.007	0.00
Year of Birth	0.001	0.65	0.013	0.00	0.002	0.19	0.015	0.00
Income from Official Position (*Ref: None*)								
1-12 taels	0.156	0.44	0.029	0.90	-0.058	0.80	-0.142	0.56
24-35 taels	0.783	0.00	0.375	0.00	0.469	0.00	0.134	0.21
36+ taels	1.075	0.00	1.007	0.00	0.567	0.08	0.842	0.01
Zuzhang	0.170	0.00	0.312	0.00	-0.018	0.76	0.226	0.00
Disabled	-0.015	0.74	0.023	0.61	-0.109	0.03	-0.027	0.56
Diminutive Name	-0.396	0.00	-0.148	0.01	-0.306	0.00	-0.074	0.18
Eldest	-0.029	0.46	-0.009	0.82	-0.076	0.07	-0.021	0.60
Number of Boys					0.732	0.00	0.415	0.00
Constant	-3.514	0.00	-12.804	0.00	-4.687	0.00	-13.760	0.00
N	22262		20988		22262		20988	

Table 6. Logistic Regression of Having Any Surviving Male Descendants 48 and 96 Years Later, 1749-1909

					w/ Control for Current Number of Boys			
	48 years		96 years		48 years		96 years	
	Coefficient	p	Coefficient	p	Coefficient	p	Coefficient	p
Region (*Ref: North*)								
Central	0.093	0.02	-0.374	0.00	0.051	0.22	-0.421	0.00
South Central	0.409	0.00	0.042	0.30	0.428	0.00	0.038	0.36
South Central	0.521	0.00	0.833	0.00	0.591	0.00	0.889	0.00
Age	0.000	0.83	0.001	0.55	-0.010	0.00	-0.008	0.00
Year of Birth	0.011	0.00	0.005	0.00	0.012	0.00	0.006	0.00
Income from Official Position (*Ref: None*)								
1-12 taels	0.150	0.44	-0.177	0.29	0.052	0.79	-0.260	0.13
24-35 taels	0.232	0.01	0.233	0.00	0.036	0.71	0.094	0.26
36+ taels	0.747	0.09	1.164	0.00	0.393	0.38	1.052	0.00
Zuzhang	0.234	0.00	0.407	0.00	0.138	0.01	0.353	0.00
Disabled	-0.170	0.00	0.043	0.19	-0.217	0.00	0.017	0.62
Diminutive Name	-0.175	0.00	-0.150	0.00	-0.129	0.00	-0.115	0.00
Eldest	0.028	0.39	0.052	0.07	0.010	0.75	0.046	0.11
Number of Boys					0.424	0.00	0.278	0.00
Constant	-7.037	0.00	-3.747	0.00	-8.098	0.00	-4.191	0.00
N	22262		20988		22262		20988	