Economics Working Paper Series

2017/001

The transmission of health across 7 generations in China, 1789-1906

Jean-François Maystadt and Giuseppe Migali

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

LUMS home page: http://www.lancaster.ac.uk/lums/

The transmission of health across 7 generations in China, 1789-1906*

Jean-François Maystadt†

Giuseppe Migali‡

December 23, 2016

Abstract

We study the intergenerational transmission of health using linked registered data from China between 1789 and 1906. We first document the intergenerational correlations across 7 generations. We then identify intergenerational causal associations comparing children born from twin mothers or fathers. In particular, we find a strong and persistent intergenerational elasticity between mothers and children of about 0.52. The intergenerational association from fathers is much weaker and seems to be largely driven by genetic factors. The estimates remain relatively stable up to generation 5 and are robust to different checks. Overall, our results highlight the nurturing role of women in explaining the intergenerational transmission of health, stressing the key role played by women in affecting children's health outcomes in developing countries.

JEL Classification: I14, I29, I3

Keywords: Intergenerational correlations, causal effects, long-term health outcomes.

*We thank participants in the International Workshop on Applied Economics of Education (Catanzaro, Italy), the Economics seminar at the University of Ferrara (Italy), the LICOS 25 years workshop at KU Leuven (Belgium) and the first Dondena Workshop on Public Policy at Bocconi University (Italy) for their comments and suggestions.

†j.maystadt@lancaster.ac.uk, Department of Economics, Lancaster University Management School, Lancaster University. LICOS, KU Leuven, Belgium

‡corresponding author: g.migali@lancaster.ac.uk, Department of Economics, Lancaster University Management School, Bailrigg Lancaster LA1 4YX, UK. Dipartimento S.G.S.E.S., Universita' Magna Graecia, Catanzaro, Italy.

1 Introduction

Equality of opportunities is often declared as a societal objective, however the extent to which individuals have the opportunity to fulfil their aspirations largely depends on the ability to overcome intergenerational constraints. Economists have long been interested in measuring the intergenerational transmission (IGT) of socio-economic outcomes (Solon, 1999; Black and Devereux, 2010). A key policy interest is to distinguish the role of the environment a child is growing in ("nurture") from the genetic transmission of parents' characteristics ("nature"). Nurture calls for targeted actions aiming at strengthening the initial endowments of individuals in terms of physical and human capital. Nature may render such interventions ineffective and potentially pave the way for policies (e.g. enhancing mobility, ...) aiming at increasing genetic diversity in the population.[1]

In this paper, we estimate the intergenerational elasticity of life expectancy between parents and children, using linked registered data from rural China between 1789 and 1906. We do find strong intergenerational transmission in health between parents and their sons, in particular for mothers. However, such elasticities are likely to be driven by unobserved genetic characteristics. To draw causal inference, we compare the outcomes of children born from parents who are same-sex twins. Interestingly, with parental and individual controls, the intergenerational transmission of health between mothers' and children's lifespans goes down to about 0.52. The equivalent association between fathers' and children's lifespans stands at 0.20, suggesting that much of the IGT was driven by genetic factors. In other words, nurture matters much more to explain the IGT between mothers and their sons, not so much for fathers. Although less relevant for contemporaneous China, our analysis helps us to understand the nature of the intergenerational elasticity in health in a highly patrilineal society, as could be found in many developing countries today. The role of mothers in explaining the nurture component of the IGT is consistent with the important role of women in within-household allocation and its consequences for children's long-term health (Duflo, 2012).

Our contribution lies at the crossroad of three strands of literature. First, while the

[1]The role of genetic diversity in economic development has been documented by Ashraf and Galor (2013) and is subject to a fierce debate in social sciences (Rosenberg and Kang, 2015).

economic literature in the IGT has focused on earnings, education or welfare dependence (Solon, 1992; Holmlund et al., 2011; Chetty et al., 2014)[2], there is still limited evidence on the intergenerational transmission of health outcomes. Existing studies establishing a positive intergenerational association in a variety of health outcomes are Currie and Moretti (2003), Classen (2009), Royer (2009), Bhalotra and Rawlings (2011), and Parman (2012). Most of these studies focus on weight, a relatively short-term and more volatile health outcome. An exception is Bhalotra and Rawlings (2011) who investigate the association between mothers' health and children's anthropometric measurements, together with neonatal, infant and under-five mortality for 38 developing countries. Beyond the scope of their analysis, a major difference with our study is that they do not attempt to disentangle nurture from nature and focus on the associations between two generations. That is also the case for most of the other studies using weight as a health outcome.[3]

We focus on a long-term health outcome, life expectancy, proxied by the lifespan, i.e. the approximated number of years between birth and death. As pointed by Parman (2012), such a proxy for long-term health outcomes has the major advantage to receive a common interpretation across contexts, time and gender. Moreover, the use of lifespan avoids the standard estimation problems of "lifecycle bias", zero income, and non-linearity encountered in the literature on income mobility across generations (Black and Devereux, 2010). Historical studies using earnings data are often restricted to investigate the IGT between fathers and sons. Mothers have long been neglected. Chadwick and Solon (2002) is a noticeable exception for the U.S. but the use of earnings overlooks the importance of within-household allocation and assortative mating in affecting women's welfare (Duflo 2012). Investigating the intergenerational correlation of education in the U.S., Sweden and Norway, Behrman and Rosenzweig (2002), Holmlund et al. (2011), and Pronzato (2012) find stronger correlations between fathers' education and children's education than the ones between mothers and children. Given our focus on life expectancy, Parman (2012) is certainly the closest to our work. He estimates the intergenerational correlation in lifespans in North Carolina and

[2]There is a large literature in other social sciences assessing the association between parents' and children's IQ scores using the so-called Behavioral Genetics Model (Herrnstein and Murray, 1994).

[3]We exploit a parent-twin approach similar to Currie and Moretti (2007) and Royer (2009) among those studies looking at the intergenerational transmission of health.

finds very strong correlations between daughters and mothers and between sons and fathers. In particular for sons, the intergenerational elasticities for fathers and mothers stand at about 0.36 and 0.16, respectively. In our study, we do find stronger elasticities for mothers. Furthermore, we do find that the association with fathers' lifespan is weaker when the specification is augmented with grand-father or father twin fixed effects, which are more likely to control for unobserved heritable traits ("nature"). The association with the mothers' lifespan remains strong with a coefficient at about 0.52.

Our results also echo the biodemographic literature that has exploited historical data on lifespan to assess the inheritance of human longevity. Gavrilov and Gavrilova (2001) provide an extensive review. The estimates of heritability in historical studies vary widely between zero and 0.89 but the majority points to an estimate of less than 0.3 (You et al., 2010). However, most of these studies do not seek to identify the intergenerational transmission of lifespan for a representative sample since they focus on a specific population.[4] A few exceptions exist but they tend to investigate the issue for very small areas.[5] Gudmundsson et al. (2000) and Mazan and Gagnon (2007) propose more general studies by focusing on the majority of the population in Iceland and Quebec, respectively. Some of these studies also assess the correlations between siblings (Swedlund et al., 1983; Perls et al., 2002; Kerber et al., 2001; Mazan and Gagnon, 2007; Salaris et al., 2013) or twins (McGue et al., 1993; Herskind et al., 2006; Yashin and Iachine, 1997). Evidence is also mixed with respect to the relative importance of the maternal or paternal lines of inheritance of human longevity (You et al., 2010). Therefore, we also contribute to the biodemographic literature. With the use of twin-parent fixed effects, we offer a more credible identification strategy to distinguish nature from nature and shed light on the maternal and paternal components of the nurturing effect.

Second, the study of intergenerational associations between parents and children is very relevant in developing countries, where imperfect credit and labor markets limit the ability to escape poverty traps across generations. While the literature has reached a relative consensus

[4]e.g. the Landed Gentry (Beeton and Pearson, 1899), some aristocratic families (Gavrilov and Gavrilova, 2001), centenarians in New England (Perls et al., 2002), Moormons in Utah (Kerber et al., 2001).

[5]e.g. the Connecticut Valley (Swedlund et al., 1983), the French Jura (Cournil et al., 2000), a Flemish village in Belgium (Matthijs et al., 2002), or a village in Sardigna (Salaris et al., 2013).

of an income intergenerational elasticity between 0.3 and 0.45 (Solon, 1999; Chetty et al., 2014), little is known about the magnitude of such correlation in developing countries. The gender dimension is particularly interesting to study, given the potential role of mothers in nurturing children (Duflo 2012). There is an emerging literature investigating economic mobility in India, Malaysia, Mexico, Nepal and Vietnam (Lillard and Willis, 1997; Binder and Woodruff, 2002; Emran and Shilpi, 2015; Hnatkovska et al., 2013; Azam, 2015, e.g.). As far as we know, only a few recent papers focus on the intergenerational transmission of health in developing countries (Bhalotra and Rawlings, 2011, 2013; Eriksson et al., 2014). As explained above with Bhalotra and Rawlings (2011), one major difference with the existing literature is that we seek to draw causal inference and to distinguish between nurture and nature, using a parent-twin approach. Eriksson et al. (2014) use the age and gender adjusted average health measures in the parent's province as an instrumental variable, to assess the transmission of healths across two generations in China. Compared to their work, we prefer the use of parent-twin fixed effects less requiring in terms of identifying assumptions. Related studies on developed countries have also focused on the IGT between two generations, the so-called AR(1) model (Lindahl et al., 2015; Clark and Cummins, 2015). Well, other studies have shown that estimating the IGT between generations may underestimate the persistence of socio-economic outcomes across generations (Long and Ferrie, 2013; Lindahl et al., 2015; Clark and Cummins, 2015). One of the strengths of our analysis is to exploit up to 7 successive generations to assess the relevance of the AR(1) model and the stability of the IGT across generations. We therefore also contribute to the literature seeking to exploit more than 2 generations since the three aforementioned historical studies on the US, the UK, and Sweden are restricted to 3 or 4 generations, and overlook the role of women given the use of earnings (or education) data.

Third, we contribute to a booming literature on the use of parent twins to distinguish nature from nurture. Behrman (2016) proposes an extensive review of the use of twins in economics. Behrman and Rozensweig (2002) seminally investigate the intergenerational transmission of education using twin mother fixed effects. They find a negative association between mother's and children's education, contrasting with the positive association for father's schooling. The interpretation of the authors is that educated women are more likely

to participate to the labor markets, with detrimental impact on the children's schooling. These results have been widely discussed, questioning the negative association for mothers (Antonovics and Goldberger, 2005). So far, there is a relative consensus in existing twin studies on a much larger impact in the intergenerational transmission of education from fathers compared to mothers (Black and Devereux 2010). However, little is known about the generalization of these results to developing countries, in contexts where women empowerment has been associated with improved education among children and economic development in other contexts (Behrman et al., 1999; Behrman and Rosenzweig, 2002; Duflo, 2012).

2 Background and data

We investigate the intergenerational transmission of health in the Liaoning province between 1789 and 1906. The Liaoning province is located in North-East China (Figure 1) and was the original home of the Manchu Qing dynasty emperors (1644-1912). The Qing dynasty was the last dynasty of Imperial China. Most of the eighteen century was seen as a period of political stability and economic expansion (Wong, 1997; Pomeranz, 2000; Meng Xue and Koyama, 2016) but signs of economic decline already emerged at the end of the eighteenth century under the Qianlong Emperor (1735-1796). The Qing ruling was then hardly challenged during the nineteenth century, which is the focus of the present study. Following the First Opium War and the Treaty of Nanking (1842), political instability exacerbated and materialized in a series of popular uprisings including the Taiping rebellion (1850-1864) and the Dungan Revolt (1862-1872). Such revolts were largely driven by opposition to an autocratic and state-controlled regime, or what Chesneaux (1973) names a "bureaucratic feudalism". "In China the state was all-powerful and the peasant was as much exploited by the public demands of the state and bureaucracy as he was by the individual greed of the landlord" (Chesneaux, 1973, 11). Paradoxically, such a bureaucracy provided a wealth of administrative data to be exploited in social sciences. We, indeed, use the China Multi-Generational Panel Dataset-Liaoning (CMGD-LN) that directly relies on population records that have been directly transcribed from the so-called Eight Banner population registers preserved in the Liaoning Provincial Archives (Lee and Campbell, 2010). The "Eight Banner" –

originally the military arm of the Manchus – was a civil and military administrative system organized by the Qing dynasty.

Being an extremely important feature of late imperial China, such registers document the demographic, economic, and social life of the population during that period in great details (Lee and Campbell, 2010). The data comprises triennial data from 29 sets of household registers with 1.5 million records of approximately 260,000 unique individuals from the Qing Imperial Household Office, between 1749 and 1909 in the Liaoning province (Lee et al., 2010; Lee and Campbell, 2011; Song et al., 2015). Missing data are recorded between 1888 and 1903 since the corresponding registers were destroyed by fire (Lee and Campbell, 2010). We only use the sample of males linked to their male and female ancestors across seven generations. We omit daughters since few female births are recorded and married or widowed women could not be traced back to their natal households (Song et al., 2015; Dong et al., 2015).

Our main information of interest consists in life expectancy, proxied by the individual lifespan. We construct the lifespan of each individual as the difference between the last year observed in the register and the year of entry into the dataset at early age. The major drawback of such a definition is the omission of children dying in infancy and in early childhood (potentially before 3 years old since registry takes place every 3 years) and the possibility of unrealistic high lifespans on the other end of the distribution. To reduce the second problem, we exclude from our analysis those being registered as unauthorised migrants (so called 'Tao') since their records seem too poor to be included in mortality analysis (Lee et al., 2012). We also restrict the lifespan below 76. We impose this restriction since mortality record has been recognized as problematic for age above 75 (Dong and Campbell, 2014). In Section 4, we will discuss several robustness checks with respect to the definition of lifespan and the sample restriction, including using a much more strict definition of lifespan, dealing with possible migration bias (without being formally recorded as migrants), or relaxing sample restrictions.

The data offers other information on individual or household characteristics, such as birth year, sex, relationships within the household in each registry, district in which the village of residence is located, migration experiences. Little information is given on socio-economic characteristics of the individuals. The patrilineality in Chinese social organization is well

documented by Song et al. (2015) where sons were seen as more valuable than daughters. Daughters were expected to leave the family after marriage while the sons would remain within the family and provide support to ageing parents. Understanding the role of women in intergenerational transmission of health in this highly patriarchal society is of key interest for contemporaneous economic development.

Originally, the data follow individuals through the registry every three years. However, since our main dependent variable is the lifespan, we transform such a dataset into a cross-section of more than 160,000 individuals that belong to 1062 distinct kin groups. The data originate from 13 districts. We further restrict the sample by excluding individuals that enter into the registers prior to 1789 and those who were alive in 1909. The reason is that households registers before 1789 do not identify residential households and do not uniformly distinguish villages (Lee and Campbell, 2011) while lifespan cannot be approximated for those still alive in 1909. The resulting data consists of information on up to seven generations of the same families.

Table 1 shows how 36,511 individuals can be linked to 5,229 fathers and 7,190 mothers, to 2,205 grand-fathers and 2,750 grand-mothers, to 1,132 great-grand-fathers and 340 great-grand-mothers, to 552 great-great-grand-fathers, and finally, to 295 great-great-great-grand-fathers.[6] For interpretative purposes, we should bear in mind that the (great-) grand-mothers are (grand-) mothers of the fathers and not of the mothers. While the gain in life expectancy of about 7 years between the individuals and the parents is an artefact of the data structure (by construction, an individual cannot die in childhood and becomes a parent), the increasing trend in subsequent positions may seem surprising. However, such a trend may be partly explained by the decreasing life expectancy observed during the period of investigation.[7]

[6]These numbers do not add up to the total number of individuals since the same individual represented as a child at one point in time may well become father, grand-father, or had another family position in subsequent years. Contrary to Lindhal et al. (2008), the family status (father, grand-father, ...) does not necessarily coincide with a particular generation. For instance, an individual recorded as a father can be in any generation of the family tree. Such a distinction is a major advantage to study the stability of the intergenerational transmission of health across generations (see Section 4.3).

[7]Another explanation is that we may oversample short-lived people in more recent registers by excluding individuals still alive in 1909. In Section 4.2, we will assess the robustness of our results to the exclusion of those born from 1830 to give all individuals the opportunity to reach 75 by 1906. We can already note that Panel B of Figure 2 give a similar decline when using a stricter definition of lifespan based on the few individuals regarding which the CMGD-LN indicates they die between two registers. We will further use that stricter definition of lifespan in Section 4.2.

Such a decline during the Qing dynasty in the standard of living has been well documented (Allen et al., 2005). Figure 2 further shows the mid-1800 as a turning point. Section 4.3 will further illustrate whether such a decline in life expectancy is associated with instability in the intergenerational transmission of health.

3 Methodology

3.1 Intergenerational correlations

To motivate our analysis beyond two generations, we will first adopt various specifications of the following model (Lindahl et al., 2015; Clark and Cummins, 2015):

$$LS_{it-l} = \alpha + \theta_t + \mu_d + \beta_{1,j} LS^m_{it-j} + \beta_{2,j} LS^f_{it-j} + \gamma \mathbf{X_i} + \epsilon_i \quad (1)$$

where $l = 0 \dots 4$ and $j = 1 \dots 5$. With $l = 0$ and $j = 1$, the main coefficients of interest $\beta_{1,1}$ and $\beta_{2,1}$ capture the elasticity between the (log) lifespan of individual i, LS_{it} and the one of his or her mother, LS^m_{it-1} and father, LS^f_{it-1}, respectively. With $l = 0$ and $j > 1$, we can assess the direct effects of grand-fathers (LS^f_{it-2}), up to great-great-great-grand-fathers (LS^f_{it-5}). Due to data constrains on linked female members of the family, data are only available up to great-grand-mothers (LS^m_{it-3}). That basically means that we will only be able to control for assortative mating up to that level. The direct effects for great-great-grand-fathers and great-great-great-grand-fathers are likely to include the indirect effect of their partners. In all our specifications, standard errors are clustered at the descent (family tree) level. We also introduce a time indicator to capture cohort- or registry-specific effects, θ_t. We also control for unobserved heterogeneity at district level with district fixed effects, μ_d, and observed heterogeneity at the individual level with control variables, X_i. We include individual characteristics such as the fact to be disabled, to be a migrant during the course of the individual's life, the birth order, the size of the household, the number of brothers and sisters at the approximate time of birth, and the occurrence of natural disasters in the year before birth and during childhood (first 10 years of life). To render our estimated coefficients comparable across specifications, we restrict the sample to those individuals for which the

lifespan is available for mothers, fathers, grand-mothers and grand-fathers. Such a restriction is relaxed when further generations are added to preserve the size of our analytical sample.

To assess the relevance of the AR(1) model, we follow Lindhal et al. (2015) by comparing the direct effects obtained in equation (1) to estimates of the intergenerational correlation across two consecutive generations (i.e $j - l = 1$). We therefore compute the predictions of intergenerational transmission of health between a child and his ancestors, up to five generations apart based on the AR(1) estimates from consecutive generations. Similar to Lindhal et al. (2015), the associated standard errors are obtained using the Delta method. The difference between the direct effects and the respective predictions will give a sense of the direct impact of ancestors on the lifespan of the child. Such comparison cannot distinguish between the nurture and nature channels of transmission, but will help us to explore the extent to which we may underestimate the intergenerational elasticity by focusing on only two generations. The recent literature has indeed shown that estimates obtained from two generations severely underestimate the long-run intergenerational persistence in socio-economic outcomes across generations (Long and Ferrie, 2013; Lindahl et al., 2015; Clark and Cummins, 2015). Based on this first empirical exploration, we will then further compare the AR(1) specification with more complex models, from AR(2) up to AR(5) of the following form:

$$LS_{it} = \alpha + \theta_t + \mu_d + \sum_{j}^{3} \gamma_{1,j} LS^m_{it-j} + \sum_{j}^{5} \gamma_{2,j} LS^f_{it-j} + \gamma \mathbf{X_i} + \epsilon_i \tag{2}$$

The same description of variables and fixed effects apply, while standard errors are also clustered at the descent level. The same sample restrictions apply.

3.2 The twin approach applied to registry data

Equation (1) can be simplified to $j = 1$ to allow for a causal identification of the nurture channel of the intergenerational transmission of health between parents and their children. The main challenge in identifying a causal relationship indeed consists in distinguishing nurture from nature. The IGT may be largely driven by inherited genetic differences across families. In absence of observed genomic information, the literature has largely relied on the

twin approach, i.e. the comparison of individuals whose fathers or mothers are twins.[8] The idea is that by comparing the IGT among children of twin parents helps to isolate the IGT from fixed family characteristics, and removes (to a large extent) the variation in IGT due to genetic differences. The twin strategy will be contrasted with a more standard approach of controlling for grandfather fixed effects, which similarly to parent-twin fixed effects also remove some unobserved differences in initial endowments. More specifically, the various specifications take the following form:

$$LS_i = \alpha + \theta_t + \mu_d + \varphi_k + \beta_1 LS_i^m + \beta_2 LS_i^f + \gamma \mathbf{X_i} + \delta \mathbf{Z_i} + \epsilon_i \tag{3}$$

LS_i still denotes the (log) lifespan of the individual i, while LS_i^m and LS_i^f, denote the ones of his mother and father, respectively. Standard errors are clustered at the descent (family tree) level. All our specifications also include time (θ_t) and district fixed effects (μ_d), together with individual control variables. The main difference with equation (1) with $j = 1$ is that we introduce additional fixed effects, φ_k. Our preferred specifications will compare children from plausibly same-sex twin parents, through the introduction of mother-twin (for β_1) or father-twin fixed effects (for β_2). The same individual controls, X_i, used in our previous models are also introduced. As mentioned above, we will also compare children sharing the same grand-father, through the introduction of grand-father fixed effects. To render our estimates comparable across specifications, we restrict our sample to those individuals with mother twins.

Our approach relies on two main identifying assumptions. First, siblings born the same year and of the same sex are considered as twins. In our analysis, we exploit about 322 and 596 pairs of father and mother same-sex twins, respectively.[9] A major drawback is

[8] Other approaches would consist in controlling for directly observed genetic differences (Beauchamp et al., 2011) or the use of adoptees (Bjorklund et al., 2006; Sacerdote, 2007). Another approach has used instrumental variables such as great-grandfathers' outcomes of interest (Clark and Cummins, 2015) or unexpected events or reforms. However, the exclusion restriction that needs to be met and the generalisability of results, often driven by local average treatment effects, constitute important drawbacks of such an approach (Black and Devereux, 2010).

[9] Twins are identified based on a common mother. Identification of the twins based on mothers is preferred since identification based on fathers may wrongly infer that one sister is a twin with her sister-in-law if they are born the same year. An indication of that drawback is the fact we identify a greater number of pairs of mother twins using common father (812 instead of 596 with mother identification), while the number of pairs of father twins remains almost identical (320 instead of 322). We nonetheless test the robustness of

that we cannot be sure whether those same-year newborns are monozygotic twins. Well, the causal interpretation of our results depends on the assumption that twin parents are identical in their endowments. Such assumption is naturally more likely with monozygotic twins, compared to dizygotic twins who share on average 50 percent of the genes. We cannot be certain to remove completely the genetic component of the IGT with our parent-twin fixed effects. We will therefore apply the bounding exercise proposed by Holmlund et al. (2008, 58-61). They indeed demonstrate that we can obtain proper identification without having information about monozygotic (MZ) and dizygotic (DZ) twins, under the assumption that the intergenerational association is identical for twins and same-sex siblings. The lower-bound estimate is defined in such a way:

$$\widehat{\beta_{TS}} = \frac{\widehat{\beta_{TW}} - \lambda\theta\widehat{\beta_{SIB}}}{1 - \lambda\theta} \tag{4}$$

where $\widehat{\beta_{TW}}$ and $\widehat{\beta_{SIB}}$ denote the estimates obtained under the parent same-sex twin fixed effects specification and under the parent same-sex siblings fixed effects specification, respectively. θ represents the share of DZ twins among all same-sex twin pairs. Similar to Holmlund et al. (2008), θ is approximately 0.5 in samples, like ours, that do not separate MZ from DZ twins. λ is an indicator for possible treatment differentials between twin and non-twin sibling parents. Under the assumption of absent treatment differentials ($\lambda = 1$) and when the sibling estimate is larger than the twin estimates ($\widehat{\beta_{SIB}} > \widehat{\beta_{TW}}$), $\widehat{\beta_{TS}}$ provides a lower-bound estimate of the causal intergenerational transmission of health between parents and children. $\widehat{\beta_{TS}}$ is nonetheless likely to underestimate the intergenerational transmission of health. Relaxing the assumption of absent treatment differentials ($\lambda < 1$), the intergenerational transmission of health between parents and sons should increase and provides an upper-bound estimates when λ tends to zero. The estimate should then converge towards the parent same-sex siblings fixed effects.

our main results to identifying twins based on common fathers in section 4. One possible caveat of using common mothers to identify twins is that the registers do not specify the identity of an individual's mother. The software used to create the dataset links children to the wife of the household head, giving priority to the link made between a child and a wife in the earliest available register. In cases of remarriage or polygyny, children may be incorrectly assigned to the wife who was observed in the register. However, Lee et al. (2010) indicate that it should be a minor issue since windower remarriage was very uncommon and polygyny was almost non-existent. We actually do not observe any polygynous household in our sample.

The second key identifying assumption is that (parent) twins are treated similarly within their family after birth. Such assumption would not be valid if (grand-)parents make compensating or reinforcing investments or if children's outcomes are affected by both the twin parent and his or her partner. We cannot exclude such unobserved heterogeneity between parent twins. However, we reduce the threat by controlling for parents' exposure to shocks, parental controls and the partner's lifespan. Parental control variables, Z_i, include the father's and mother's fact to be disabled, the age at approximated birth, and the fact to be a migrant during his or her course of life. Parental exposure to shocks include the number of shocks (natural disasters) occurring during their childhood.

3.3 The stability of the intergenerational transmission of health

We will investigate the stability of the parent-child causal transmission of health, across 7 generations within the same family tree. We use a similar specification than in equation (3):

$$LS_i^g = \alpha + \theta_t + \mu_d + \beta_1 LS_i^{m,g-1} + \beta_2 LS_i^{f,g-1} + \gamma \mathbf{X_i} + \delta \mathbf{Z_i} + \epsilon_i \tag{5}$$

The main differences are that we cannot include parent-twin or grand-father fixed effects and estimate the association between the lifespan of parents and sons by pairs of generations. The index g denotes the generation of the sons, from 2 to 7. In such a way, we can assess the stability of the IGT across generations. We should nonetheless be cautious in comparing our results across generations since the sample of our extended families with multiple generations (e.g. those with a seventh generation) may differ from other households for unobserved reasons. As seminally shown in biodemography by Swedlund et al. (1983), it is likely the population of interest will become increasingly homogeneous across generations. We discuss that issue in more details in Section 4.3.

4 Results

4.1 Intergenerational correlations across generations

Table 2 provides the AR(1) estimates of the intergenerational elasticity between two consecutive generations, together with the estimates of the direct impacts of ancestors on the child lifespan. Controlling for assortative mating, the intergenerational correlation for father and mother stand at about 0.30 and 0.51, respectively (columns (1) and (2)). The estimates provided by the AR(1) model are in the same range of what has been found for earnings in the US, the UK or Nordic countries in more contemporaneous times (Solon, 1999; Holmlund et al., 2011; Parman, 2012; Chetty et al., 2014). However, the higher elasticities between sons and mothers, compared to those between sons and fathers contrast very much with the existing literature. For instance, Holmlund et al. (2011) find a higher association between fathers' and children's education (0.25), than the one for mothers (0.20). Compared to a closer literature on the intergenerational transmission of health, Parman (2012) also finds higher elasticities between sons and fathers (0.359) than between sons and mothers (0.157). Our results are nonetheless largely consistent with the strong correlation between mothers' and children's health found in developing countries (Bhalotra and Rawlings, 2011, 2013; Classen, 2009). However, the comparison is limited since those papers do not provide estimates for the patrilineal linkage. One exception is Eriksson et al. (2014) who find higher intergenerational correlations for fathers (0.298) in rural China between 1991 and 2009, although the difference is relatively small with the one for mothers (0.272).

The rest of Table 2 reveals a few interesting results. First, we do not find any direct effect from grand-father (column 3), great-grand-father (column 5), and great-great-grand-father (column 7). We are therefore unlikely to underestimate the intergenerational transmission of health between fathers and sons in case we use the AR(1) model. On the contrary, column (4) reveals that there is a statistically significant estimate of the association between grand mothers' lifespan and that of grandchild. The associated elasticity of 0.63 is much larger than the predicted association of 0.23 (0.515×0.463) based on the correlations between mothers and sons (0.515) and mothers and grand-mothers (0.463). The direct effect does not seem to be at play when investigating the direct role of great-grand-mother (column

6). Overall, the AR(1) model seems to be appropriate for fathers, whereas, given the direct role of grand-mothers, it is likely to underestimate the intergenerational correlation of health for mothers across generations. Although we cannot control for assortative mating due to sample size constraints in columns (5) to (7), that should not alter our main conclusions since neglecting assortative mating usually tends to inflate the estimated coefficients.[10]

The direct effect played by grand-mothers is further confirmed when we extend the analysis by implementing AR(2), AR(3), and AR(4) models in Table 3. The estimated coefficients for grand-mothers stand in a range between 0.37 (column 2) and 0.57 (column 4). The fact that the estimated coefficient is not significant at any reasonable level of confidence in the AR(5) is certainly due to the reduction of sample size to 139 observations.

Overall, Table 2 and Table 3 provide suggestive evidence that women matter much more than men for the intergenerational transmission of health in rural China between 1789 and 1906. In particular, mothers and grand-mothers seem to bear an important direct role on children's well-being. Of particular interest is the fact that grand-mothers are not the mothers of the mothers but their mothers-in-law. Such distinction suggests that women may be decisive in explaining the intergenerational transmission of health through a nurture channel. Nonetheless, such evidence does not allow us to claim any causal relationships since inherited genetic differences may inflate the intergenerational transmission of health. However, it is very informative in shedding light on the possible cost associated with limiting our analysis to an AR(1) model. Even with this upward endogeneity bias, we know that the AR(1) model is not underestimating the intergenerational elasticity between fathers and sons but may well underestimate that between mothers and sons.

4.2 Nurture versus nature in the intergenerational transmission of health

To distinguish nurture from nature, we restrict our analysis to the association in health between children and their parents. In Table 4, we report the estimated coefficients β_1 and

[10]Introducing great-grand-mothers in the estimation corresponding to column (5) and great-great-grand-mothers in the one corresponding to column (7) would reduce the sample size to 293 and 22 observations, respectively.

β_2 of equation (3) using several specifications augmented with a large set of fixed effects, more likely to isolate the effect of nurture from the role of nature.[11] In Panels A and B, we introduce the lifespans of mothers and fathers, separately. To account for assortative mating, Panel C includes both variables of interest. Column (1) of Panels A and B only includes district and cohort fixed effects, and indicates an intergenerational transmission of health of 0.92 and 0.67 for mothers and fathers, respectively. When we turn to the equivalent estimations in Panel C, where the effect of assortative mating is controlled for, the coefficient of the lifespan of the father is further reduced to 0.34, while the intergenerational association between mothers and sons stands at about 0.80.[12]

Assortative mating clearly matters. We further interpret the results in Panel C of Table 4. When the associations parents-children in health are compared among individuals sharing the same grand-father, in column (2), the magnitude of the coefficient for the lifespan of the mother and the father is reduced to 0.53 and 0.15, respectively. The coefficient of 0.15 associated with the lifespan of the father becomes statistically insignificant. In columns (3) and (4), we show the results obtained implementing our preferred identification strategy of the intergenerational causal effects. We estimate more requiring models introducing mothers and fathers' twin fixed effects. Such estimations are more likely to isolate the effect of nurture. Comparing children whose mothers share the same genetic background (column 3), the intergenerational elasticity between mothers and sons stands at 0.57. In column (4), the coefficient of 0.174 obtained for the father lifespan is statically different from zero but its magnitude is only about one third of the one of mothers. Mothers seem to matter much more in the nurturing transmission of health in our study. In column (5) and (6), such interpretation receives further support when we control for individual characteristics, including our proxy for pre-natal shock. As expected given the potential influence of pre-natal environment, the estimate of the intergenerational transmission of health between mothers and sons is reduced to around 0.52. In theory, such estimation should capture the nurture

[11] We provide the descriptive statistics of our control variables and the detailed results of Panel C of Table 4 in Tables A1 and A2 in the Appendix.

[12] The coefficients equivalent to the same regressions (with district and cohort fixed effects) with the full set of individual and parental controls stand at about 0.75 (0.104) and 0.31 (0.091) for mothers and fathers, respectively. It only differs slightly compared to column (1) of Table 3 since all specifications of Table 4 are restricted to the sample of individuals with twin mothers (except for the models using twin fathers).

effect, net of the pre-natal environmental conditions. On the contrary, the introduction of parent twin fixed effects slightly increases the coefficient for the father's lifespan to 0.19.[13]

The causal interpretation of the estimated coefficients in columns (5) and (6) of Table 4 strongly relies on the assumption that twin parents would be treated in a similar way after birth. The identification strategy indeed rests on the assumption that twin parents differences are uncorrelated with other twins differences. We cannot directly observe investment made by grand-parents into parent twins' human capital. However, columns (7) to (8) of Table 4 show that our conclusions remain qualitatively unchanged when adding parental control variables. The intergenerational association between sons and mothers stands at 0.52 in our preferred specifications with parent-twin fixed effects. A ten percent increase in the mother's lifespan (equivalent to about 3.4 years of life at the mean) would translate into a rise of about 5 percent in life expectancy for her child (about 1.7 years at mean value). The association between fathers and sons seems largely explained by unobserved genetic factors. The association stands at about 0.20 with the father twin fixed effects. The difference in IGT from fathers and from mothers is likely to be a lower-bound, since we have demonstrated in Section 3.1 the likely underestimation of the IGT between mothers and sons in the AR(1) model and the opposite between fathers and sons. Nurture matters much more for the intergenerational transmission of health between mothers and sons, than between fathers and sons.

Similar to the elasticities shown in Section 4.1, our results on the importance of mothers in explaining the nurture component seem to contrast very much with the results found in the literature, in particular regarding the intergenerational transmission of education in developed countries (Behrman and Rosenzweig, 2002; Black and Devereux, 2010; Holmlund et al., 2011; Pronzato, 2012). One explanation might be that mothers have been found to impact greatly children's welfare in developing countries in particular for health, in contrast to education outcomes (Duflo 2012).

As described in Section 3, one remaining concern is that we cannot be certain that

[13]The coefficient of the lifespan of the father would be reduced to the point where it becomes insignificant when twins are identified based on common fathers. The intergenerational transmission of health between mothers and sons would range between 0.39 and 0.43. Table A5 in the Appendix replicates Table 4 based on this alternative identification of twin parents.

our twin parents are effectively twins, and even less monozygotic twins. To deal with this issue, we implement the bounding exercise proposed by Holmlund et al. (2008). Similar to these authors, in Table A3, we only focus on the specification without controls. Results with controls are qualitatively similar. In Panel A of Table A3, we compute the lower-bound estimate of the intergenerational transmission of health between mothers and sons based on equation (4), while the siblings estimate constitutes an upper-bound.[14] In Panel B of Table A3, the sibling estimates of the IGT between fathers and sons is obtained using grand-father fixed effects since grand-fathers in our data are fathers of the fathers.

Assuming absence of treatment differentials between twin and sibling parents ($\lambda = 1$), Panel A of Table A3 indicates the lower-bound of the intergenerational transmission of health between mothers and sons to stand at about 0.54 (column 3). The equivalent bounding exercise implemented with individual controls would give a lower-bound estimate of 0.51 for the intergenerational transmission of health between mothers and sons. We can therefore conclude that had we been able to identify and use only MZ twins, our analysis would have produced a statistically significant causal intergenerational elasticity of at least 0.51. As expected, if we relax the assumption of absent treatment differentials ($\lambda < 1$) in columns (4) to (7), the IGT increases in magnitude and converge towards the father twin fixed effect estimate. A similar bounding exercise provides an estimated IGT between fathers and sons bounded between 0.17 and 0.20 (Panel B of Table A3).

Finally, our main results strongly rely on the way we approximate the lifespan of the individuals recorded in the population registers. We have mainly two concerns. First, we may wrongly infer that someone is dead when he is not observed anymore in a register. However, the coefficient of correlation between the approximated lifespan and the lifespan computed for those we know they die between two registers stands at 0.98, and this is a strong indication that attrition is likely to be a minor issue. The graphical representation in Figure 2 is also almost identical when using that stricter definition.[15] We also investigate the importance of

[14]Such a sibling estimate is equivalent to a maternal grand-mother fixed effects. Mother-twin fixed effects estimates of the association between mothers and sons cannot be directly compared to the grand-father fixed effects since grand-fathers are patrilineal.

[15]In the CMGD-LN, we can use the information according to which the observed individual was annotated in the next available register as having died during the three years covered by that register. The information is only available for about 14% of the analytical sample. When such information is used to construct the lifespan, we obtain similar results in the AR(1) model using the grand-father fixed effects (similar to e.g.

migration. While we control for *observed* migration at the individual and parental levels and exclude unauthorised migrants whose records are recognized as being very poor, we cannot exclude that *unobserved* migration biases our results. To make an educated guess of the likely bias, we first use the characteristics at birth of the *observed* migrants to identify potential migrants (*unobserved*). Using a simple nearest-neighbor matching technique (Heckman et al., 1997), about 7 percent of our analytical sample is matched to *observed* migrants.[16] We then replicate our main results from Table 4 in Table A4, excluding these potential migrants. Our results are largely unchanged and confirm the low level of mobility reported in the historical background of this study.

Second, our results may be biased in favour of shorter-lived persons in more recent years. Since we exclude those who were alive in 1909, the lifespan data for longer-lived people were not yet available by the time of data collection, a well-known problem in biodemography (Gavrilov and Gavrilova, 2001). Nonetheless, our main results do not seem to be a consequence of the exclusion of the individuals still alive in 1909 and the subsequent risk of oversampling shorter-lived individuals for more recent years. We replicate the main results, excluding in a backward fashion those born before 1906 up to 1830. In the most extreme case, when we exclude those born after 1830, we give the same chance to all individuals to reach the maximum lifespan by 1906. Figure 3 shows that the intergenerational transmission of health between mothers and sons remains fairly stable when we exclude cohorts born between 1850 and 1906. When we also exclude those born before 1850 up to 1830, we witness a sharp reduction in the coefficient but still standing above the level of 0.2 in the most extreme case. Such reduction may be partly explained by the reduction by about half in sample size but also by some changes occurring around 1850 in the intergenerational transmission of health. We investigate further that issue in Section 4.3. The intergenerational transmission of health between fathers and sons remains at around 0.2 following the same exercise.

column (2) of Table 4). However, the sample size becomes too small to draw causal inference with the use of mother-twin (467 observations) or father-twin fixed effects (99 observations).

[16]The propensity score is based on the following co-variates: the household size at birth, the number of brothers and sisters at birth, birth order, and earlylife and antenatal shocks.

4.3 Is the intergenerational transmission of health stable overtime?

After establishing the nurturing role of mothers in the intergenerational transmission of health, we investigate the stability of such estimates across generations. To that purpose, we assess equation 5 by pairs of generations. Table 5 suggests that the estimated coefficients remain relatively stable up to generation 5, although with an increasing trend between generations 3 and 5. We should already acknowledge that such an increasing trend cannot be explained with the likely homogenisation of the population that record several successive generations in our sample. However, a strong jump is observed for the association between generations 5 and 6. An increase by about one quarter to one third suggests that the society has witnessed greater equality of opportunities for the generation born on average in 1865 and whose mothers were born on average in 1836. Combined with the fast decline in lifespan (Figure 2) and the confirmation of our main results when cohorts born after 1830 are sequentially excluded (Figure 3), our results point to a possible structural break in the second half of the nineteenth century. Such a structural break is confirmed when estimating all preferred specifications with parent-twin fixed effects, augmented with an interaction term between the lifespan of the mothers and a dummy equal to 1 for all years after 1850. Even if not precisely estimated, the intergenerational transmission of health is found to be on average 50 percent higher after 1850.

The social conditions of the latest period of the Qing dynasty could partly explain such a structural break. Society was organized along distinct social groups. The upper gentry ("Shen-Shih" or "Shen-Chin") derived great power, not only through the ownership of land, but also by the organization of local corps, the administration of justice, the control of the economy (salt monopoly), and the imposition of taxes (Chang, 1967; Chesneaux, 1973). Access to the priviledged elite was highly regulated through a system of examinations and degrees controlled by the government (Chang, 1967). An exception to that regulation is the access to the lower gentry, with the "shen-yuan" and the "chien-sheng" whose titles were acquired by military education. The title of "chien-sheng" could be bought by men of military education (Chang, 1967). Since access to these groups were not hereditary, the

levels of social mobility were high for preindustrial standards (Jiang and Kung 2015). As a result, the share of the upper and lower gentry was relatively stable during most of the Qing dynasty. However, such a relative stability was greatly disturbed during the Taiping war (1850-1864). Chang (1967, p.83) indicates that "Regulations were established providing for the contribution of money to the military fund by local people who were, in turn, to be rewarded with increases in the shen-yuan quota of their native places. Such a regulation first appeared in 1853 when there was an urgent need to increase the public revenue in order to meet the ever increasing expenses of war against the Taipings." As a result, the number of people with a lower gentry title increases by about 50 percent, from an estimated 1,094,734 to about 1,443,900 with a strongly reduced total population. The Taiping rebellions mainly took place in South China and led to the loss of at least 25 million lives (Yu, 2012). Such a change in the size of the gentry seems to have been associated with a strengthening in the intergenerational transmission of health.[17] Our data are too limited to shed further light on the main channels underlying the structural change in the intergenerational transmission of health around 1850. Understanding in depth the nature of such a structural break is certainly a path for further research.

5 Conclusions

Identifying the nurturing effect in the intergenerational transmission of welfare is key to guide policies aiming at promoting equality of opportunity. In this paper, we study the intergenerational transmission of health using linked registered data from China between 1789 and 1906. Our preferred specifications rely on comparing children from the same twin mothers or fathers. We find a strong intergenerational elasticity between mothers and their children, standing at about 0.52. Based on the evaluation of the direct impacts of further ancestors, we know that the related two-generations model is likely to underestimate the

[17] In Hao and Xue (2016), the Taiping rebellion is also associated with structural changes such as the implementation of an education reform, the creation of new schools, or the resulting formalization of local self-governance. However, such changes took place in early century. It was indeed "only with the abolishment of the exam system in 1905 that modern education began to expand" (Hao and Xue, 2016, 6). However, we cannot exclude that the dismantlement of kindship networks described by Hao and Xue (2016) may be associated with a change in the intergenerational transmission of health, even if the channel and the resulting bias are far from obvious to conjecture.

true intergenerational elasticity between women and their descendants. The elasticity of 0.52 is therefore likely to be a lower-bound estimate. On the contrary, the intergenerational association from fathers is much weaker and seems to be largely driven by genetic factors. At best, the intergenerational transmission of health between fathers and sons stands at about 0.20. The weaker nurturing effect on the father side cannot be explained in our analysis by the limit of the two-generations model, since grand-fathers and further ancestors do not seem to have any direct impact on the children's health.

Our results contrast with the existing literature that seems to point to a stronger role played by fathers in the intergenerational transmission of socio-economic outcomes, such as earnings, education or health. A contrast which suggests that the existing literature cannot necessarily be generalized to developing countries, where women play a prominent role in affecting children's welfare, in particular health outcomes. More research, using a twin-parent approach, would be needed to be more affirmative on the subject. It would also be interesting to know whether the higher intergenerational elasticity between women and sons is confirmed in developing countries when a twin approach is applied to other outcomes than health and when the approach is extended to the parents-daughters relationships.

References

Allen, R. C., Bengtsson, T., and Dribe, M. (2005). *Living standards in the past: New perspectives on well-being in Asia and Europe*. Oxford University Press on Demand.

Antonovics, K. L. and Goldberger, A. S. (2005). Does Increasing Women's Schooling Raise the Schooling of the Next Generation? Comment. *American Economic Review*, 95(5):1738–1744.

Ashraf, Q. and Galor, O. (2013). The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development. *American Economic Review*, 103(1):1–46.

Azam, M. (2015). Intergenerational Occupational Mobility among Men in India. *The Journal of Development Studies*, 51(10):1389–1408.

Beauchamp, J. P., Cesarini, D., Johannesson, M., van der Loos, M. J. H. M., Koellinger, P. D., Groenen, P. J. F., Fowler, J. H., Rosenquist, J. N., Thurik, A. R., and Christakis, N. A. (2011). Molecular Genetics and Economics. *Journal of Economic Perspectives*, 25(4):57–82.

Beeton, B. and Pearson, K. (1899). Data for the problem of evolution in man, II: A first study of the inheritance of logevity and the selective death rate in man. *Proceedings of the Royal Society of London*, 65:290–305.

Behrman, J. R. (2016). Twins Studies in Economics. In Komlos, J. and Kelly, I., editors, *Oxford Handbook of Economics and Human Biology*, chapter 19. Elsevier, 1 edition.

Behrman, J. R., Foster, A. D., Rosenzweig, M. R., and Vashishtha, P. (1999). Women's Schooling, Home Teaching, and Economic Growth. *Journal of Political Economy*, 107(4):682–714.

Behrman, J. R. and Rosenzweig, M. R. (2002). Does Increasing Women's Schooling Raise the Schooling of the Next Generation? *American Economic Review*, 92(1):323–334.

Bhalotra, S. and Rawlings, S. (2011). Intergenerational persistence in health in developing countries: The penalty of gender inequality. *Journal of Public Economics*, 95:286–299.

Bhalotra, S. and Rawlings, S. (2013). Gradients of the intergenerational transmission of health in developing countries. *Review of Economcis and Statistics*, 95(2):660–672.

Binder, M. and Woodruff, C. (2002). Inequality and Intergenerational Mobility in Schooling: The Case of Mexico. *Economic Development and Cultural Change*, 50(2):249–267.

Bjorklund, A., Lindhal, M., and Plug, E. (2006). The origins of intergenerational associations: Lessons from Swedish adoption data. *Quarterly Journal of Economics*, 121(3):999–1028.

Black, S. and Devereux, P. J. (2010). Recent developments in intergenerational mobility. Working Paper 15889, NBER.

Chadwick, L. and Solon, G. (2002). Intergenerational Income Mobility Among Daughters. *American Economic Review*, 92(1):335–344.

Chang, C.-L. (1967). *The Chinese Gentry. Studies on Their Role in Nineteenth-Century Chinese Society*. Third edition. Seattle and London: University of Washington Press.

Chesneaux, J. (1973). *Peasants revolts in China*. London: Thames and Hudson.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Clark, G. and Cummins, N. (2015). Intergenerational Wealth Mobility in England 1858 − 2012: Surnames and Social Mobility. *The Economic Journal*, 125(582):61–85.

Classen, T. (2009). Measures of the intergenerational transmission of body mass index between mothers and their children in the United States, 1981 2004. *Economics and Human Biology*, 8:30–43.

Cournil, A., Legay, J.-M., and Schachter, F. (2000). Evidence of Sex-Linked Effects on the Inheritance of Human Longevity: A Population-Based Study in the Valserine Valley (French Jura), , 18-20 Centuries. *Royal Society*, 267(1447):1021–1025.

Currie, J. and Moretti, E. (2003). Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *The Quarterly Journal of Economics*, 118(4):1495–1532.

Dong, H. and Campbell, C. D. (2014). Kindship Matters: Long-term mortality consequences of childhood migration, historical evidence from northeast China, 1792-1909. *Social Science and Medecine*, 119:274–283.

Dong, H., Campbell, C. D., Kurosu, S., Yang, W., and Lee, J. Z. (2015). New Sources for Comparative Social Science: Historical Population Panel Data from East Asia. *Demography*, 52:1061–1088.

Duflo, E. (2012). Women Empowerment and Economic Development. *Journal of Economic Literature*, 50(4):1051–1079.

Emran, M. S. and Shilpi, F. (2015). Gender, Geography, and Generations: Intergenerational Educational Mobility in Post Reform India. *World Development*, 72(C):362–380.

Eriksson, T., Pan, J., and Qin, X. (2014). The intergenerational inequality in China. *China Economic Review*, 31:392–409.

Gavrilov, L. and Gavrilova, N. (2001). Biodemographic Study of Familial Determinants of Human Longevity. *Population: An English Selection*, 13(1):197–222.

Gudmundsson, H., Gudbjartsson, D., Kong, A., Gudbjartsson, H., Frigge, M., Gulcher, J., and Stefansson, K. (2000). Inheritance of human longevity in Iceland. *European Journal of Human Genetics*, 8:743–749.

Hao, Y. and Xue, M. M. (2016). Friends from Afar: The Taiping rebellion, Cultural proximity and primary School in the Lower Yangzi, 1850-1949. *mimeo*.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, 64(4):605–654.

Herrnstein, R. J. and Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. The Free Press.

Herskind, A., McGue, M., Holm, N., Sorensen, T., Harvald, B., and Vaupel, J. (2006). The heritability of human longevity: A population-based study of 2,872 Danish twin pairs born 1870-1900. *Human Genetics*, 97:319–323.

Hnatkovska, V., Lahiri, A., and Paul, S. (2013). Breaking the caste barrier: Intergenerational mobility in india. *Journal of Human Resources*, 48(2):435–473.

Holmlund, H., Lindahl, M., and Plug, E. (2008). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. Discussion Paper 3630, IZA.

Holmlund, H., Lindahl, M., and Plug, E. (2011). The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods. *Journal of Economic Literature*, 49(3):615–651.

Kerber, R., O'Brien, E., Smith, K., and Cawthon, R. (2001). Longevity in Utah Genealogies. *The Journals of Gerontology. Sciences and Medical Sciences*, 56(3):B130–B139.

Lee, Z. L. and Campbell, C. D. (2010). China multi-generational panel dataset, Liaoning (CMGPD), 1749-1909. User Guide. User guide, Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Lee, Z. L. and Campbell, C. D. (2011). China multi-generational panel dataset, Liaoning (CMGPD), 1749-1909. Dataset 2011-09-02, Inter-university Consortium for Political and Social Research.

Lee, Z. L., Campbell, C. D., and Chen, S. (2010). China multi-generational panel dataset, Liaoning (CMGPD), 1749-1909. User guide, Inter-university Consortium for Political and Social Research.

Lee, Z. L., Campbell, C. D., and Chen, S. (2012). China multi-generational panel dataset, Liaoning (CMGPD), 1749-1909. Training guide, Inter-university Consortium for Political and Social Research.

Lillard, L. A. and Willis, R. J. (1997). Motives for intergenerational transfers: Evidence from Malaysia. *Demography*, 34(1):115–134.

Lindahl, M., Palme, M., Massih, S. S., and Sjogren, A. (2015). Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations. *Journal of Human Resources*, 50(1):1–33.

Long, J. and Ferrie, J. (2013). Intergenerational Occupational Mobility in Great Britain and the United States since 1850. *American Economic Review*, 103(4):1109–1137.

Matthijs, K., van de Putte, B., and Vlietinck, R. (2002). The Inheritance of Longevity in a Flemish Village (18th-20th Century). *European Journal of Population*, 18(1):59–81.

Mazan, R. and Gagnon, A. (2007). Familial and Environmental Influences on Longevity in Historical Quebec. *Population: English Edition*, 62(2):271–292.

McGue, M., Vaupel, J. W., Holm, N., and Harvald, B. (1993). Longevity is moderately heritable in a sample of danish twins born 1870-1880. *Journal of Gerontology*, 48(6):B237–B244.

Meng Xue, M. and Koyama, M. (2016). Autocratic rule and social capital: Evidence from Imperial China. Technical report, Mimeo.

Parman, J. (2012). Gender and Intergenerational Mobility: Using Health Outcomes to Compare Intergenerational Mobility Across Gender and Over Time. Working Paper 122, College of William and Mary.

Perls, T. T., Wilmoth, J., Levenson, R., Drinkwater, M., Brewster, M. S., Kunkel, L., and Puca, A. (2002). Life-Long Advantage of Siblings of Centenarians. *Proceedings of Sciences of the United States of America*, 99(12):8442–8447.

Pomeranz, K. (2000). *The Great Divergence, China, Europe and the making of the modern world economy*. Princeton, New Jersey: Princeton UP.

Pronzato, C. (2012). An examination of paternal and maternal intergenerational transmission of schooling. *Journal of Population Economics*, 25(2):591–608.

Rosenberg, N. A. and Kang, J. T. (2015). Genetic Diversity and Societally Important Disparities. *Genetics*, 201(1):1–12.

Royer, H. (2009). Separated at Girth: US Twin Estimates of the Effects of Birth Weight. *American Economic Journal: Applied Economics*, 1(1):49–85.

Sacerdote, B. (2007). How Large are the Effects from Changes in Family Environment? A Study of Korean American Adoptees. *The Quarterly Journal of Economics*, 122(1):119–157.

Salaris, L., Tedesco, N., and Poulain, M. (2013). Familial transmission of human longevity: a population-based study in an inland village of Sardinia (Italy), 1850-2010. *Vienna Yearbook of Population Research*, 11:325–349.

Solon, G. (1992). Intergenerational Income Mobility in the United States. *American Economic Review*, 82(3):393–408.

Solon, G. (1999). Intergenerational mobility in the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3, Part A, chapter 29, pages 1761–1800. Elsevier, 1 edition.

Song, X., Campbell, C. D., and Lee, J. Z. (2015). Ancestry Matters: Patrilineage Growth and Extenction. *American Sociological Review*, 80(3):574–602.

Swedlund, A., Meindl, R., Nydon, J., and Gradie, M. (1983). Family Patterns in Longevity and longevity Patterns of the Family. *Human Biology*, 55(1):115–129.

Wong, R. B. (1997). *China Transformed: Historical Change and the Limits of European Experience*. Ithaca, NY: Cornell University Press.

Yashin, A. and Iachine, I. (1997). How frailty models can be used for evaluating longevity limits: Taking advantage of an interdisciplinary approach. *Demography*, 34:31–48.

You, D., Gu, D., and Yi, Z. (2010). Familial Transmission of Human Longevity Among the Oldest-Old in China. *Journal of Applied Gerontology*, 29(3):308–332.

Yu, M. (2012). *The Taiping Rebellion: A Military Assessment of Revolution and Counter-revolution*. University Press of Kentucky.

Figure 1: China Multi-Generational Panel dataset-Lioning

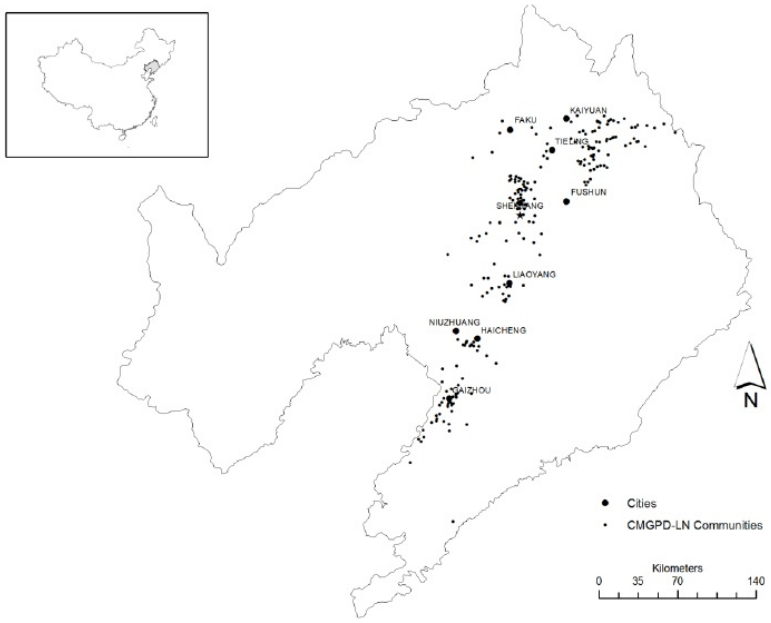

Figure 2: Lifespan Males - 1789-1910

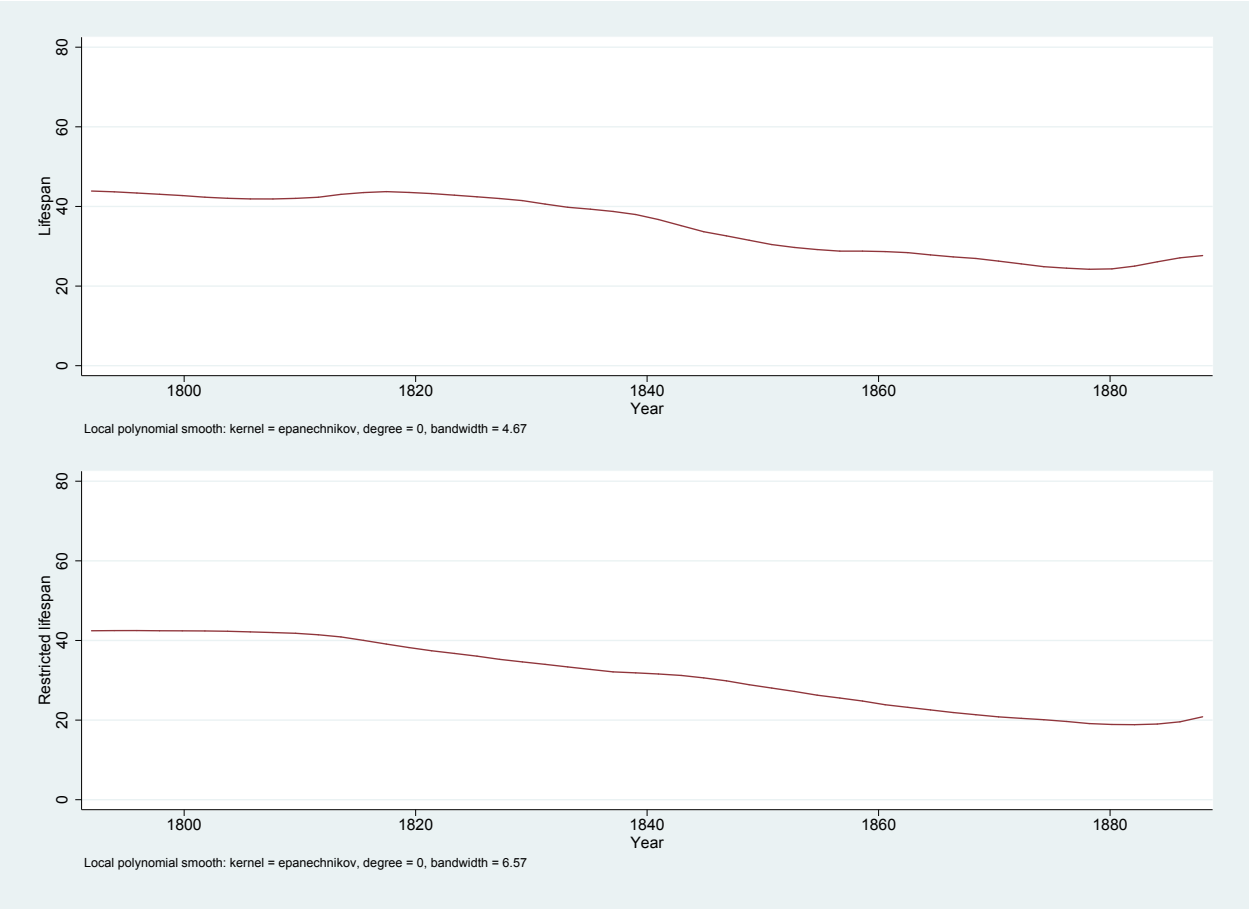

Figure 3: IGT across cohorts 1830-1906

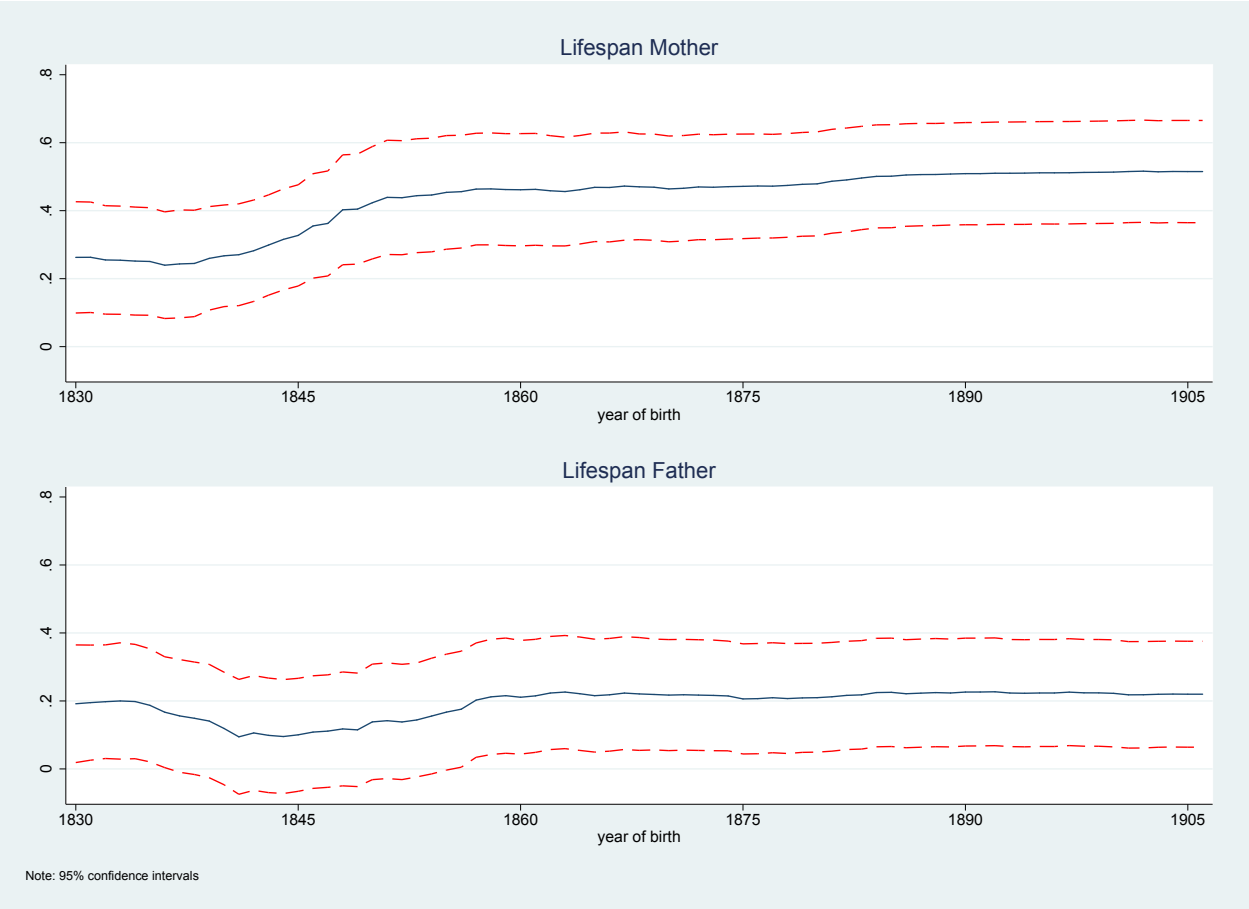

Table 1: Descriptive statistics

Variable	*Mean*	*Std. Dev.*	*N*
lifespan males	33.852	21.625	36,511
Lifespan father	46.243	15.488	5229
Lifespan mother	43.197	14.297	7190
Lifespan Gfather	55.175	12.879	2205
Lifespan Gmother	60.803	9.548	2750
Lifespan GGfather	57.481	12.243	1132
Lifespan GGmother	68.021	6.876	340
Lifespan GGGfather	59.089	11.578	552
Lifespan GGGGfather	59.858	10.859	295

Table 2: IGT across generations

	LogLifespan						
	father	mother	Gfather	Gmother	GGfather	GGmother	GGGfather
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables:							
LogLifespan child	0.299***	0.515***	0.055	0.634***	0.044	0.688	-0.054
se	(0.041)	(0.049)	(0.040)	(0.074)	(0.048)	(0.418)	(0.066)
N	10518	10518	10518	10518	6814	612	4211
LogLifespan father			-0.004				
se			(0.026)				
N			1843				
LogLifespan mother				0.463***			
se				(0.048)			
N				2462			
LogLifespan Gfather					0.118**		
se					(0.057)		
N					765		
LogLifespan Gmother						0.261	
se						(0.172)	
N						144	
LogLifespan GGfather							-0.033
se							(0.063)
N							204
predictions			-0.001	0.233***	0.000	0.062***	0.000
se			(0.008)	(0.033)	(0.002)	(0.019)	(0.000)
Year and District FE	✓	✓	✓	✓	✓	✓	✓
Individual characteristics	✓	✓	✓	✓	✓	✓	✓
Assortative mating	✓	✓	✓	✓			

Each estimates is from a separate regression of the loglifespan of a family member on the lifespan of an older member.
Predictions are obtained by the product of the IGC estimates of consecutive generations.
Standard errors of prediction are computed using the Delta method.
Sample size restricted to available lifespan for M, F, GF and GM.
Standard errors clustered at descent level.
Significance levels: *** 1% ** 5% * 10%.

Table 3: AR models of IGT across generations

Dependent variable:	LogLifespan				
	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
	(1)	(2)	(3)	(4)	(5)
LogLifespan mother	0.519***	0.498***	0.669***	0.471*	0.861*
	(0.050)	(0.049)	(0.182)	(0.250)	(0.453)
LogLifespan father	0.300***	0.271***	0.263*	0.431**	0.612*
	(0.041)	(0.041)	(0.155)	(0.204)	(0.355)
LogLifespan Gmother		0.373***	0.366	0.569*	0.801
		(0.069)	(0.311)	(0.340)	(0.855)
LogLifespan Gfather		0.052	-0.009	-0.060	0.268
		(0.039)	(0.122)	(0.132)	(0.369)
LogLifespan GGmother			0.247	-0.010	0.407
			(0.522)	(0.511)	(0.777)
LogLifespan GGfather			0.304*	0.024	-0.227
			(0.176)	(0.212)	(0.344)
LogLifespan GGGfather				-0.142	-0.278
				(0.239)	(0.349)
LogLifespan GGGGfather					-0.127
					(0.245)
N	10475	10475	436	328	139

All models include Year and district FE and individual characteristics.
Standard errors clustered at descent level.
Significance levels: *** 1% ** 5% * 10%.

Table 4: IGT estimates with grand-father and twin-parent fixed effects

Dependent variable:	Log Lifespan							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: mother LS only								
LogLifespan mother	0.920***	0.556***	0.617***		0.545***		0.539***	
	(0.065)	(0.099)	(0.068)		(0.076)		(0.076)	
Panel B: father LS only								
LogLifespan father	0.667***	0.295***		0.296***		0.298***		0.257***
	(0.067)	(0.098)		(0.079)		(0.078)		(0.079)
Panel C: father and mother LS								
LogLifespan mother	0.802***	0.533***	0.567***	0.533***	0.521***	0.481***	0.522***	0.397***
	(0.059)	(0.102)	(0.068)	(0.082)	(0.075)	(0.083)	(0.075)	(0.098)
LogLifespan father	0.337***	0.150	0.235***	0.174**	0.118*	0.192**	0.094	0.197**
	(0.052)	(0.103)	(0.056)	(0.078)	(0.063)	(0.076)	(0.068)	(0.077)
N	9187	9187	9113	3305	6871	3305	6871	3305
Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual characteristics					✓	✓	✓	✓
Parental characteristics							✓	✓
GF FE		✓						
Twin mother			✓		✓		✓	
Twin father				✓		✓		✓

Twins identified using mother id.
Standard errors clustered at descent level.
Significance levels: *** 1% ** 5% * 10%.

Table 5: IGT across 7 generations

Dependent variable:	LogLifespan son					
	g2	g3	g4	g5	g6	g7
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: no controls						
LogLifespan mother (g1)	0.467***					
	(0.172)					
LogLifespan mother (g2)		0.430**				
		(0.170)				
LogLifespan mother (g3)			0.476***			
			(0.117)			
LogLifespan mother (g4)				0.498***		
				(0.123)		
LogLifespan mother (g5)					0.667***	
					(0.168)	
LogLifespan mother (g6)						0.419*
						(0.225)
Panel B: including individual and parental characteristics						
LogLifespan mother (g1)	0.399***					
	(0.181)					
LogLifespan mother (g2)		0.331*				
		(0.176)				
LogLifespan mother (g3)			0.386***			
			(0.122)			
LogLifespan mother (g4)				0.485***		
				(0.122)		
LogLifespan mother (g5)					0.703***	
					(0.189)	
LogLifespan mother (g6)						0.391*
						(0.202)
Average birthweight						
	g2	g3	g4	g5	g6	g7
son	1795	1805	1823	1842	1865	1886
	g1	g2	g3	g4	g5	g6
mother	1764	1775	1793	1813	1836	1861
N	378	724	1,385	1,082	535	233

Each estimates is from a separate regression of loglifespan of son on lifespan of mother
All models include Year and District FE and control for Assortative mating
Standard errors clustered at descent level.
Significance levels: *** 1% ** 5% * 10%.

Table A1: Descriptive Statistics of control variables

	Mean	**Std. Dev.**
disabled	0.057	0.232
hh size	14.749	13.694
N. brothers	0.849	1.166
N. sisters	0.16	0.476
migrant	0.089	0.285
early life shocks	2.206	2.26
antenatal shocks	0.349	0.743
birth order 1	0.346	0.476
birth order 2	0.225	0.417
birth order 3	0.145	0.352
birth order 4	0.097	0.296
birth order 5	0.064	0.244
birth order 6	0.041	0.198
father disabled	0.116	0.321
father migrant	0.132	0.339
mother disabled	0	0.012
mother migrant	0.085	0.279

Sample size 36,511.

Table A2: IGC estimates with twin-parent fixed effects*

Dependent variable:	Log Lifespan			
	(1)	(2)	(3)	(4)
LogLifespan mother	0.521***	0.481***	0.522***	0.397***
	(0.075)	(0.083)	(0.075)	(0.098)
LogLifespan father	0.118*	0.192**	0.094	0.197**
disabled	0.486***	0.396***	0.484***	0.388***
	(0.036)	(0.034)	(0.036)	(0.034)
Log(household size)	-0.010	-0.011	-0.008	-0.016
	(0.022)	(0.025)	(0.021)	(0.025)
n. of brothers	0.051***	0.012	0.050***	0.013
	(0.013)	(0.016)	(0.013)	(0.016)
n. of sisters	0.043	0.045	0.043	0.043
	(0.029)	(0.029)	(0.028)	(0.029)
migrant	0.397***	0.402***	0.518***	0.439***
	(0.037)	(0.044)	(0.048)	(0.052)
earlylife shock	-0.019**	-0.026**	-0.019**	-0.025**
	(0.009)	(0.010)	(0.009)	(0.010)
antenatal schocks	-0.029*	-0.019	-0.028*	-0.020
	(0.015)	(0.024)	(0.015)	(0.024)
birth order 1	0.336***	0.100	0.326***	0.085
	(0.069)	(0.063)	(0.069)	(0.062)
birth order 2	0.339***	0.132**	0.331***	0.119*
	(0.069)	(0.065)	(0.069)	(0.064)
birth order 3	0.298***	0.096	0.292***	0.087
	(0.062)	(0.069)	(0.062)	(0.068)
birth order 4	0.343***	0.148**	0.341***	0.144**
	(0.061)	(0.058)	(0.061)	(0.059)
birth order 5	0.171***	0.018	0.170***	0.007
	(0.062)	(0.073)	(0.062)	(0.074)
birth order 6	0.165**	0.119	0.166**	0.110
	(0.068)	(0.083)	(0.068)	(0.084)
father disabled			-0.000	-0.000
			(0.036)	(0.036)
father migrant			-0.084	-0.077
			(0.052)	(0.092)
mother disabled			0.000	-0.583***
			(.)	(0.161)
mother migrant			-0.138**	-0.004
			(0.066)	(0.089)
father age at birth			0.003	
			(0.004)	
father age^2 at birth			-0.000	
			(0.000)	
mother age at birth				0.008
				(0.010)
father age^2 at birth				-0.000
				(0.000)
N	6871	3305	6871	3305
Year and District FE	✓	✓	✓	✓
Twin mother	✓		✓	
Twin father		✓		✓

Twins identified using mother id.
Full table (panel C, col.5-8, Table 4)*.
Standard errors clustered at descent level.
Significance levels: *** 1% ** 5% * 10%.

Table A3: IGT estimates adjustment

Dependent variable:	Log Lifespan						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			$\lambda = 1$	$\lambda = 0.75$	$\lambda = 0.50$	$\lambda = 0.25$	$\lambda = 0.1$
Panel A: mother twin FE							
LogLifespan mother	0.595**	0.567***	0.539***	0.550*	0.558	0.563	0.565
	(0.244)	(0.068)	[0.197]	[0.315]	[0.553]	[1.269]	[3.421]
GM FE	✓						
Twin mother		✓					
$\theta = 0.5$			✓	✓	✓	✓	✓
Panel B: father twin FE							
LogLifespan father	0.150	0.174**	0.198*	0.188	0.182	0.177	0.175
	(0.103)	(0.078)	[0.119]	[0.182]	[0.310]	[0.697]	[1.860]
GF FE	✓						
Twin father		✓					
$\theta = 0.5$			✓	✓	✓	✓	✓

Standard errors in parentheses clustered at descent level. Standard errors in brackets are computed following **?** using the pooled variance and taking the square root of $V(\widehat{\beta_{TS}} = \frac{V(\widehat{\beta_{TW}})(\frac{1}{1-\theta})^2(N_{TW}-1)+V(\widehat{\beta_{SIB}})(\frac{\theta}{1-\theta})^2(N_{SIB}-1)}{N_{TW}+N_{SIB}})$
θ is the share of dizygote over monozygote twins.
Year and District FE included. Control for assortative mating.
Significance levels: *** 1% ** 5% * 10%.

Table A4: IGT estimates - excluding potential migrants

Dependent variable:	Log Lifespan							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: mother LS only								
LogLifespan mother	0.797***	0.513***	0.543***		0.566***		0.553***	
	(0.084)	(0.115)	(0.081)		(0.080)		(0.080)	
Panel B: father LS only								
LogLifespan father	0.547***	0.241*		0.269***		0.274***		0.231***
	(0.092)	(0.134)		(0.088)		(0.085)		(0.086)
Panel C: father and mother LS								
LogLifespan mother	0.711***	0.499***	0.508***	0.499***	0.534***	0.452***	0.530***	0.352***
	(0.074)	(0.117)	(0.080)	(0.085)	(0.079)	(0.087)	(0.079)	(0.103)
LogLifespan father	0.276***	0.149	0.175**	0.161*	0.154**	0.180**	0.124	0.186**
	(0.069)	(0.136)	(0.070)	(0.087)	(0.070)	(0.084)	(0.076)	(0.085)
N	6275	6024	6176	2939	6019	2939	6019	2939
Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual characteristics					✓	✓	✓	✓
Parental characteristics							✓	✓
GF FE		✓						
Twin mother			✓		✓		✓	
Twin father				✓		✓		✓

Twins identified using mother id.
Standard errors clustered at descent level.
Significance levels: *** 1% ** 5% * 10%.

Table A5: IGT estimates with grand-father and twin-parent fixed effects

Dependent variable:	Log Lifespan							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: mother LS only								
LogLifespan mother	0.953***	0.606***	0.451***		0.409***		0.435***	
	(0.107)	(0.168)	(0.134)		(0.142)		(0.145)	
Panel B: father LS only								
LogLifespan father	0.626***	0.306*		0.222		0.092		0.060
	(0.111)	(0.175)		(0.233)		(0.233)		(0.235)
Panel C: father and mother LS								
LogLifespan mother	0.845***	0.580***	0.433***	0.302*	0.394***	0.303*	0.418***	0.232
	(0.095)	(0.172)	(0.136)	(0.167)	(0.143)	(0.173)	(0.145)	(0.204)
LogLifespan father	0.312***	0.186	0.094	0.163	0.079	0.031	0.106	0.024
	(0.083)	(0.178)	(0.115)	(0.232)	(0.126)	(0.233)	(0.139)	(0.238)
N	3444	3444	3357	1151	2981	1151	2981	1151
Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual characteristics					✓	✓	✓	✓
Parental characteristics							✓	✓
GF FE		✓						
Twin mother			✓		✓		✓	
Twin father				✓		✓		✓

Twins identified using father id.
Standard errors clustered at descent level.
Significance levels: *** 1% ** 5% * 10%.