The Role of Family Background for Earnings in Rural China

Tor Eriksson (Aarhus School of Business, Aarhus University)
Yingqiang Zhang (Beijing Jiaotong University)

Abstract
This paper provides estimates of brother income correlations for rural China. Brother correlations are a parsimonious measure of the importance of family and community background as determinants of individuals’ economic status. We find internationally high levels of income similarity for brothers and siblings: 0.57 and 0.58, respectively. We show that compared with the 1990s, income correlations have decreased in more recent years, but remain high. Furthermore, we document virtually no differences between the coastal and interior provinces and by father’s education. The high brother correlations imply that the high level of income inequality in China is likely to persist.

JEL Codes: D31, J62

Keywords: Family background, sibling correlation, income inequality

Acknowledgement: Work on this paper began when Yingqiang Zhang was a visiting doctoral student at Aarhus School of Business, Aarhus University during the academic year 2008-9, funded by a grant from the China State Scholarship Fund. He is also grateful for a grant from the General Motor Corporation’s China Development Research Foundation. We have received valuable comments from Anders Björklund and participants at a seminar at Renmin University on an earlier version of the paper.

Corresponding author. Address: Department of Economics, ASB, Aarhus University, Hermodsvej 22, DK-8230 Aabyhoej, Denmark. E-mail: tor@asb.dk
1 Introduction

As is well known, the two previous decades of persistent high economic growth in China has been accompanied by growing income inequality. Due to her size and huge differences in natural conditions for economic activities, China had already before the economic reforms large differences in average incomes between the more developed urban areas and the rural areas (Knight and Song (1999)), and the concentration of economic development to certain special economic zones and the coastal regions has led to a further increase in the urban-rural income gap. As a consequence, the level and growth of Chinese income inequality to a high extent reflect rural-urban differences; see e.g., Sicular et al. (2006), Wan (2007) and Yang (2002). The preponderance of research on income inequality has, however, predominantly focused on urban areas and has mainly dealt with determinants of cross-sectional inequality or aggregate trends therein.

Considerably less attention has been paid to issues like: to what extent is the increase in inequality reflecting changes in the relative role of permanent and transitory income components; for two exceptions, see Jalan and Ravallion (1998) and Whalley and Yue (2009). Another central question is: to what extent has increased cross-sectional inequality been accompanied by changes in income mobility? This question is important because increased inequality is a less severe problem if it is associated with an increase in individuals’ income mobility over their life cycle. Several studies have shown – see e.g., Khor and Pencavel (2006) and Ding and Wang (2008) – that income mobility is indeed high in China, and actually higher than in many industrialized countries. On the other hand, recent studies (see Shi, Liu, Nueth and Xin (2010) and references cited therein) have also documented that income mobility and its equalizing effect on inequality has gradually been decreasing in rural China.

In this paper we will be concerned with another, longer-term aspect of income inequality. More precisely, we will examine the level and changes in intergenerational economic mobility in rural areas in nine provinces during the period 1989 to 2006.\(^1\) The three main questions addressed are:

\(^1\) The analysis in Shi et al. (2010) builds on the same data set as the current paper.

\(^2\) We know of only three earlier studies on China that focus on intergenerational mobility: Chen and Feng (2009), Zhang and Eriksson (2010) and Gong, Leigh and Meng (2010). The first studies the relationship between parental education and adult children’s wages, while the second uses the same data source as the current paper to study the importance of a set of family background factors for the income of the offspring as adults. The third estimates parent-offspring income elasticities on data from urban areas in twelve provinces.
How important is family and community background as a determinant of incomes for people living in the rural areas? Is there a difference between regions which differ with respect to the timing and extent of economic reforms, that is, between the coastal and the interior provinces? Are there signs of changes over time?

The scarcity of studies of income mobility and intergenerational mobility in particular, for China, is chiefly due to the lack of adequate data. Ideally, one would like to have access to long panel data that encompass observations on the parents as well as their offspring as adults and which enable the researcher to link economic outcomes of parents and their children. These types of data sets are rare also in other countries, whereas data sets containing information about economic outcomes of siblings are more common and are for many purposes of intergenerational study at least as useful as panels on two generations. It is also important to notice that the sibling correlation is a broader measure of the impact of family and neighbourhood background than the parent-offspring correlation as it accounts not only for the factors directly related to parents’ income, but also factors like values and aspirations; for a formal analysis, see Solon (1999). In a recent review of the literature, Björklund and Jäntti (2009) make a plausible case for father’s earnings accounting for less than half of the sibling correlation. Remarkably, sibling correlation estimates are only available for the four Nordic countries (Denmark, Finland, Norway, and Sweden) and the United States.

It seems reasonable to expect that family and community background could be at least as important in China since in the Chinese society family ties have by tradition (only temporarily broken during the Maoist era from 1949 to 1977) been very important in transferring economic and social advantages between generations. Thus, it is commonplace to hear of children benefitting from their parents’ social networks in getting access to jobs or educational opportunities. The household registration system (hukou) that was introduced by the Communist Party to restrict geographic mobility, and which has been gradually abolished only towards the end of the period under study, is

---

3 Other demands on an ideal data set include: several observations over time on income for both generations in order to reduce the noise in the measure of lifetime income; offspring’s income should be measured after the beginning of their labour market careers, again to reduce noise in the lifetime income measure; little attrition. The two first problems are likely to give rise to a downward bias in the sibling correlation estimates, and hence lead to an upward bias in mobility; see Hertz (2007) for a discussion. Gong et al. (2010) use a survey from 2004 to obtain information about the offspring’s incomes and their parents. Parents’ incomes are predicted from repeated cross-section data from the period 1986-2004.

4 The main limitation of the siblings approach is that it allows for making inferences about the effect of family background but not always about the effect of the parents’ economic position.

likely to have reduced possibilities for people from the rural areas to access better opportunities and hence contributed to lower intergenerational mobility.

Owing to the one child policy implemented in the urban regions in China since 1978, the applicability of a siblings approach to intergenerational mobility is limited to the rural areas. As the rural population makes up as much as 54.3 per cent of the total population (in 2008), and since much of the poverty is concentrated in rural areas, policies to reduce overall inequality needs to target these areas, it is particular interesting to learn more about the drivers of income inequality in the rural population.

2 Statistical model
We estimate the sibling correlations following the literature beginning with Solon (1992) and Zimmermann (1992); see Solon (1999), and Björklund and Jäntti (2009) and Black and Devereux (2010) for two reviews of the more recent literature. Let

\[ y_{ijt} = \beta'X_{ijt} + \epsilon_{ijt} \] (1)

where \( y_{ijt} \) denotes the logarithm of annual income for the \( j \)th sibling in family \( i \) in year \( t \). \( X \) is a vector of exogenous variables (age and time effects) and \( \beta \) is the associated vector of coefficients. The main object of the analysis is the error term \( \epsilon_{ijt} \). This has three components:

\[ \epsilon_{ijt} = a_i + u_{ij} + v_{ijt} \] (2)

the first of which is a permanent component shared by all siblings of family \( i \), the second is an individual-specific permanent component not shared by siblings of the same family, and the third is a transitory component reflecting temporary shocks to earnings or measurement error. Under the assumption that all three error components are orthogonal (i.e., that \( u_{ij} \) is purely individual) the variance of the error term is

\[ \sigma^2(\epsilon) = \sigma^2(a) + \sigma^2(u) + \sigma^2(v) \] (3)

---

6 Even if migrant workers are not included in the rural population (as is commonly done in the national household surveys), rural poor account for about 90 per cent of all those living below the poverty line. See World Bank (2009).
where the first left hand side term is the variance in income due to differences between families and second term captures the variance within families.

Further, the covariance of a pair \((i, j')\) of randomly drawn siblings’ incomes (purged of age and time effects) is

\[
\text{Cov}(\varepsilon_{ijt}, \varepsilon_{ij't}) = \text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) = \sigma^2(a)
\]

and the correlation of long-run incomes among siblings is

\[
\rho = \frac{\sigma^2(a)}{\sqrt{\sigma^2(a) + \sigma^2(u)}}
\]  

(4)

Thus, the sibling correlation is the proportion of the population variance in long-run incomes that can be attributed to the factors that siblings have in common, both within the family (e.g., common genes, resources in the home, parental aspirations, cultural inheritance) and in the surrounding neighbourhood (e.g., schools) of the family. In a Chinese context, in addition to the normal factors you would think work via family background social connections (guanxi) are likely to be very important. The measure of intergenerational mobility is 1-\(\rho\).

We proceed as follows. First Eq. (1) is estimated by OLS. The dependent variable is the log of annual individual total income\(^7\) deflated by the 2006 consumer price index. We include a cubic in age and year dummies as \(X\)-variables and compute the residuals from (1) to estimate the variance of the error components in (2) using \textit{xtmixed} in Stata (version 9). This information is next used to estimate the sibling correlation \(\rho\). We do this for brothers and for siblings irrespective of gender\(^8\), separately for the coastal and interior provinces, the 1990s and the period 2000-6, and by father’s education and employer. It should be noted that the data set is rather small and hence dividing it into sub-samples may exacerbate the sample size problem. Furthermore, comparisons between

\(^7\) The income measure is quite comprehensive as it includes not only income from work but also income from business, farming, fishing, gardening, livestock, subsidies and a number of other minor sources (such as leased land, non-land rental, and remittances). Note, that agricultural income also encompasses the value of self-consumed products.

\(^8\) The number of sisters is considerably smaller and hence the estimates for them are not reported.
groups should be interpreted with due caution as the extent of noise in income variable and attrition may differ between the sub-samples.

3 Data description

The data set used in this study originates from the China Health and Nutrition Survey (CHNS), which is an ongoing international collaborative project between the Carolina Population Center at University of North Carolina at Chapel Hill and the Chinese Center for Disease Control and Prevention. The CHNS is a longitudinal survey which includes seven waves in 1989, 1991, 1993, 1997, 2000, 2004 and 2006, and it has been conducted by an international team of researchers representing various disciplines. The first survey included 3,795 households and in 1993, 1997, 2000 and 2004, 3,616, 3,441, 3,875 and 4,403 households respectively, participated in the survey.

The number of individuals covered by the surveys is approximately 19,000. All individuals in each household were surveyed in 1991, 1993, 1997 and 2000 for all data. This is important for the current paper as it provides us with data from two generations. The surveys were conducted in three coastal provinces: Liaoning, Shandong and Jiangsu, and in five inland provinces: Henan, Hubei, Hunan, Guangxi and Guizhou. Within each province, counties are stratified by income and a weighted sampling scheme has been adopted to randomly select a number of counties within which in turn a number of villages and townships are randomly selected. The data set contains rich information on individuals and households in terms of economic, demographic and social factors including age, gender, household income, individual income, and parental background.

This paper focuses on the offspring of the households. The offspring are defined as the children of the parents in the surveyed households who are above 19 years of age and are biological, step or foster children of the father, mother, or both. Thus, the siblings are identified through their parents (one or both). We have applied the following sample restrictions on the original data set. First, all observations where information regarding family background is lacking are discarded. Second, for the offspring only individuals in the age range 20 to 50 years with positive annual incomes are included. This gives us 1,765 offspring observations. It is also worth pointing out that we observe

---

9 For a detailed description of the project and the data created within it, see http://www.cpc.unc.edu/china.
10 In the 1997 wave Liaoning was replaced by Heilongjiang, but returned to the survey in 2000.
60 per cent of the sample individuals at least twice, which is helpful as it contributes to a reduction of some of the transitory variation in individuals’ incomes.\textsuperscript{11}

\textit{Table 1} gives some descriptive statistics of our estimation sample plus some corresponding information for the fathers of the offspring. These largely reflect the fact that this emanates from the rural and hence less developed parts of China. Thus, the share employed in private firms is relatively low and the proportion with high school education or above is lower than the national average. The same applies to average individual and household incomes. As for the differences between the generations, one may note the higher educational level of the offspring and the lower fraction of the offspring working in the governmental and state-owned enterprise sector.

\begin{table}
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Variable} & \textbf{Offspring} & \textbf{Father} \\
\hline
Gender: male (%) & 67.9 & \\
\hline
\textbf{Province}: & & \\
Coastal (%) & 35 & 26 \\
Interior (%) & 65 & 74 \\
\hline
\textbf{Education} (%): & & \\
Primary school or less & 15.7 & 26.1 \\
Middle school & 47.8 & 41.9 \\
High school, vocational education, and college and above & 36.5 & 31.9 \\
\hline
\textbf{Employer} (%): & & \\
Agriculture & 18.2 & 9.9 \\
Co-operatives, private firms (incl self-employed) & 33.6 & 20.7 \\
Government sector, state-owned firms & 48.2 & 69.3 \\
\hline
Age (years) & 32.7 & \\
& (6.3) & \\
\hline
Individual income (RMB) & 7,403.9 & \\
& (10,202.1) & \\
\hline
Household income (RMB) & 27,575.0 & \\
& (33,849) & \\
\hline
Number of observations & 1,765 & 473 \\
\hline
\end{tabular}
\caption{Descriptive statistics (standard error in parentheses)}
\end{table}

\textsuperscript{11} As most adjacent observations are three or four years apart, the likelihood of serial correlation which would lead to a downward bias in the sibling correlation is low. With the data at hand it is not possible to incorporate serial correlation in the estimations. If possible and if accounting for serial correlation would make a difference, the estimates would be larger than those reported below.
4 Estimates

In our presentation and discussion of the sibling correlations estimates we will henceforth concentrate on the estimates for brothers. This is primarily motivated by the absence of income for some years for the women in the data (most likely when they are taking care of their small children) and by the fact that sibling correlations will also be affected by gender wage differentials. However, we have also computed sibling correlations and will occasionally mention them as we discuss the brother correlations. The standard errors are computed by bootstrapping.\(^{12}\)

As can be seen from Table 2, the brother correlation for the nine rural provinces is 0.57. This is from an international perspective a high correlation; it is for instance about a third higher than the U.S. estimates which are from the higher end of the international league.\(^{13}\) A correlation of 0.57 implies an \(R^2\) of approximately 33 per cent, which is more than what is typically explained by a standard specification of a Mincerian earnings equation. The corresponding correlation for siblings is 0.58. Thus, the importance of family background (including neighbourhood) for persons living in Chinese rural areas is very high. Gong et al. (2010) obtain income elasticity estimates from urban areas in China of 0.74, 0.84, 0.33 and 0.47 for father-son, father-daughter, mother-son, and mother-daughter, respectively.\(^{14}\) So, mobility is low also in the urban parts of the country.

What are the effects of governmental policies that have both spread the economic reforms to a growing number of regions and increasingly focussed on the rural areas and on improving the conditions and incomes of workers in the agricultural sector? Dividing the time period into two parts – the 1990s and the period 2000-6, respectively – shows that the brother correlation has decreased over time: from 0.61 in the nineties to 0.54 in more recent years, while the siblings correlation has remained unchanged.

\(^{12}\) The standard errors are fairly small as compared to those obtained from studies of small samples from another large country, the United States.

\(^{13}\) Solon, Corcoran, Gordon and Laren (1991) obtain estimates around 0.45 for the U.S.; Björklund, Eriksson, Jäntti, Raaum and Österbacka (2002) replicate Solon et al. (1991) on a larger sample from the same data source and obtain an estimate of 0.43. They also get comparable and considerably lower estimates for the four Nordic countries: 0.23 for Denmark, 0.26 for Finland, 0.14 for Norway, and 0.25 for Sweden. Mazumder (2008) estimate sibling correlations on two larger US samples using different estimation techniques and finds brother correlations in the interval 0.43 to 0.54. Corresponding estimates for sister correlations are lower: 0.32 to 0.36.

\(^{14}\) These elasticities are likely to be somewhat lower if transformed into correlations (because of the higher income inequality for the offspring generation).
Next we turn to look for possible differences between different sub-samples. The first difference we consider is that between the coastal and interior provinces. This distinction is potentially important since the coastal provinces (Liaoning, Shandong and Jiangsu) were affected earlier and to a higher extent by the Chinese government’s economic reforms and regional preferential policies; see Démurger (2000). The share of brothers (in our sample) of which at least one has a high school or higher level education is higher in the coastal provinces; 43.9 and 36.4 per cent in the coastal and interior provinces, respectively. It is possible that the higher and faster development in the coastal provinces has weakened the importance of family and community background for individuals’ incomes more than in the provinces in the interior.\(^{15}\) The estimates in Table 2 do not lend strong support to this conjecture, however. Although the brother and sibling correlations are higher in the coastal than in the interior provinces, the differences are small and are not statistically significant.\(^{16}\)

<table>
<thead>
<tr>
<th>Category</th>
<th>Brothers</th>
<th>Siblings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All</strong></td>
<td>0.57 (0.026)</td>
<td>0.58 (0.018)</td>
</tr>
<tr>
<td><strong>Period:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990s</td>
<td>0.61 (0.023)</td>
<td>0.55 (0.026)</td>
</tr>
<tr>
<td>2000-6</td>
<td>0.54 (0.031)</td>
<td>0.56 (0.031)</td>
</tr>
<tr>
<td><strong>By area:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coastal provinces</td>
<td>0.58 (0.027)</td>
<td>0.58 (0.027)</td>
</tr>
<tr>
<td>Interior provinces</td>
<td>0.55 (0.032)</td>
<td>0.57 (0.030)</td>
</tr>
<tr>
<td><strong>By father’s education:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary school or lower</td>
<td>0.56 (0.039)</td>
<td>0.62 (0.029)</td>
</tr>
<tr>
<td>High school or above</td>
<td>0.54 (0.031)</td>
<td>0.57 (0.024)</td>
</tr>
<tr>
<td><strong>By father’s employer:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government, state-owned enterprise</td>
<td>0.60 (0.024)</td>
<td>0.86 (0.019)</td>
</tr>
<tr>
<td>Agriculture</td>
<td>0.83 (0.047)</td>
<td>0.72 (0.050)</td>
</tr>
<tr>
<td>Private enterprise (incl. self-employed), co-operative</td>
<td>0.55 (0.043)</td>
<td>0.56 (0.031)</td>
</tr>
</tbody>
</table>

\(^{*}\)Bootstrapped standard errors in parenthesis

\(^{15}\) We might also expect that the coastal provinces would be the main source of the decline in the aggregate brother correlation over time. Unfortunately, because of the limited size of the estimation sample we cannot divide the data into coastal/interior provinces by time periods.

\(^{16}\) Zhang and Eriksson (2010) find a similar pattern for differences in the more narrow measure of intergenerational mobility, father-son income elasticities.
How important is the father’s education as an element in the family background shared by brothers? To shed some light on this, we estimate brother correlations for two sub-samples: the offspring of fathers with primary or lower education and those whose fathers have/had a middle school or higher level of education. As can be seen from the table, the difference is slightly larger than the one observed for coastal/interior provinces: 0.62 for sons of low educated father and 0.54 for the others. A similar pattern is found for siblings. Thus, parental education does seem to play a role for differences in offspring’s income similarity. This is a noteworthy finding since for the parental generation a person’s economic status was only relatively weakly related to his or her level of education. Rather, a person’s socio-economic status was more influenced by his or her employer. Chen and Feng (2009) find that the effect of father’s education on individuals’ earnings is quite sensitive to the jobs the father has held. Moreover, Zhang and Eriksson (2010) find that the father’s employer contributed substantially more than father’s education to the inequality of opportunity in China. When we divide the data set into sub-samples according the father’s employer, we find for the brothers that the estimate for the agricultural sector is higher (0.83), whereas the correlations for the governmental or private/co-operatives sectors do not differ much from the overall estimate: 0.60 and 0.55, respectively. As for siblings, the correlation is also high for those whose father worked in the state sector.

All in all, our estimates by sub-samples indicate that there is a stronger similarity in brothers’ economic status when their father has a primary school education (or less) and when their father was employed in the agricultural sector. Notably, whether they are living in a province in the interior or in the coastal areas does not make a difference. It should be emphasized, however, that on average the correlations are high and that the differences between sub-groups are small. By and large, a similar pattern is discernible for siblings.

5 Conclusions

In this paper we have added to the tiny stock of brother income correlations that have been estimated outside U.S. and the Nordic countries.¹⁷ These are computed from a sample from the rural areas in a survey data set collected from nine Chinese provinces during the period 1989 to 2006.

¹⁷A recent working paper by Schnitzlein (2010) gives a brother correlations estimate of 0.37 for Germany.
Overall, we find high income correlations – 0.57 for brothers and 0.58 for siblings – and rather small differences between sub-groups: coastal/interior regions, father’s education and father’s employer (except agriculture). A recent study by Gong et al. (2010) estimated parent-offspring income elasticities for urban China and found that these are remarkably high from an international perspective. Our study documents that the low income mobility is also present in the rural parts of China\textsuperscript{18}, which is where the bulk of poverty in China is located. Although we find some traces of a decline over time in the importance of family and community background, the high brother correlations observed, overall as well as for sub-groups, are (or should be), given the high level of income inequality in China, a major cause for concern as they imply that the current high level of income inequality will persist also in the future.

\textsuperscript{18} Unfortunately, the data set is not sufficiently large to allow closer examination of the relative role of different pathways of intergenerational transmission.
References


Gong, C., A. Leigh, and X. Meng (2010), Intergenerational Income Mobility in Urban China. IZA Discussion Paper No. 4811


World Bank (2009), From Poor Areas to Poor People: China’s Evolving Poverty Reduction Agenda. An Assessment of Poverty and Inequality in China. World Bank: Poverty Reduction and Economic Management Department, East Asia and Pacific Region

