Family, Community and Long-Term Earnings Inequality[§]

Paul Bingley
SFI Copenhagen

Lorenzo Cappellari
Università Cattolica Milano

Konstantinos Tatsiramos
University of Nottingham

November 2015

Abstract

Correlations between the earnings of siblings reflect shared family and community background, but evidence is mixed on the relative magnitudes of these influences. Using administrative data on the Danish population we link brothers, schoolmates and teenage neighbors and estimate a model of multi-person earnings dynamics to measure jointly the relative importance of family, neighborhoods and schools for long-term earnings. We find that: (1) family is by far the most relevant factor; (2) the influence of neighborhoods and schools falls rapidly, becoming insignificant by age 30; and (3) community effects are persistent and upward biased by a factor of five if family effects are ignored.

Keywords: Sibling correlations; Neighborhoods; Schools; Life-cycle earnings; Inequality

JEL codes: D31, J62

[§] We thank Anders Björklund, Markus Jännti, Andreas Peichl, Josef Zweimüller, seminar audiences at IZA in Bonn, SOFI in Stockholm, University of Padua, University of Venice, ZEW in Mannheim, European University Institute in Florence, IAB in Nuremberg, and conferences at Royal Economic Society in Manchester, European Society of Population Economics in Braga, 16th IZA/CEPR European Symposium in Labour Economics (ESSLE) and Fourth SOLE/EALE World Conference in Montreal for their comments. The usual disclaimer applies. The research for this paper was supported by a British Academy Research Grant (award SG113192) and by a Danish Strategic Research Council Grant (DSF-09-065167).

Addresses for correspondence: Bingley, SFI, The Danish National Centre for Social Research, Herluf Trolles Gade 11, DK-1052, Copenhagen, pbingley@sfk.dk; Cappellari, Department of Economics and Finance, Università Cattolica Milano, Largo Gemelli 1, I-20123 Milano, lorenzo.cappellari@unicatt.it; Tatsiramos, University of Nottingham, School of Economics, University Park, Nottingham, NG7 2RD, UK, konstantinos.tatsiramos@nottingham.ac.uk.

1. Introduction

Family and community background are generally considered as two of the most important factors that determine socioeconomic outcomes, including earnings. Families can determine earnings by transmitting abilities, preferences and resources, while communities can influence earnings through neighborhood quality, school quality and peers. Existing research measuring the influence of community background on earnings through the neighborhood of residence provides mixed evidence of either substantial positive (e.g. Page and Solon, 2003; Raaum, Sörensen and Salvanes, 2006; Chetty, Hendren and Katz, 2015), or zero neighborhood effects (e.g. Oreopoulos, 2003; Ludwig et al., 2013). Besides this mixed evidence, because community background has been associated with the neighborhood, the influence of school quality on earnings – relative to neighborhood quality – has not received much attention. Also, mostly due to data limitations, little is known about the influence of community background on *long-term earnings*. Understanding the relative magnitude of the influences of family, neighborhoods and schools on earnings in the long-run is important for identifying the driving forces of existing inequalities and for interventions that aim to reduce them, especially because any influences early in life may not be long-lasting.

In this paper, using very rich administrative data on the Danish population we link brothers, teenage neighbors and schoolmates and estimate a model of *multi-person* earnings dynamics to measure *jointly* the relative importance of family, neighborhoods and schools for long-term earnings. Using the parameter estimates of the model we decompose for the first time the sibling correlation of earnings into the three components of interest – family, neighborhood and school – allowing for sorting of families across neighborhoods and schools. Exploiting the extraordinary information available in the Danish registers we can observe siblings and their peers over their life-cycle, which ensures we can address measurement error biases and estimate the influence of the determinants of permanent earnings in the very long run, and up to age 51.

The correlation of sibling earnings, which measures the fraction of the variation in permanent earnings that can be attributed to both observed and unobserved factors shared by siblings during childhood, has been widely used in the literature as a broad measure of the influence of *both* family and community background (for reviews see Solon, 1999; Björklund and Jännti, 2009; Black and Devereux, 2011). By observing siblings but also their teenage neighbors and schoolmates we are able to separately identify the influence of families, neighborhoods and schools on permanent earnings. While siblings share both the family and the community, neighbors and schoolmates share only their specific community factor (neighborhood or school, respectively) but do not share the family. We can also separate neighborhood from school effects because there is no full overlap between neighbors and schoolmates in Denmark. That is, different students in the same neighborhood attend different schools and schools enroll students from different neighborhoods. Following Baker and Solon (2003), we exploit this rich information within a model of earnings dynamics, which we extend it to account for the joint earnings dynamics of multiple persons, and decompose the sibling correlation of earnings into family, neighborhood and school effects allowing for sorting of families across neighborhoods and schools.[1]

We find that family is by far the most relevant factor determining long-term earnings. The influence of neighborhoods and schools falls rapidly, becoming insignificant by age 30. This implies on the average a zero influence of community background over the life-cycle, which means that any community effects early in life are relatively small and are not long-lasting. These findings highlight the importance of considering the long-term effects of community background on earnings beyond the first years of working life. We also show that, if family effects are ignored, the influence of neighborhoods and schools are persistent and upward biased by a factor of five. This suggest that linking siblings and their peers and jointly

[1] We discuss formally identification of the model in section 5.3.

estimating the family and community factors within our model captures correlated effects due to residential sorting that would otherwise lead to biased estimates of community influences.

The paper is structured as follows. In Section 2 we present the theoretical background and discuss the related literature. In Section 3 we describe the data and the definition of teenage neighbors and schoolmates, while in Section 4 we present descriptive statistics on earnings of siblings and peers over the life-cycle. In Section 5 we develop the econometric model for assessing the relative importance of families, neighborhoods and schools within the sibling correlation, based on the joint analysis of life-cycle earnings for brothers, neighbors and schoolmates. The main results are presented in Section 6 together with a sensitivity analysis and evidence of heterogeneity by family, neighborhood and school types. We conclude in the last section.

2. Background and Related Literature

The aim of this paper is to identify the determinants of long-term earnings inequality and, in particular, the extent to which earnings inequality can be explained by differences in family and social background. Based on the analysis by Becker and Tomes (1979), families by transmitting abilities, preferences and resources to their offspring can influence their human capital investment and, therefore, their earnings. Community background can also influence individual outcomes through institutions such as the school and its quality (e.g. Hanushek, 2006), or through the quality of neighborhood, or peer influences, social norms and role models in the neighborhood (e.g. Wilson, 1987). Differences between families in the availability of these traits, resources and exposure to the community environment would lead to differences in human capital accumulation. According to human capital theory, differential investments of human capital would generate heterogeneity of both initial earnings and earnings growth (Mincer, 1958, Ben-Porath, 1967). These models predict that heterogeneous investments in human capital induce a negative correlation between initial earnings and earning growth rates. This negative correlation arises because investors trade off lower initial earnings with higher earnings growth in later parts of their working life. The prediction is that inequality of earnings should follow a *u-shape pattern* by age because earnings profiles would exhibit a cross-over property.

The correlation of sibling earnings (or other outcomes) has been used as a way of measuring the *joint* influence of family and community background shared by the siblings (see the reviews in Solon, 1999; Björklund and Jännti, 2009; and Black and Devereux, 2011). To disentangle family from community effects, where community is defined by the neighborhood of residence, a common approach followed in the literature is to compare the correlation of sibling earnings with the correlation of earnings among unrelated neighbors. The idea is that while siblings share both the family and the neighborhood, unrelated neighbors share only the neighborhood but not the family. Following this approach, Page and Solon (2003) using data from the PSID, and Raaum, Sörensen and Salvanes (2006) using administrative data from Norway, find a substantial or non-negligible effect of neighborhoods on earnings. However, the estimate of neighborhood effect is recognized to be an upper bound because of non-random sorting of families into neighborhoods, which leads to a positive correlation between the two factors. The correlation of neighbors' earnings will correctly measure the proportion of variance due to neighborhood effects only if sorting is random. Addressing sorting by exploiting quasi-random assignment of families to public housing projects in Toronto, Oreopoulos (2003) finds a zero influence of neighborhood quality in the total variance of income and wages. Instead, the neighborhood effect is found to be positive and significant for the whole population of Toronto, where assignment to neighborhoods is not random.

Outside the sibling correlation literature, the evidence from social experiments such as the Moving to Opportunity experiment, which offered to eligible families living in high poverty neighborhoods randomly a voucher to move to better neighborhoods, suggests that changes in neighborhood quality had on average little impact on economic outcomes including earnings (e.g. Ludwig et al., 2013). However, Chetty, Hendren and Katz (2015) using administrative data from tax returns find that moving to a lower-poverty neighborhood improves earnings in their mid-twenties for children who were below age 13 when their families moved. Focusing on long term effects, Gould, Lavy and Paserman (2011) using the airlift of Yemenite immigrants as a natural experiment find effects of early childhood environment almost 60 years after only on education but not on other economic outcomes.[2]

Within the sibling correlation literature, because community background has been associated with the neighborhood, the influence of school quality on earnings – relative to neighborhoods – has not received much attention. However, within the school literature there is a long interest on the effect of school quality on earnings. Exploiting variation of school quality across cohorts within U.S. regions Card and Krueger (1992) find that higher quality as measured by a lower pupil/teacher ratio increases the rate of return to schooling and earnings. Examining the effect of pupil/teacher ratio for the UK, Dearden, Ferri and Meghir (2002) find no significant relationship with earnings. More recently, linking the data from the Tennessee STAR experiment, which randomly assigned students and their teachers to classrooms of different size with tax return data, Chetty et al. (2011) find no effect of class size on earnings at age 27, but they find a positive effect of teacher quality. Exploiting a maximum class size rule in Sweden, Fredriksson, Öckert and Oosterbeek (2013) find a positive effect of smaller class size on adult earnings at ages 27 to 42.

In this study, we contribute to these different strands of the literature by developing a unified framework which allows decomposing directly the influence of families, neighborhoods and schools on long-term earnings inequality allowing for sorting of families into communities. The two unique features of this study are the following: (1) we provide the

[2] Studies focusing on educational achievement outside the sibling correlation framework but using quasi-experimental variation of neighborhood quality have also found no impact of neighborhoods (e.g. Jacob, 2004; Gibbons, Silva and Weinhardt, 2013).

first joint decomposition of sibling correlation between family and community effects distinguishing the influence of neighborhoods from schools , and (2) we estimate these effects in the very long run (up to age 51), while most studies measure outcomes up to age 30.

3. Data

We use data from administrative registers of the Danish population. The civil registration system was established in 1968 and everyone resident in Denmark then and since has been registered with a unique personal identification number which has subsequently been used in all national registers enabling accurate linkage. In outline, construction of our dataset proceeds as follows: First we create our sample of brothers by sampling fathers and finding their first and second born sons. Second we link sons to their teenage communities of neighbors and schoolmates.

In order to establish our dataset of brothers we consider men who are first born in the years 1960-1982 and their immediately younger brothers born 1962-1983. This selection is because of completeness of registered parentage and the small number of first sons observed born before 1960.[3] We consider full biological brothers who share both parents. Following the tradition in the sibling literature, we keep men without a younger brother in the sample. The robustness analysis in Section 6.3 shows that their exclusion does not affect results. Next we link our sampled brothers (and singletons) to men in their teenage communities (schools and neighborhoods). School attendance rules were such that pupils should start in first grade in the August of the calendar year they turn seven. The national pupil database was established along with a school reform that made attendance in 9th grade compulsory from the academic

[3] Subsequent sons beyond the first two are very few (4 percent) and are not considered in the analysis. The son birth order is determined irrespective of daughters present in the family. We also exclude from the sample sons who were adopted before age 17; sons who are themselves observed as fathers; brothers born less than 12 months apart; and second sons if they are born more than 12 years after the first or later than 1983.

year beginning August 1973.[4] The database links pupils to the schools they are enrolled from 8th grade and above. School identifiers are consistent over time and schools are classified according to whether they are publicly run (77% of schools and 89% of pupils in our estimation sample) or privately run, and whether they are exclusively for pupils with special educational needs (10% of schools and 1% of pupils in our gross sample).[5]

We link pupils to schoolmates on 31 October of the calendar year they turn 15, which is in the academic year they would normally attend 9th grade.[6] During our sample period, pupils were assigned to public schools on a catchment area basis according to place of residence. Our sample contains 2657 schools with males attending 9th grade. They have on average 14.7 schoolmates. Primary and lower secondary education usually takes place in the same school and most pupils attend the same school for all grades. For example, in 2007, the first year that the pupil database was extended down to grade 1, 90% of pupils in grades 1-8 were enrolled in the same school the following year. Due to the organization of primary and secondary schools largely as a single unit, there is likely to be less pupil mobility between schools than in other countries. The institutional setting makes Denmark a good place to look for school effects, because of the coherence of the schoolmate group.

Address of residence is obtained from the central person register which was established in 1968. Individuals are required to report changes of address to the municipal person register within two weeks. Precision of historical address registration has improved over time and we use parish of residence which is recorded consistently throughout our sample period. Similar to schools, our census point is 31 October of the calendar year a male turns 15. There are 1905 parishes covering on average an area of 22.4 km^2 and containing 19.4 teenage neighbors.

[4] Early or late school start and grade retention were uncommon (less than 10 percent), meaning most pupils begin the final year of compulsory schooling in the calendar year they turn fifteen.

[5] We exclude special schools from our estimation sample.

[6] In robustness checks we consider 8th grade attachment or both 8th and 9th grade attachment as definitions of schoolmates.

An important feature of the data is that there is no full overlap between neighbors (at the level of parish) and schoolmates. This means that different students in the same neighborhood attend different schools and schools enroll students from different neighborhoods. In particular, different students in the same parish attend about 15 different secondary schools, while secondary schools enroll students from around 18 different parishes.

Finally, for both brothers and for peers we use pre-tax annual labor earnings measured in 2005 prices. Table 1 presents the cohorts we include in the sample, the first year we start observing their earnings, the total number of year observed, and the last age observed. Following Baker and Solon (2003) we group data in 2-years birth cohorts and we compute age by imputing each cohort with its earlier year of birth. The selection of birth cohorts ensures that each cohort is observed for at least 6 years (cohort 1982) up to 28 years (cohort 1960).

4. Descriptive statistics on earnings of siblings and peers

In this section we provide a description of the interpersonal covariance structure of earnings. There are two types of cross-person relationships that are of interest to our analysis: 1) between members of the same family (brothers), and 2) between peers attending the same school at age 15 and/or residing in the same neighborhood at the same age.

The covariance of earnings among brothers is computed from families with at least two male children. We group non-sibling peers in clusters depending on whether they shared the school and the neighborhood, only the school, or only the neighborhood. We obtain the between-peers covariance of earnings (at each relevant age) by first computing the within-cluster covariance and then averaging covariances between clusters using the weighting scheme of Page and Solon (2003, pp. 840), which gives more importance to more populated clusters.

We begin by describing the sibling earnings correlation by age in Figure 1. The plot labeled "At same age" reports the computed correlation when the brothers are at the same point in their life-cycle, a counterfactual that is available in our data. The earnings correlation declines between age 24 and 30, and remains stable after age 30. The decline suggests that sources of initial earnings heterogeneity that are shared between brothers are negatively correlated with heterogeneity in earnings growth. As discussed in Section 2, human capital models predict investments in education or training to induce such a negative correlation. The second plot fixes the age of the older among the two brothers at 35 and reports the sibling correlation by age of the younger brother. In this case, the earnings correlation is relatively low at age 24 (actually close to zero) and increases sharply so that by the early-30s it matches the "same age" correlation. This pattern illustrates that the earnings correlation computed between siblings of different ages is an underestimate of the correlation one would obtain observing siblings at the same point in their life-cycle. This is a form of life-cycle bias as the ones discussed in Jenkins (1987) and Haider and Solon (2006). The figure shows that we can observe this bias in the data, which suggests that we have the information required for controlling it in estimation.

Besides human capital investments, the large contemporaneous associations at the early stage of the life-cycle in Figure 1 may also reflect the correlation of transitory shocks. It is well known that earnings instability is large for young cohorts (see e.g. Baker and Solon, 2003). It is also plausible that siblings are subject to common shocks, for example, because of similar local economic conditions at labor market entry. As a way to assess if the relatively large sibling correlation at young ages is driven by permanent earnings differences or transitory fluctuations, we also computed sibling correlations for brothers born at least five, eight or ten years apart, which are shown in Figure 2. The larger the age difference, the less likely it is that brothers entered the same labor market and shared transitory shocks at entry, so that these samples are less likely to be influenced by transitory fluctuations compared with the samples underlying Figure 1. A declining pattern of the sibling correlation between the mid-20s and the early-30s persists even after excluding closely spaced brothers that most likely share transitory earnings fluctuations. This suggests that the source of the convex evolution of sibling correlations is in the permanent earnings component.

In Figure 3 we plot the earnings correlations for non-relative peers at the same point in their life-cycle distinguishing between those sharing both the school and the neighborhood, sharing only the school, or only the neighborhood. These empirical correlations pick-up all sources of peer similarities, both those correlated with family effects and those independent of them. A few points are worth mentioning in this graph. The first is the magnitude of the peer earnings correlation, which is roughly one tenth of the correlation of sibling earnings reported in Figures 1 and 2. Second, the earnings correlation is higher at the beginning of the life-cycle and up to age 30, which implies that after that age the influence of peers appears to be negligible. Third, schools seem to exhibit stronger influence compared to neighborhoods. Finally, the graph also reports the correlation of earnings for "Unrelated" peers, i.e. non-relatives that share neither the school nor the neighborhood. This correlation is computed by randomly matching each individual in the sample with 1000 unrelated peers of the same age. We find this correlation to be equal to zero at each age, which suggests that the evolution of sibling and peer correlations over age is picking up some underlying forces due to families, schools and neighborhoods, and is not simply an artifact of age effects.

5. Econometric model

To estimate the independent contributions of family and community background on permanent earnings we exploit the rich information on the earnings of siblings, teenage neighbors and schoolmates within a model of multi-person earnings dynamics, distinguishing permanent from transitory earnings and allowing for heterogeneous earnings growth.

In particular, the logs of age- and time- adjusted gross annual earnings, denoted by w, are assumed to be the sum of two components, a permanent one denoted by y and a transitory one denoted by v, which are orthogonal by definition:

$$w_{ifsna} = y_{ifsna} + v_{ifsna}\ ;\ E\left(y_{ifsna}v_{ifsna}\right) = 0, \tag{1}$$

where the indices i, f, s, n and a stand for individual, family, school, neighborhood and age.[7]

The model extends the joint earnings dynamics model of Bingley and Cappellari (2013) for three persons (father and sons) to a multi-person setting. The model tackles the two measurement error biases in the estimation of correlations in permanent earnings between persons highlighted in the literature of earnings correlation between family members, particularly fathers and sons. The first source of bias addressed by the model is related to transitory income shocks which makes current earnings a poor measure of permanent earnings (Solon, 1992; Mazumder, 2005). Separate identification of permanent and transitory earnings is granted by the availability of individual level panel data. The second source of bias addressed by the model is related to life-cycle bias due to age differences between family members and the heterogeneous earnings variation over individual life-cycles (Jenkins, 1997; Haider and Solon, 2006).

5.1 Specification of permanent earnings

We allow permanent earnings (y) in equation (1) to depend on both *shared* and *idiosyncratic* components. Shared components capture those determinants of permanent earnings that are common between brothers, schoolmates and neighbors. The idiosyncratic component represents individual-specific sources of variation in permanent earnings. We model life-cycle dynamics of shared components using a specification based on *heterogeneous income profiles* (HIP), which is also known as a *random growth* model. We augment this with a *restricted*

[7] Age is measured in deviation from the life cycle starting point, which is set at 24.

income profile (RIP) process for individual-specific components, which is an idiosyncratic unit root (*random walk*) shock.

As discussed in Section 2, the heterogeneous income profiles specification is inspired by human capital models in which heterogeneity of initial earnings and heterogeneous earnings growth are generated by differential investments (Mincer, 1958; Ben-Porath, 1967). These models predict that heterogeneous investments in human capital induce a negative correlation between initial earnings and earnings growth rates, because investors trade off initial earnings against earnings growth throughout the life-cycle. The resulting negative covariance of intercepts and growth rates would generate a *u-shaped* evolution of earnings dispersion by age due to the 'Mincerian cross-overs' of earnings profiles. These observations, combined with insights from the model of Becker and Tomes (1979) on parental preferences for child human capital, motivate our specification choice for shared earnings determinants, which reflect the idea that resemblance of earnings across individuals stems from similarities in social background and human capital investments. As shown in Section 4, the life-cycle patterns of earnings correlations between siblings and peers are consistent with these mechanisms.

Besides the earnings profile shared by siblings, neighbors and schoolmates, we assume permanent earnings to follow a unit root in age (ω_{ia}), which captures long-term individual deviations from the shared profile. This represents idiosyncratic ability revealed over time, either to the labor market or to individuals themselves. Overall, our permanent earnings model is specified as follows:

$$y_{ifsna} = \pi_t\left[\left(\mu_f + \mu_s + \mu_n\right) + \left(\gamma_f + \gamma_s + \gamma_n\right)a + \omega_{ia}\right]; \qquad (2)$$

$$\omega_{ia} = \omega_{i(a-1)} + \xi_{ia};\ t = c(i) + 24 + a,$$

where $c(i)$ is the birth cohort of person i and π_t is a calendar time shifter allowing for the possibility of aggregate changes of the permanent earnings process over time.

The parameters of the individual-specific linear profile of earnings are factored into three zero-mean components, with their variances capturing family (*f*), school (s) and neighborhood (*n*) heterogeneity in *initial earnings* (denoted by μ_f, μ_s, μ_n) and life-cycle *earnings growth* (denoted by $\gamma_f, \gamma_s, \gamma_n$). We allow for arbitrary correlation of initial and growth rate heterogeneity within each of the shared components. We also allow for arbitrary correlation across each of the shared components, which is important for taking into account sorting of families across communities (schools and neighborhoods). While previous studies comparing neighbor and sibling correlations have acknowledged the importance of sorting of family into communities (see Page and Solon, 2003; Oreopoulos, 2003; Raaum, Sörensen and Salvanes, 2006), the modeling approach followed in this study is arguably the first attempt to actually estimate these sorting correlations.

The assumptions on the variance-covariance structure of permanent earnings are as follows:

$$\left(\mu_f, \gamma_f\right) \sim \left(\sigma^2_{\mu\Phi}, \sigma^2_{\gamma\Phi}, \sigma_{\mu\gamma\Phi}\right) \tag{3.a}$$

$$\left(\mu_s, \gamma_s\right) \sim \left(\sigma^2_{\mu\Sigma}, \sigma^2_{\gamma\Sigma}, \sigma_{\mu\gamma\Sigma}\right) \tag{3.b}$$

$$\left(\mu_n, \gamma_n\right) \sim \left(\sigma^2_{\mu\mathrm{N}}, \sigma^2_{\gamma\mathrm{N}}, \sigma_{\mu\gamma\mathrm{N}}\right) \tag{3.c}$$

$$\left(\mu_f, \mu_s, \mu_n\right) \sim \left(\sigma_{\mu\Phi\Sigma}, \sigma_{\mu\Phi\mathrm{N}}, \sigma_{\mu\Sigma\mathrm{N}}\right) \tag{3.d}$$

$$\left(\omega_{i24}, \xi_{ia}\right) \sim \left(0,0; \sigma^2_{\omega_{24}b}, \sigma^2_{\xi b}\right), b = 1,2 \text{ ,} \tag{3.e}$$

where idiosyncratic parameters are allowed to vary by birth order (denoted by b in eq. 3(e)), while capital Greek letters indicate the dimension of heterogeneity a variance-covariance parameter refers to (family Φ, school Σ, neighborhood N, or their combination). Correlation across family and community effects is allowed through the intercepts of the individual-specific profiles (assumption (3.d)). This choice is made because empirically most of the community effects vanish after two or three years (see Figure 4), and for not overcrowding the parameter space.

5.2 Specification of transitory earnings

We model transitory earnings (v) in equation (1) to capture any serial correlation of transitory shocks using an autoregressive AR(1) process. We allow brothers to draw shocks from birth-order-specific distributions and we account for age effects in the variance of these shocks through an exponential spline. Our model for transitory earnings can be summarized as follows:

$$v_{ifsna} = \eta_t u_{ifsna};\ u_{ifsna} = \rho_b u_{ifsn(a-1)} + \varepsilon_{ifsna};$$
$$\varepsilon_{ifsna} \sim \left(0, \sigma^2_{\varepsilon b} \exp\left(g_b(a)\right)\right),\ u_{ifsn24} \sim \left(0, \sigma^2_{u_{24}b}\right), \tag{4}$$

where η_t is a time loading factor and u_{ifsna} is the birth-order-specific AR(1) process (note the index b). The autoregressive process begins at age 24 and we specify the variance of the initial condition denoted by $\sigma^2_{u_{24}b}$. The process evolves through the arrival of white noise shocks (denoted by ε) whose variance is age-and-brother-specific ($\sigma^2_{\varepsilon b} \exp\left(g_b(a)\right)$, with $g_b(a)$ denoting a linear spline in age with knots at 28, 33, 38 and 43.

We allow transitory earnings to be correlated across individuals. The specific way in which we model such correlation depends on the type of relationship between individuals. For brothers, the use of birth order specific distributions of shocks enables identifing the contemporaneous correlation of AR(1) innovations. Let i and i' index two individuals; the brother correlation of AR(1) innovation is specified as follows:

$$E\left(\varepsilon_{ifsna}\varepsilon_{i'fs'n'a'}\right) = \sigma_f,\ \ \forall\ s, s', n, n', a = a' \pm |c(i) - c(i')|. \tag{5}$$

That is, when the two individuals belong to the same family and when their age difference is such that the two shocks belong to the same time period, then these shocks are allowed to be correlated with covariance denoted by σ_f. This correlation of shocks between siblings does not depend on whether the two brothers attended the same school, or lived in the same parish when they were aged 15 and is transmitted to non-contemporaneous time periods through the autoregressive structure of the model.

Due to dimensionality issues, we cannot follow a similar approach for modeling the correlation of shocks across community members belonging to different families (f and f'). Instead, we allow for catch-all “mass-points” covariances (λ) collapsing all the parameters of the underlying stochastic processes, and allow such covariances to fade away over time. For any two non-necessarily different age levels a and a', correlations of transitory shock across non-sibling peers are specified as follows:

$$E(u_{ifsna}\, u_{i'f'sna'}) = \lambda_{sn}^{1+|t-t'|}, E(u_{ifsna}u_{i'f'sn'a'}) = \lambda_{s}^{1+|t-t'|} \;\; \forall\, n \neq n', \tag{6}$$

$$E(u_{ifsna}u_{i'f's'na'}) = \lambda_{n}^{1+|t-t'|} \;\forall\, s \neq s'.$$

5.3 Identification of permanent earnings components and decomposition of the sibling correlation

Assumptions (3.a) – (3.e) fully specify the intertemporal and interpersonal distribution of *permanent* earnings.[9] Identification of parameters is achieved by exploiting different types of moment restrictions generated by the model. For a given individual, moment restrictions for two time periods are a function of all sources of earnings heterogeneity, which include the idiosyncratic component, as well as the components due to the influences from the family, the school and the neighborhood. The moment restrictions for a single individual for two non-necessarily different age levels a and a' can be written as follows:

$$E(y_{ifsna}, y_{ifsna'}) = \tag{7}$$

$$\{\sigma_{\mu\Phi}^2 + \sigma_{\mu\Sigma}^2 + \sigma_{\mu\mathrm{N}}^2 + (\sigma_{\gamma\Phi}^2 + \sigma_{\gamma\Sigma}^2 + \sigma_{\gamma\mathrm{N}}^2)aa' + (\sigma_{\mu\gamma\Phi} + \sigma_{\mu\gamma\Sigma} + \sigma_{\mu\gamma\mathrm{N}})(a + a') + 2\sigma_{\mu\Phi\Sigma}$$
$$+ 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \sigma_{\omega_{24}b}^2 + \sigma_{\xi b}^2 min(a, a')\}\pi_t\pi_{t'}$$

[9] Parameter identification of transitory earnings is discussed in the Appendix.

Cross-persons moments (across siblings, neighbors, or schoolmates) do not depend on idiosyncratic heterogeneity. Moment restrictions between siblings (different i but same *f*) depend on the family effects. Moreover, they are also functions of school effects, neighborhood effects, both, or none, depending on the extent to which siblings shared schools and/or neighborhoods.[10] Moment restrictions for siblings can be written as follows:

$$E\left(y_{ifsna}, y_{i'fs'n'a'}\right) = \{\sigma_{\mu\Phi}^2 + \sigma_{\gamma\Phi}^2 aa' + \sigma_{\mu\gamma\Phi}(a+a') + I(s=s')\left[\sigma_{\mu\Sigma}^2 + \sigma_{\gamma\Sigma}^2 aa' + \sigma_{\mu\gamma\Sigma}(a+a')\right] + I(n=n')\left[\sigma_{\mu\mathrm{N}}^2 + \sigma_{\gamma\mathrm{N}}^2 aa' + \sigma_{\mu\gamma\mathrm{N}}(a+a')\right] + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}}\}\, \pi_t \pi_{t'}, \quad (8)$$

where $I(\quad)$ is an indicator function. Equation (8) nests moments restrictions for four types of siblings, corresponding to the four elements of the set generated by intersecting $I(s=s')$ and $I(n=n')$. These types include siblings who: (1) share both the school and the neighborhood; (2) share only the school; (3) share only the neighborhood; and (4) share only the family but neither the school nor the neighborhood.

The above moment conditions are sufficient for identifying family, school, and neighborhood effects, because school and neighborhood effects are identified by the presence of siblings that went to different schools or grew up in different neighborhoods due to family mobility. However, the cross-effects covariances are not identified. This is evident from the fact that the term $2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}}$ enters equation (8) irrespective of whether siblings went to the same school or lived in the same parish. Because families sort across schools and neighborhoods, school and neighborhood effects are always correlated between brothers, and such covariance is not separable from the variance of family effects $\sigma_{\mu\Phi}^2$. To identify the sorting parameters $\sigma_{\mu\Phi\Sigma}$, $\sigma_{\mu\Phi\mathrm{N}}$ and $\sigma_{\mu\Sigma\mathrm{N}}$, we exploit moment restrictions for non-sibling peers that *do not* share the family effect. Using these restrictions is also helpful for estimating community effects without relying exclusively on family mobility across communities.

[10] This is one difference with PSID-based studies (e.g. Page and Solon, 2003) in which all siblings share the neighborhood by sampling design.

Moment restrictions for peers belonging to different families f and f' can be written as follows:

$$E\left(y_{ifsna}, y_{i'f's'n'a'}\right) = \left\{I(s = s')\left[\sigma_{\mu\Sigma}^2 + \sigma_{\gamma\Sigma}^2 aa' + \sigma_{\mu\gamma\Sigma}(a + a') + 2\sigma_{\mu\Phi\Sigma}\right] + I(n = n')\left[\sigma_{\mu\mathrm{N}}^2 + \sigma_{\gamma\mathrm{N}}^2 aa' + \sigma_{\mu\gamma\mathrm{N}}(a + a') + 2\sigma_{\mu\Phi\mathrm{N}}\right] + 2\sigma_{\mu\Sigma\mathrm{N}}\right\} \pi_t \pi_{t'} \tag{9}$$

Equation (9) nests moment restrictions for three types of peers depending on them sharing the school, the neighborhood or both. This identifies the three sorting parameters, where the covariance is zero for those who do not share any community effect. Note that the covariance between family and a given community effect (school or neighborhood) enters the moment restrictions in (8) only for peers sharing that specific effect.

Using parameter estimates from the model we can predict the contributions of each factor to the sibling correlation of permanent earnings over the life-cycle as follows:

$$r^F(a) = \frac{E\left(y_{ifsna}, y_{i'fs'n'a}\right)}{E\left(y_{ifsna}, y_{ifsna}\right)},\ r^S(a) = \frac{E\left(y_{ifsna}, y_{i'f'sn'a}\right)}{E\left(y_{ifsna}, y_{ifsna}\right)},\ r^N(a) = \frac{E\left(y_{ifsna}, y_{i'f's'na}\right)}{E\left(y_{ifsna}, y_{ifsna}\right)}\ , \tag{10}$$

where r denotes correlation coefficients of permanent earnings, F, S and N denote the three relevant dimensions of heterogeneity (family, school, neighborhood). It should be emphasized that correlations vary with age because they are estimated from a model of life cycle earnings. Given the model assumptions, the sibling correlation of permanent earnings is the sum of the three components:

$$r^B(a) = r^F(a) + r^S(a) + r^N(a) \tag{11}$$

5.4 Estimation

The model is estimated by Minimum Distance matching moment restrictions implied by the model to the empirical moments derived from the data.[11] Empirical moments are based on the

[11] Moment restrictions for transitory earnings are given in the Appendix. The orthogonality assumption between permanent and transitory earnings in equation (1) implies that moment restrictions of the full model are the sum of moment restrictions for permanent and transitory earnings. We use Equally Weighted Minimum Distance

residuals after regressing log real gross annual earnings on time dummies and a quadratic age trend by birth cohort. There are three types of empirical moments entering into the estimation. First, there are *individual moments* which include the variances and inter-temporal covariances of individual earnings. Second there are *sibling moments* which are defined only in families where there are at least two brothers. This implies that each family contributes at most one observation in the estimation of sibling empirical moments, where families with only one son do not contribute to such estimation.[12] We estimate separate empirical moments for siblings depending on whether they shared the school, the neighborhood, both or none, so as to match the four different moment restrictions that are nested in equation (8). Finally, there are empirical moments for *non-sibling peers* who shared the community. Differently from the case of families, the numerosity of peers within community clusters do vary. We account for such varying importance of community clusters using the weighting scheme proposed by Page and Solon (2003, pp. 841). In particular, we first estimate the within-cluster covariances and then we take the between-clusters weighted average of within-cluster covariances using weights that are proportional to the number of individuals in that cluster. Similar to the case for siblings, we estimate empirical moments distinguishing whether peers shared the school, the neighborhood, or both.

6. Results

We concentrate our discussion of the results on estimates of the 'core' parameters of the permanent and transitory components.[13] We present the results for the parameter estimates of the permanent component in Section 6.1 and those for the transitory component in Section

which does not weight the minimization problem but adjusts parameter variance post estimation using the empirical fourth moments matrix (see, for example, Haider, 2001).

[12] As explained in the data section, we focus on the first two brothers because third or younger brothers are a tiny proportion (4 percent) in the population of families with more than two male sons.

[13] Parameter estimates of the time effects on both components are reported in Table A1 of the Appendix.

6.2. Sensitivity analysis and heterogeneous effects are discussed in Sections 6.3 and 6.4, respectively.

6.1 Permanent earnings correlation between siblings, schoolmates and neighbors

According to equation (2), permanent earnings depend both on *shared* and *idiosyncratic* components. The parameters estimates of the shared components reported in Table 2 (Panel A) show that family is by far the most relevant factor for long-term earnings. This is true both for initial earnings (intercepts) and for earnings growth rates (slopes). The other relevant source of permanent inequality in earnings is the individual idiosyncratic component reported in Panel B of Table 2.

The covariances across the three components of shared earnings determinants capture the sorting of families into schools and neighborhoods. The estimates in Panel A suggest that these sorting effects are relevant, as the covariance of family effects with both school and neighborhood effects is positive, sizeable and statistically significant. These effects imply that a high draw from the distribution of family effects in permanent earnings is associated with similarly high draws in the distributions of school and neighborhood effects. We also find a positive covariance among community effects, which suggests that school and neighborhood effects are positively correlated. Once these sorting effects are taken into account there is no remaining statistically significant heterogeneity in initial earnings related to school and neighborhood effects. Estimating the model imposing zero cross-component covariances, in other words ignoring sorting, yields estimates of the intercepts variances of community effects that are statistically significant and of about the same size as the covariances in the unconstrained model.

All shared components of long-term earnings in Table 2 display the Mincerian cross-over property. This is apparent by noting that all covariances between intercepts and slopes of earnings profiles are negative. This indicates that families associated with low earnings at age 24 are also associated with faster growth in life-cycle earnings. This implies that the variance of permanent earnings across families is *u-shaped in age* because it falls in the years of catch up and increases after the cross over point. The point of cross over can be computed as the year in which the earnings variance is minimized, and it is located at age 34 for the between-families earnings distribution. A similar u-shape pattern of the variance of earnings over age is also observed across schools and across neighborhoods. The cross over point is age 36 for the between-neighbors earnings distribution, and age 38 for the between-schools earnings distribution.

We use these parameter estimates to generate predictions, based on the formulae provided in Section 5.3 (equations 10 and 11), of the sibling correlation and its decomposition into the three factors of interest: family, school and neighborhood. In particular, we consider the case of two brothers who attended the same school and lived in the same neighborhood when they were 15, so that the resulting sibling correlation is the sum of family, school and neighborhood effects. As shown in Figure 4, the life-cycle pattern of the sibling correlation is u-shaped in age. More specifically, the estimated correlation is about 0.6 at age 24, drops to 0.15 at age 37, and rises back to 0.34 by age 51, which is the last age we observe younger brothers. The average sibling correlation is 0.28 (s.e. 0.012), which is in line with previous estimates for Denmark.[16] As mentioned earlier, the u-shaped pattern is a symptom of "Mincerian cross-overs" of earnings profiles. That is, the negative estimates of the covariance between intercepts and slopes for all the shared factors of earnings profiles imply that the distribution of shared components, and therefore the sibling and peer correlations, first shrink and then fan out over the life-cycle. The same u-shaped pattern was also a feature of the raw cross-person covariances in Figures 1 to 3, and in particular Figure 2, which depicted siblings' earnings covariances for brothers with few years of age spacing.

[16] Using a model without community effects, Bingley and Cappellari (2013) report an average sibling correlation of 0.23 between ages 25 and 48. Using our sample to estimate a model without community effect in the 25-48 age range we obtain an average sibling correlation of 0.25. Differences between the two estimates are due to the different age range investigated, different specifications and different sample selections.

Considering the decomposition of the sibling correlation in Figure 2, it is evident that family accounts for most of the dispersion of permanent earnings over the life-cycle. The community effects are very small and are only significant at the beginning of the life-cycle, while by age 30 they become negligible and not significantly different from zero. On average, over the life-cycle, we estimate the correlation in permanent earnings across schoolmates to be 0.004 (s.e. 0.010), and across neighbors to be 0.009 (s.e. 0.010). These results indicate that the only factor that generates a substantial correlation in permanent earnings between brothers is the family. Instead, there is not much room for community effects in shaping the sibling correlation.

Our findings are in line with those of Oreopoulos (2003) who used quasi-random assignment of neighbors and showed that the neighbors correlation in adult earnings was virtually zero in a variance decomposition exercise similar in spirit to ours. Page and Solon (2003), instead, found in the PSID that the neighbors correlation (0.16) was about half of the sibling correlation (0.34). By formulating a model for the joint estimation of family and community effects, allowing for sorting of families into communities, we can replicate the previous approaches by estimating community effects on a sample that excludes siblings and constraining family-related model parameters to be zero. The idea of this exercise is that, by ignoring family effects and their sorting into communities, community effects would capture not only the effects of communities but will also pick up the influence of families. The results of this exercise are reported in Figure 5 in which we plot community effects (the sum of school and neighborhood effects) from the model that *ignores* family effects and siblings data, alongside community effects estimated from our full model. The comparison is striking. When family effects are ignored, we find a sizeable correlation among members of the youth community, which is significant throughout the life-cycle. The average correlation in this model is 0.071 (s.e. 0.001), which amounts at 25% of the sibling correlation.[17] As we have seen in Figure 4, the model that controls for family effects tells a radically different story about the relevance of community effects, with a correlation of permanent earnings between members of the same youth community of just 0.013 (=0.004+0.009, s.e. 0.009) and insignificant. This suggests that when sorting is not allowed for the community effects are upward biased by a factor of five. The results of the full model are close to those from quasi-randomized variation of families across communities that aim to control for that type of selection.

6.2 Transitory earnings

Parameter estimates of transitory earnings in Table 3 show a clear age pattern of transitory shocks, whose variance decreases between the mid-20s and the mid-30s, while the decrease slows down after age 35. The sharp decline followed by a leveling-off is consistent with the patterns reported by Baker and Solon (2003) who find the variance of transitory shocks to be declining at decreasing rates between the ages of 25 and 45. These patterns look similar between brothers. Also, the autoregressive coefficient (roughly 0.5) is very similar between brothers and of a moderate size. Table 3 shows that transitory shocks are contemporaneously correlated between brothers. However, compared to the variance of the shocks, the size of the covariance is negligible. The model also yields estimates of the covariance in transitory earnings between non-relative peers, which turns out to be negligible and imprecisely estimated.

6.3 Sensitivity analysis

We subject our results to several sensitivity checks by estimating the model for different family sizes (up to 2 or up to 3 children), excluding singletons, and by varying the degree of

[17] The ratio between neighbor and sibling correlations was 0.47 in Page and Solon (2003).

exposure to communities. We report in Table 4 the average sibling correlation and its decomposition by family, school and neighborhood factors. Overall, the findings from these additional estimations do not change the main conclusion from the baseline model that family accounts for most of the variation in permanent earnings, while the influence of the community factors is overall negligible and only significant early in the working life and before age 30.

More specifically about the definition of youth communities, the concern with the baseline model might be that it is based on membership only at age 15, which might miss part of the effects of communities due to potentially limited exposure (see also Gibbons, Silva, Weinhardt, 2013 and Chetty et al., 2015 for similar discussions). To address this concern, we re-estimated the model using two alternative criteria to define community membership, which are characterized by greater exposure to communities relative to the one-year definition used in the baseline model. First, we define schoolmates and neighbors as individual sharing schools and neighborhoods, respectively, for two years during both ages 14 and 15. Second, we define the neighborhood as the prevalent parish of residence between ages 14 and 18.[18] As reported in Table 4, none of these alternative definitions alters our finding that community effects account for only a limited share of the sibling correlation in earnings. Defining peers as those sharing schools and neighborhoods both at age 14 and 15 yields an average correlation of permanent earnings between schoolmates equal to 0.002 (s.e. 0.009), and an average correlation between neighbors equal to 0.010 (s.e. 0.010). Similarly, using the parish in which individuals lived most frequently between the ages of 14 and 18 as identifier of youth neighborhoods, we find the average earnings correlation among neighbors to be 0.008 (s.e. 0.010), and the correlation among schoolmates to be 0.006 (s.e. 0.009). Based on this evidence it seems plausible to conclude that our finding of negligible community effect is not

[18] More than three quarters of individuals in our sample (76.5%) do not change parish of residence between ages 14 and 18, and an additional 22% changed parish of residence only once or twice. We cannot apply a similar definition to schools because of compulsory schooling ending typically when individuals are aged 15.

driven by the specific community definition that we adopt. However, due to data limitations we are not able to consider school and neighborhood effects at ages earlier than 14, so we cannot test for differential exposure effects at a younger age. However, we argue that at least for the school effects this is less likely to be the case because of the coherence of the schoolmate groups in Denmark since primary and secondary education usually takes place in the same school and most pupils attend the same school for all grades.

6.4 Heterogeneous effects

We now turn to potential heterogeneity by the type of family, school or neighborhood. For families, we distinguish between high and low educated fathers; for schools between large and small classes; and for neighborhoods between high and low density areas (urban vs. rural).

Starting with family heterogeneity, Figure 6a shows that the share of the variation in permanent earnings accounted for by the family is much higher among families with a high educated father compared to a low educated father. This is not surprising as families with higher education are more likely to transmit resources to their children which influence their earnings capacity, consistent with the idea that intergenerational transmission is stronger at the top of the income distribution.

With regard to school heterogeneity, we split observations depending upon whether school enrollment in the grade attended at age 15 was above or below the threshold of 24 pupils used in Denmark for splitting classes. Pupils who attended schools at age 15 with total enrollment below 24, between 37 and 48, between 61 and 72, and so on, were grouped in the "Large Class" group. On the other hand, pupils who attended schools at age 15 with enrollment between 25 and 36, between 49 and 60 and so on, were grouped in the "Small Class" sample. This classification is based on the fact that larger cohorts exceeding the class size threshold were split into smaller classes. Figure 6b shows that the school effect is much higher for small classes compared to large classes. This finding is consistent with the literature on the effect of class size on earnings, which suggests a positive effect from smaller classes (e.g. Chetty et al., 2011; Fredriksson Öckert and Oosterbeek, 2013). However, the school effect is not persistent and becomes insignificant after age 30. This suggests that although schools resources seem to play a larger role at the beginning of the working life, these effects are not long lasting.

The last dimension of heterogeneity that we take into account is urbanicity by exploiting information on population density (measured in 1976) in the parishes individuals lived in when they were 15. Specifically, we cut the density distribution across parishes at the upper third, and define urban individuals living in parishes that are above this threshold, and rural all remaining individuals. Figure 6c shows that early in the working life higher school effects for urban and neighborhood effects for rural pupils. However, similar with the school effects, in both cases the effects vanish after age 30.

7. Conclusion

This paper develops a unified framework which enables disentangling the contribution of families, schools and neighborhoods in labor earnings over the life-cycle. This is achieved within a model of multi-person earnings dynamics distinguishing permanent from transitory earnings and allowing for heterogeneous earnings growth. The analysis is based on administrative registers from the Danish population which we use to link brothers, schoolmates and teenage neighbors and follow them over their life-cycle and up to age 51.

Our analysis indicates that family is by far the most relevant factor that shapes long-term earnings. The contribution coming from schools and neighborhoods on long-term earnings is overestimated if the family component is ignored, which suggests that not accounting for sorting leads to an upward bias in the estimated influence of community background. Despite the negligible average community effects, we find that both schools and neighborhoods exhibit a positive and significant effect at the beginning of working life. However, these effects are not long-lasting as by age 30 they become close to zero and insignificant. These results contribute to our understanding about the effects of family and community background on labor market outcomes showing that while family influences are long-term, community influences do not have very long-term consequences lasting beyond age 30. This has implications for the design of policies aiming at reducing inequalities in the long-run suggesting that resources aimed at improving the situation of families are likely to be more effective in the very long-term than resources devoted to transforming communities.

References

Baker, Michael, and Gary Solon. 2003. “Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records.” *Journal of Labor Economics* 21 (2): 267-88.

Becker, Gary S., and Nigel Tomes. 1979. “An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility.” *Journal of Political Economy* 87 (6): 1153-89.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings.” *Journal of Political Economy* 75 (4): 352-65.

Bingley, Paul, and Lorenzo Cappellari. 2013. “Correlations of Brothers’ Earnings and Intergenerational Transmission.” Dipartimento di Economia e Finanza, Università Cattolica, Working Paper No. 6.

Björklund, Anders, and Markus Jännti. 2009. “Intergenerational Income Mobility and the Role of Family Background.” in *Oxford Handbook of Economic Inequality*, edited by Wiemer Salverda, Brian Nolan, and Timothy Smeeding, 491-521. Oxford: Oxford University Press.

Black, Sandra E., Paul J. Devereux. 2011. “Recent Developments in Intergenerational Mobility.” in *Handbook of Labor Economics, Vol. 4A*, edited by Orley Ashenfelter and David Card, 1487-541. Amsterdam: Elsevier Science, North Holland.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. “The More the Merrier? The Effect of Family Size and Birth Order on Children's Education.” *The Quarterly Journal of Economics* 120 (2): 669-700.

Card, David, and Alan B. Krueger. 1992. “Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States.” *Quarterly Journal of Economics* 100 (1): 1-40.

Chetty, Raj, Nathalien Hendren, and Lawrence F. Katz. 2015. “The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment” NBER Working Paper No. 21156.

Chetty, Raj, John N, Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR.” *Quarterly Journal of Economics* 126 (4): 1593-660.

Dearden, Lorraine, Javier Ferri, and Costas Meghir. “The Effect of School Quality on Educational Attainment and Wages.” 2002. *Review of Economics and Statistics* 84: 1-20.

Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2013. "Long-Term Effects of Class Size." *Quarterly Journal of Economics* 128 (1): 249-85.

Gibbons, Stephen, Olmo Silva, and Felix Weinhardt. 2013. "Everybody Needs Good Neighbours? Evidence from Students' Outcomes in England." *Economic Journal* 123: 831-74.

Gould, Eric D., Victor Lavy, and M. Daniele Paserman. 2011. "Sixty Years after the Magic Carpet Ride: The Long-Run Effect of the Early Childhood Environment on Social and Economic Outcomes" *Review of Economic Studies* 78: 938-73.

Haider, Steven J. 2001. "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991." *Journal of Labor Economics* 19 (4): 799-836.

Haider, Steven J., and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96 (4): 1308-20.

Hanushek, Erik A. 2006. "School Resources" in *Handbook of the Economics of Education, Volume* 2, edited by Eric A. Hanushek, and Finis Welch, 865-908. Amsterdam: Elsevier Science, North Holland.

Jacob, Brian. 2004. "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review* 94 (1): 233-58.

Jenkins, Stephen J. 1987. "Snapshots vs. Movies: 'Lifecycle biases' and the Estimation of Intergenerational Earnings Inheritance." *European Economic Review* 31 (5): 1149–58.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jefrey R. Kling, and Lisa Sanbonmatsu. 2013. "Long-term Neighborhood Effects on Low-Income Families: Evidence form Moving to Opportunity." *American Economic Review: Papers & Proceedings* 103 (3): 226–31.

Mincer, Jacob. 1958. "Investment in Human Capital and Personal Income Distribution" *Journal of Political Economy* 66 (4): 281-302.

Oreopoulos, Philip. 2003. "The Long-Run Consequences of Living in a Poor Neighborhood" *Quarterly Journal of Economics* 118 (4): 1533-75.

Page, Marianne, and Gary Solon. 2003. "Correlations between Brothers and Neighboring Boys in their Adult Earnings: The Importance of Being Urban." *Journal of Labor Economics* 21 (4): 831-56.

Raaum, Oddbjörn, Erik Sörensen, and Kjell G. Salvanes 2006. "The Neighbourhood Is Not What It Used to Be" *Economic Journal* 116 (1): 200-22.

Solon, Gary. 1999. “Intergenerational Mobility in the Labor Market” in *Handbook of Labor Economics, Vol. 3A*, edited by Orley Ashenfelter, and David Card, 1761-1800. Amsterdam: Elsevier Science, North Holland.

Wilson, William J. 1987. “The Truly Disadvantaged.” Chicago: Chicago University Press.

Table 1

Cohorts included in the sample

Birth Year	First Year Observed	Number of Years Observed	Last Age Observed
1960-61	1984	28	51
1962-63	1986	26	49
1964-65	1988	24	47
1966-67	1990	22	45
1968-69	1992	20	43
1970-71	1994	18	41
1972-73	1996	16	39
1974-75	1998	14	37
1976-77	2000	12	35
1978-79	2002	10	33
1980-81	2004	8	31
1982-83	2006	6	29

Table 2

Parameter estimates of permanent earnings

Panel A - Shared components (heterogeneous income profile –random growth)

	Coef.	s.e.
Variance of intercepts		
Family ($\sigma^2_{\mu\Phi}$)	0.0633	0.0109
School ($\sigma^2_{\mu\Sigma}$)	0.0022	0.0031
Neighborhood ($\sigma^2_{\mu\mathrm{N}}$)	0.0034	0.0035
Variance of slopes		
Family ($\sigma^2_{\gamma\Phi}$)	0.0003	0.00006
School ($\sigma^2_{\gamma\Sigma}$)	0.00003	0.00002
Neighborhood ($\sigma^2_{\gamma\mathrm{N}}$)	0.0001	0.00002
Covariance intercepts-slopes		
Family ($\sigma_{\mu\gamma\Phi}$)	-0.0039	0.0006
School ($\sigma_{\mu\gamma\Sigma}$)	-0.0005	0.0002
Neighborhood ($\sigma_{\mu\gamma\mathrm{N}}$)	-0.0010	0.0003
Covariance between components		
Family-School ($\sigma_{\mu\Phi\Sigma}$)	0.0037	0.0012
Family-Neighborhood ($\sigma_{\mu\Phi\mathrm{N}}$)	0.0037	0.0013
School- Neighborhood ($\sigma_{\mu\Sigma\mathrm{N}}$)	0.0011	0.0002

Panel B - Idiosyncratic components (restricted income profile-random walk)

	Coef.	s.e.
Initial condition (age 24)		
Brother 1 ($\sigma^2_{\omega 24,1}$)	0.0542	0.0091
Brother2 ($\sigma^2_{\omega 24,2}$)	0.0374	0.0067
Variance of innovations		
Brother 1 ($\sigma^2_{\xi 1}$)	0.0066	0.0011
Brother 2 ($\sigma^2_{\xi 2}$)	0.0071	0.0012

Table 3

Parameter estimates of transitory earnings

	Coef.	s.e.
Initial condition (age 24)		
Brother 1 ($\sigma^2_{24,1}$)	0.6613	0.0447
Brother 2 ($\sigma^2_{24,2}$)	0.6477	0.0459
Variance of innovations at 25		
Brother 1 ($\sigma^2_{\varepsilon 1}$)	0.4935	0.0357
Brother 2 ($\sigma^2_{\varepsilon 2}$)	0.4731	0.0341
Age splines in variance of innovations		
Brother 1		
26-28	-0.1370	0.0083
29-33	-0.1016	0.0058
34-38	-0.0244	0.0076
39-43	-0.0358	0.0100
44-51	-0.0153	0.0110
Brother 2		
26-28	-0.1515	0.0089
29-33	-0.1122	0.0066
34-38	-0.0364	0.0092
39-43	-0.0184	0.0125
44-51	0.0080	0.0175
Autoregressive coefficient		
Brother 1 (ρ_1)	0.4979	0.0049
Brother 2 (ρ_2)	0.5164	0.0053
Cross-person associations in transitory earnings		
Sibling covariance of innovations (σ_f)	0.0072	0.0006
Peers covariance of transitory earnings (catch-all components)		
Sharing both school and neighborhood (λ_{sn})	-0.0003	0.0006
Sharing only school (λ_s)	0.0027	0.0007
Sharing only neighborhood (λ_n)	-0.0006	0.0007

Table 4

Sensitivity analysis – Decomposition of sibling correlation (Average)

	Sibling	Family	Neighborhood	School	Community (N+S)
Baseline	0.282	0.269	0.009	0.004	0.013
	(0.012)	(0.012)	(0.010)	(0.010)	(0.009)
Families with up to 2 Children	0.329	0.311	0.021	-0.004	0.018
	(0.016)	(0.022)	(0.014)	(0.013)	(0.013)
Families with up to 3 Children	0.293	0.292	0.020	-0.019	0.001
	(0.014)	(0.018)	(0.015)	(0.013)	(0.012)
Excluding singletons	0.286	0.275	0.008	0.003	0.011
	(0.012)	(0.012)	(0.009)	(0.009)	(0.009)
Peers at age 14 and age 15	0.282	0.270	0.010	0.002	0.012
	(0.012)	(0.012)	(0.010)	(0.009)	(0.009)
Main parish of residence (age 14 -18)	0.280	0.267	0.008	0.006	0.014
	(0.012)	(0.012)	(0.010)	(0.009)	(0.009)

Figure 1

Sibling correlation of annual earnings

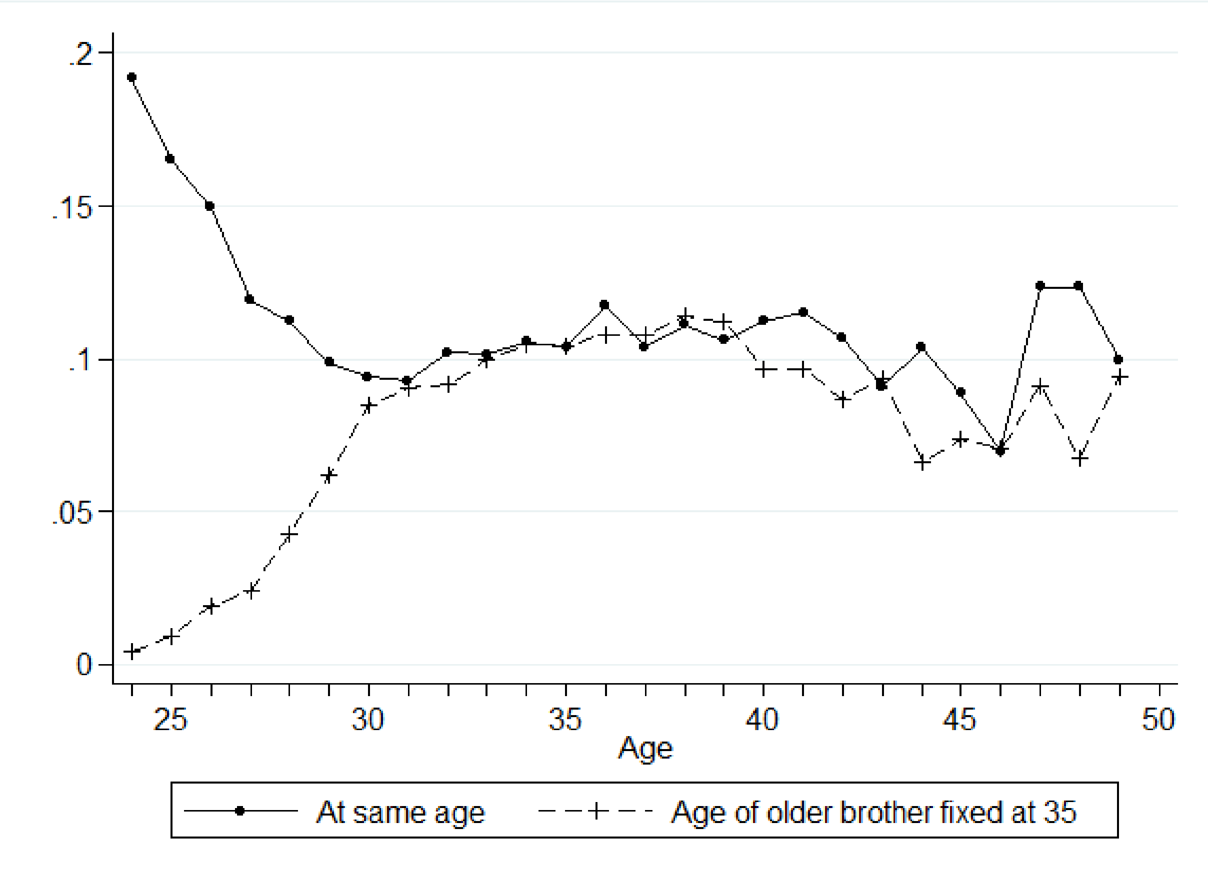


Figure 2

Sibling correlation of annual earnings by siblings' age gap

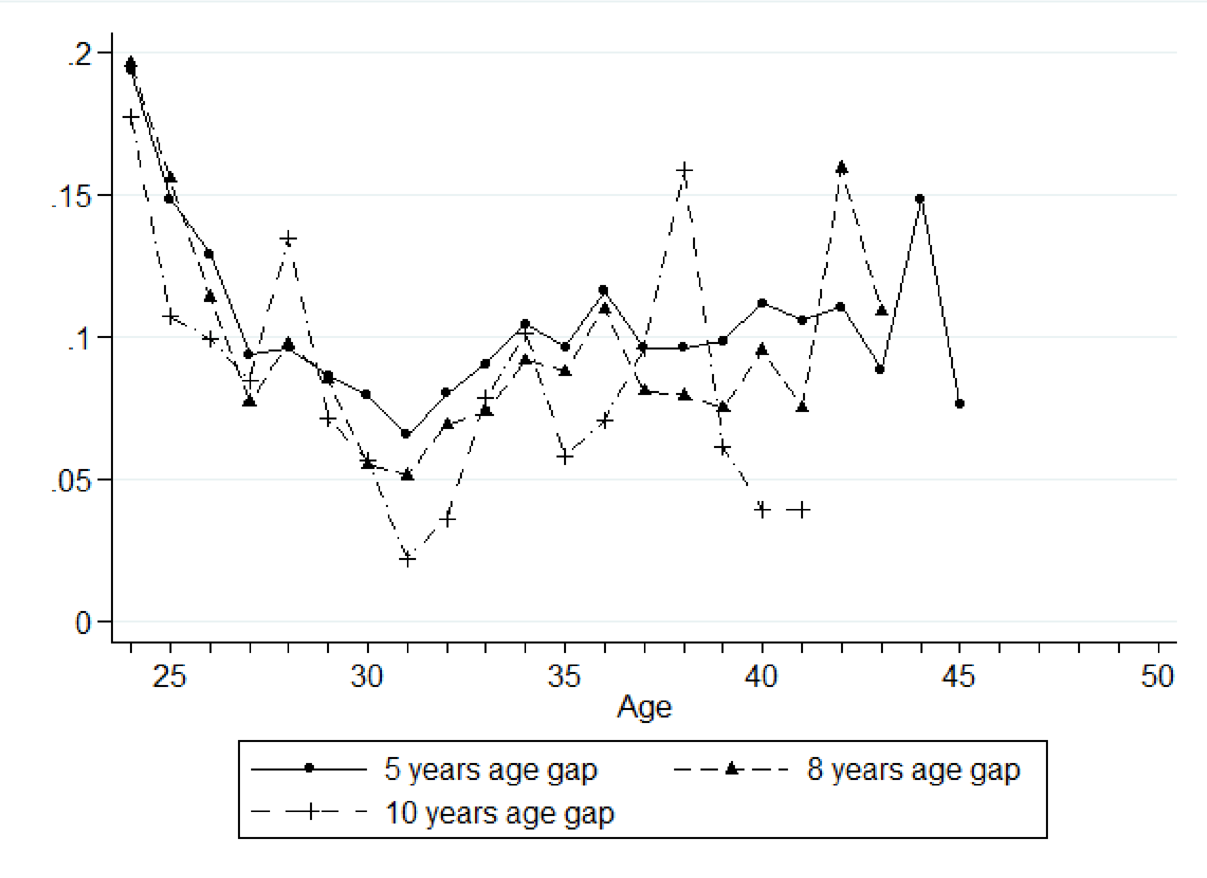


Figure 3

Correlation of annual earnings for members of youth communities

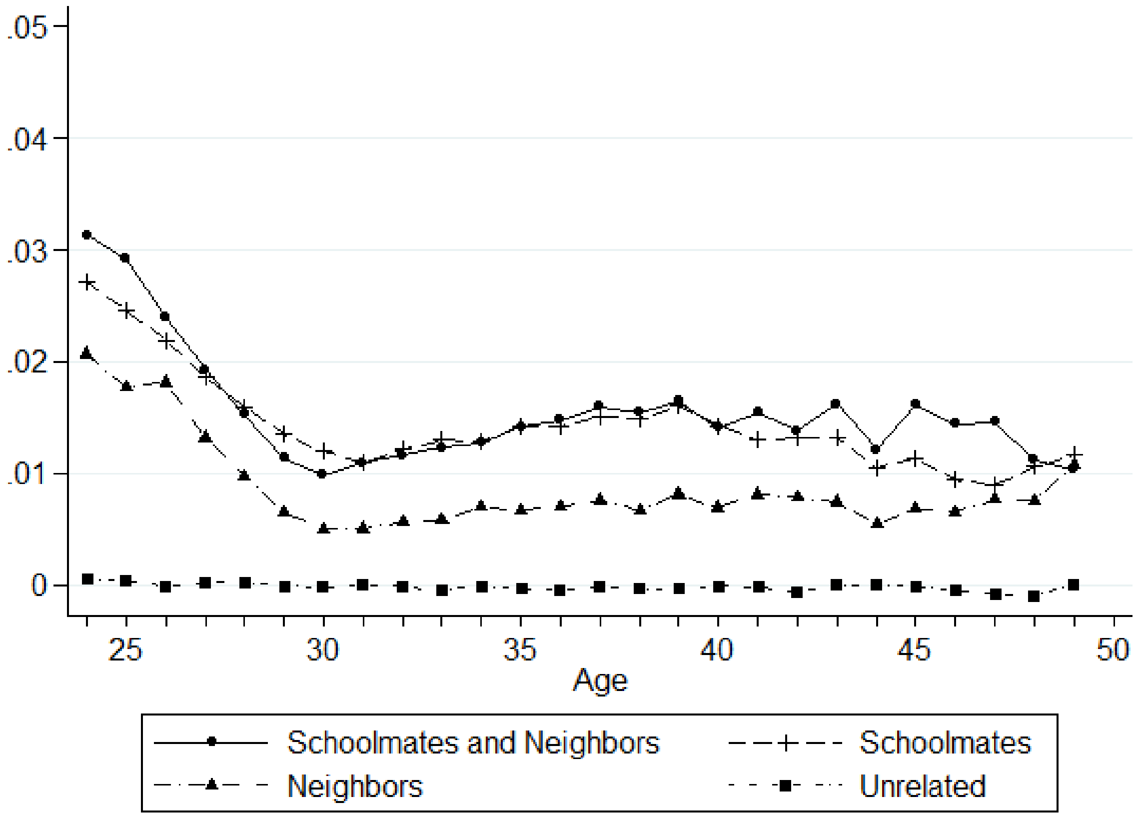


Figure 4

Predicted sibling correlation of permanent earnings and factor decomposition

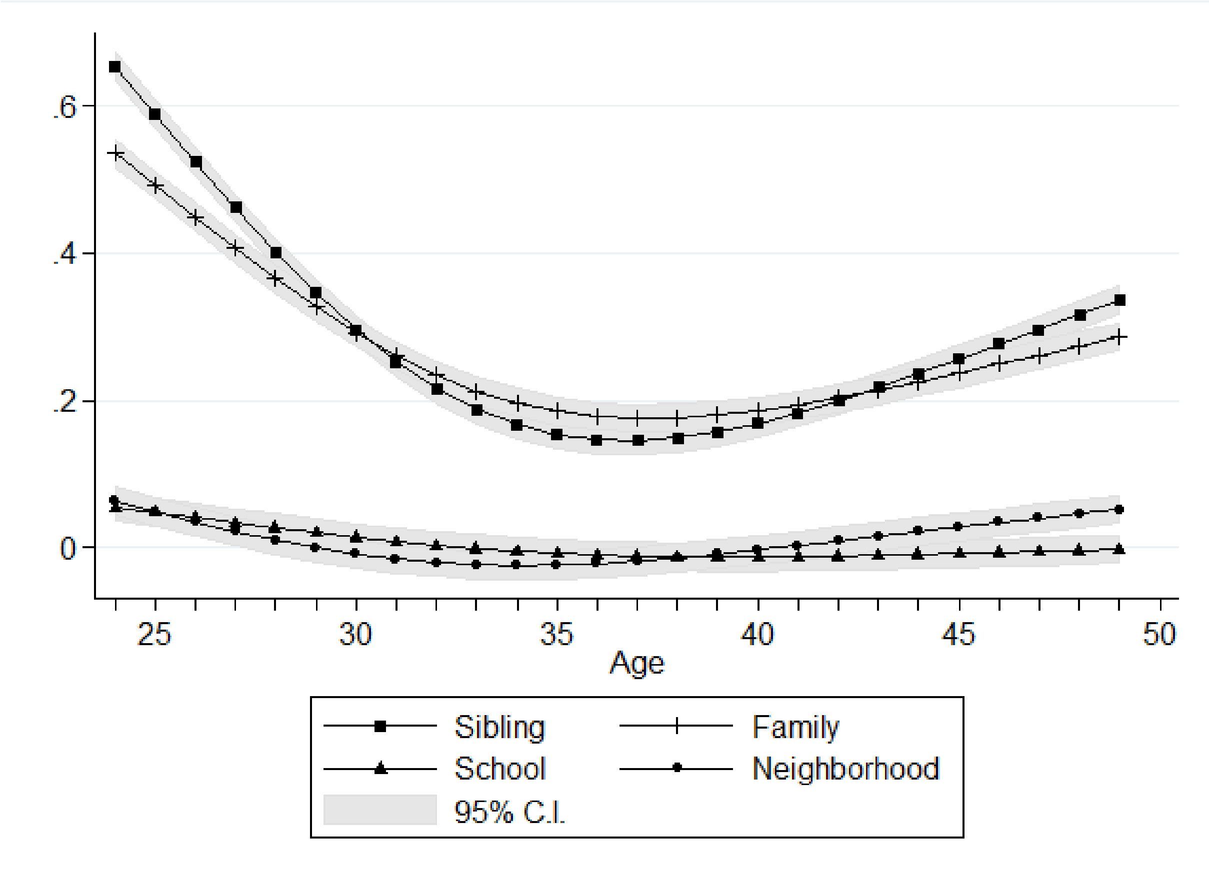


Figure 5

Predicted correlations of permanent earnings between members of youth communities
Comparison of models with and without family effects

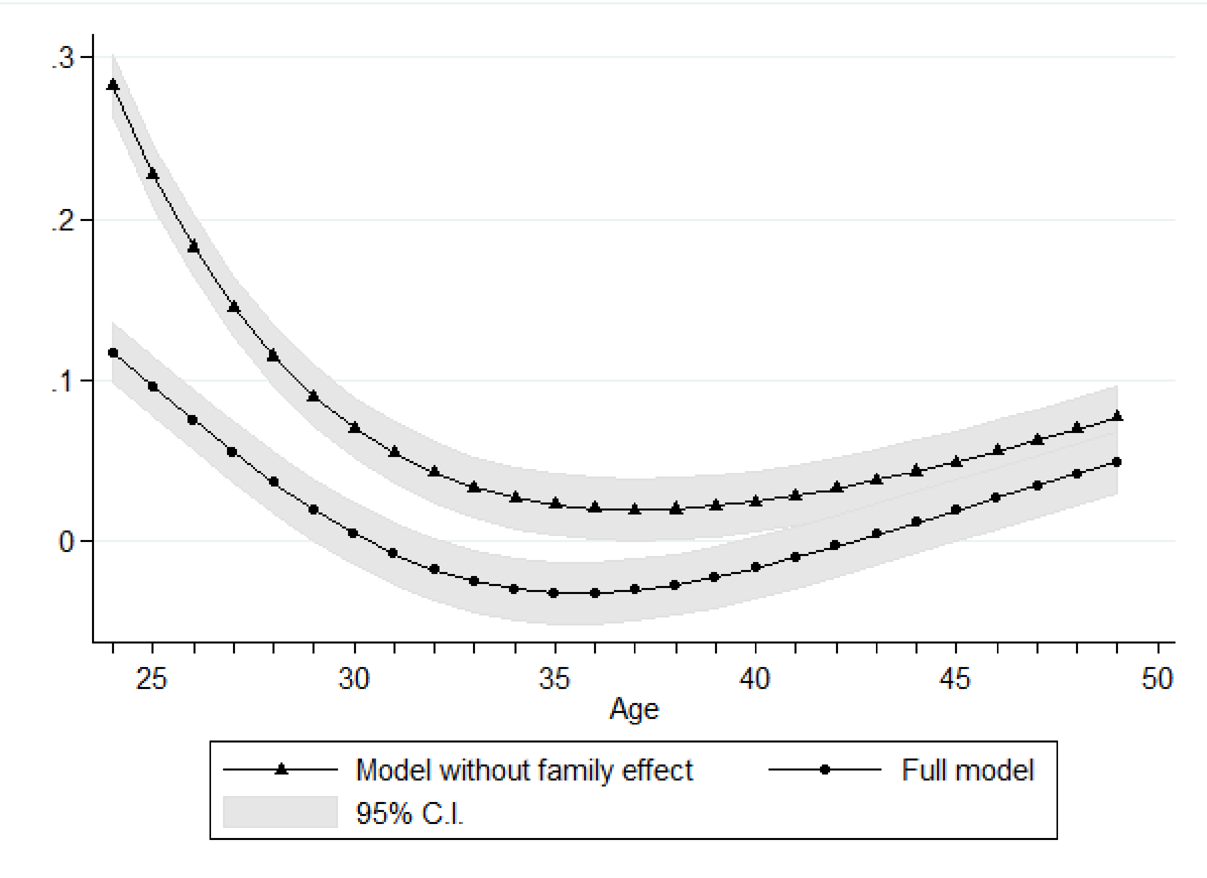


Figure 6a

Predicted correlations of permanent earnings by father's education

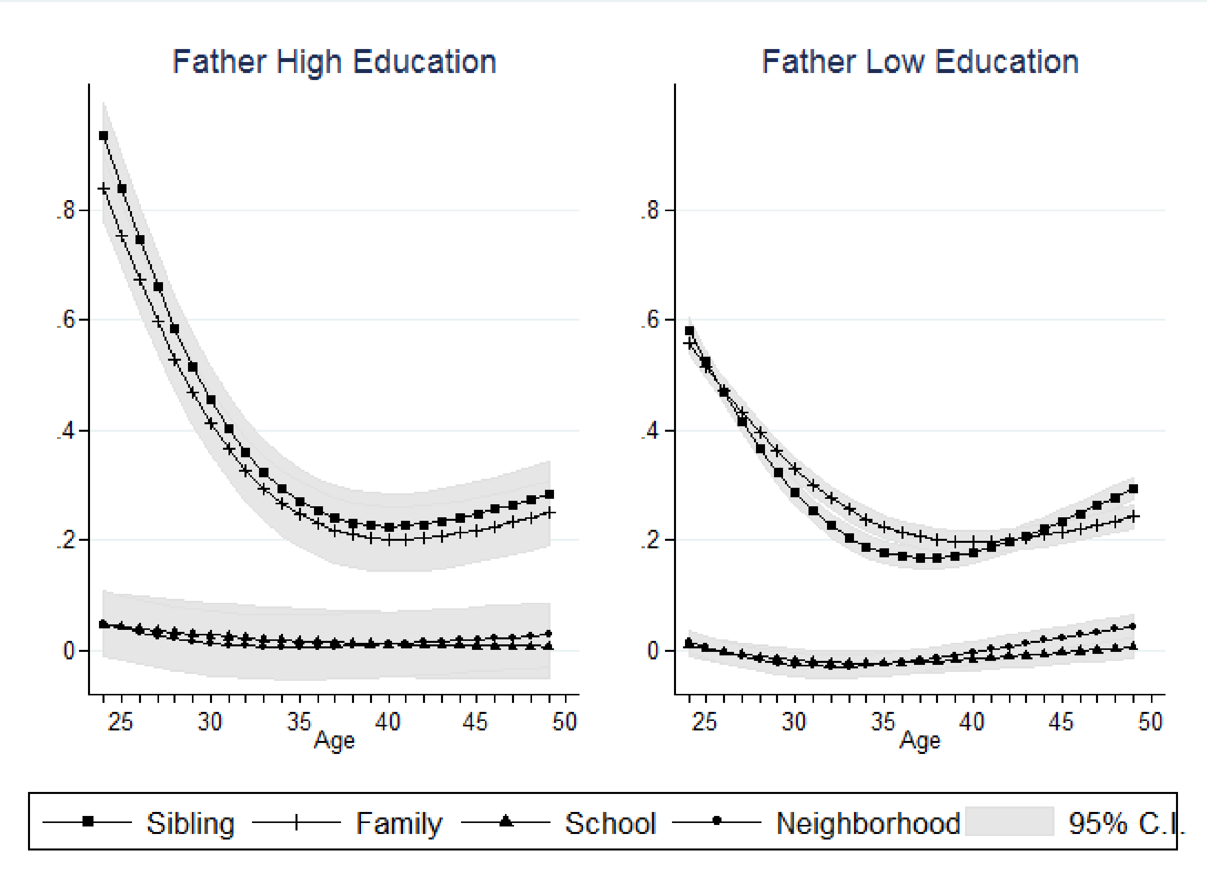


Figure 6b

Predicted correlations of permanent earnings by class size

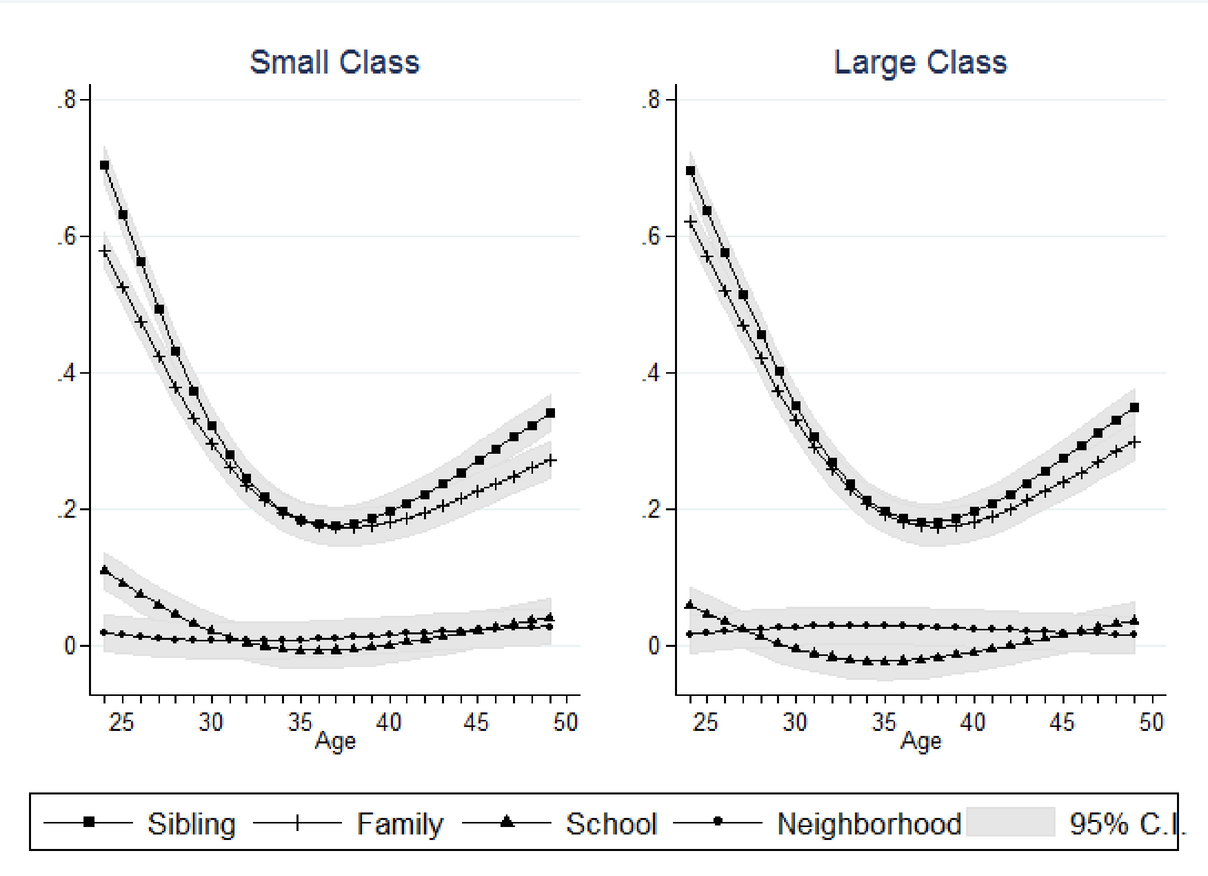
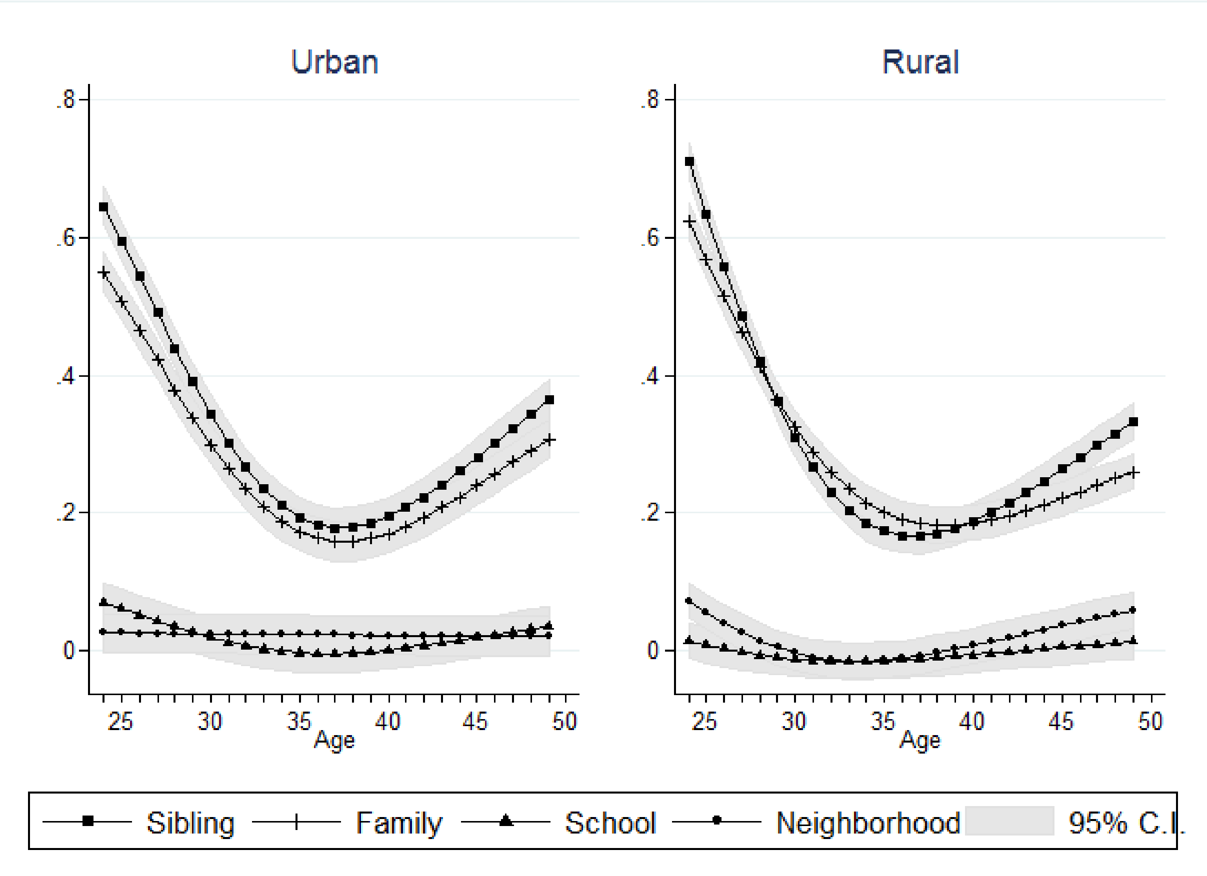


Figure 6c

Predicted correlations of permanent earnings by neighborhood density



Appendix A

Moment restrictions for transitory earnings

Considering two non-necessarily different age levels a and a', the intertemporal covariance structure of the transitory component of *individual* earnings from the birth order specific AR(1) process is as follows:

$$\begin{aligned} E\left(v_{ifsna}v_{ifsna'}\right) = [&I(a = a' = 24)\sigma_{24b}^2 \\ &+ I(a = a' > 24)\left(\exp\left(g_b(a)\right) + var\left(u_{ifsn(a-1)}\right)\rho_b^2\right) \\ &+ I(a \neq a')\left(E\left(u_{ifsn(a-1)}u_{ifsna'}\right)\rho_b\right)]\eta_t\eta_{t'}. \end{aligned} \tag{A.1}$$

Allowing for correlation of AR(1) innovations across brothers, the model yields restrictions on transitory earnings also for cross-brothers moments:

$$E\left(v_{ifsna}v_{i'fs'n'a'}\right) =$$

$$\sigma_f\left(\frac{\left(1-\left(\rho_1\rho_2^{|t-t'|}\right)^P\right)}{1-\rho_1\rho_2^{|t-t'|}}\right)^{I(t\leq t')}\left(\frac{\left(1-\left(\rho_2\rho_1^{|t-t'|}\right)^P\right)}{1-\rho_2\rho_1^{|t-t'|}}\right)^{I(t>t')}\eta_t\eta_{t'};\ \ \forall\, s,s',n,n', \tag{A.2}$$

where *P* is the number of overlapping years the two brothers are observed in the data.

We also model the correlation of transitory earnings across *non-sibling peers*. Differently from the case of brothers, we do not model the correlation of AR(1) innovations among peers because it would require distinguishing idiosyncratic components of transitory earnings for each member of school or neighborhood clusters, generating dimensionality issues. We, therefore, collapse all the cross-peers covariance structure of the transitory component into catch-all “mass point” factors absorbing all the parameters of the underlying stochastic process. For any two non-necessarily different age levels a and a', correlations of transitory earnings across non-sibling peers are as follows:

$$E\left(v_{ifsna}, v_{i'f'sna'}\right) = \lambda_{sn}^{1+|t-t'|}\eta_t\eta_{t'} \tag{A.3}$$

$$E\left(v_{ifsna}, v_{i'f'sn'a'}\right) = \lambda_s^{1+|t-t'|}\eta_t\eta_{t'}\ \ \forall\, n \neq n'$$

$$E\left(v_{ifsna}, v_{i'f's'na'}\right) = \lambda_n^{1+|t-t'|}\eta_t\eta_{t'}\ \ \forall\, s \neq s'$$

The moment restrictions above characterize the inter-temporal distribution of transitory earnings for each individual and between siblings and peers. The orthogonality assumption between permanent and transitory earnings in equation (1) implies that moment restrictions of the full model are the sum of moment restrictions for permanent and transitory earnings, the former being discussed in Section 5.3 of the paper. In general, these restrictions are a non-linear function of a parameter vector θ. We estimate θ by Minimum Distance (see Chamberlain, 1984; Haider, 2001). We use Equally Weighted Minimum Distance (EWMD) and a robust variance estimator $Var(\theta) = (G'G)^{-1}G'VG(G'G)^{-1}$, where V is the fourth moments matrix and G is the gradient matrix evaluated at the solution of the minimization problem.

Table A1: Parameter estimates of time effects (1984=1)

	Permanent Component (π_t)		Transitory Component (η_t)	
	Coeff	se	Coeff	se
t=				
1985	0.9212	0.0807	0.9492	0.0209
1986	0.9129	0.0830	0.9755	0.0231
1987	0.9653	0.0890	0.9591	0.0245
1988	1.0513	0.0960	0.9972	0.0248
1989	1.0356	0.0963	1.0511	0.0264
1990	1.1306	0.1027	1.0688	0.0265
1991	1.2037	0.1083	1.0692	0.0271
1992	1.1435	0.1016	1.1319	0.0271
1993	1.1594	0.1044	1.1397	0.0280
1994	1.2067	0.1070	1.1285	0.0274
1995	1.1561	0.1031	1.0581	0.0263
1996	1.2260	0.1078	1.0605	0.0261
1997	1.1714	0.1024	1.0543	0.0257
1998	1.2097	0.1053	1.0442	0.0254
1999	1.1890	0.1043	1.0723	0.0261
2000	1.2250	0.1073	1.0935	0.0265
2001	1.1845	0.1036	1.1162	0.0269
2002	1.2459	0.1092	1.1406	0.0276
2003	1.2526	0.1100	1.2066	0.0291
2004	1.2350	0.1085	1.1690	0.0281
2005	1.1710	0.1029	1.1628	0.0280
2006	1.1024	0.0971	1.1335	0.0270
2007	1.0175	0.0898	1.1243	0.0268
2008	0.9650	0.0856	1.1467	0.0276
2009	0.9457	0.0840	1.3173	0.0312
2010	0.9263	0.0826	1.3721	0.0325
2011	0.9135	0.0810	1.3731	0.0324