# Family, Community and Long-Term Earnings Inequality§


**Paul Bingley**
SFI Copenhagen

**Lorenzo Cappellari**
Università Cattolica Milano and IZA

**Konstantinos Tatsiramos**
University of Nottingham and IZA


**July 2016**


## Abstract

This paper studies the influence of family, schools and neighborhoods on life-cycle earnings inequality. We develop an earnings dynamics model linking brothers, schoolmates and teenage parish neighbors using population register data for Denmark. We exploit differences in the timing of family mobility and the partial overlap of schools and neighborhoods to separately identify sorting from community and family effects. We find that family is far more important than community in influencing earnings inequality over the life cycle. Neighborhoods and schools influence earnings only early in the working life and this influence falls rapidly and becomes negligible after age 30.



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

JEL codes: D31, J62


---


§ We thank Anders Björklund, Stephen Gibbons, Markus Jännti, Bash Mazumder, Andreas Peichl, Josef Zweimüller, seminar audiences at IZA in Bonn, SOFI in Stockholm, Lund University, Paris School of Economics, JRC in Ispra, University of Padua, University of Venice, ZEW in Mannheim, European University Institute in Florence, IAB in Nuremberg, and conferences audiences at Royal Economic Society in Manchester, European Society of Population Economics in Braga, 16th IZA/CEPR European Symposium in Labour Economics (ESSLE), Fourth SOLE/EALE World Conference in Montreal, DynIPer in Marseille and the Intergenerational Mobility workshop in Madrid for their comments. The usual disclaimer applies. The research for this paper was supported by a British Academy Research Grant (award no. SG113192) and by Danish Strategic Research Council Grants (codes DSF-09-070295 and DSF-10-093105).


Addresses for correspondence: Bingley, SFI, The Danish National Centre for Social Research, Herluf Trolles Gade 11, DK-1052, Copenhagen, pab@sfi.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

That the environment in which persons grow up and live in the early stages of their life is an important determinant of lifetime socioeconomic outcomes has been well established in the recent economic literature. Francesconi and Heckman (2016) report that at least 50% of the variability of lifetime earnings across persons is due to differences in attributes determined by age 18. Family and community background, which includes the neighborhood where children grow up and the school they attend, are considered important attributes determining later outcomes (e.g. Page and Solon, 2003; Raaum, Sørensen and Salvanes, 2006; Chetty and Hendren, 2015). Families can determine earnings by transmitting abilities, preference and resources, while communities can influence earnings through neighborhood quality, school quality and peers.

In this paper, we study the relative influence of family, schools and neighborhoods on earnings inequality over the life cycle. While there is a large and growing body of evidence on the influence of each of these attributes on adult outcomes, very little is known about their *relative* influence on earnings and how it evolves over the *life cycle*. Understanding the relative magnitude of these initial conditions on earnings throughout the life cycle is important for identifying the driving forces of existing inequalities and for interventions that aim to reduce them, especially because some early influences may be longer lasting than others.

There are two important challenges to consider in this analysis. The first challenge is related to data availability, since tracing the relative influence of these attributes on earnings over the life cycle requires almost complete earnings trajectories. We address this first challenge by using data from administrative registers of the Danish population, which due to their longitudinal dimension allow us to create individual earning histories based on tax records up to age 51 and avoid the well-known measurement error and life cycle biases. Long earnings histories enable us to assess the importance of families and communities in shaping the inequality of *permanent* earnings, which are generally considered to be the most relevant for individual welfare because *transitory* earnings shocks are more insurable (Blundell, Pistaferri and Preston, 2008).

The second challenge is that families with better resources or abilities tend to select better schools and neighborhoods. Because of sorting, separating the influences of family and community requires variation of the community environment that is exogenous to the outcome. We address this second challenge by observing siblings, their schoolmates and their close neighbors – sharing the same parish of residence – in the year they turn 15, which is the last year of compulsory education in Denmark.[1] In our research design, since communities are defined on the basis of individual year of birth, siblings can be exposed to different communities (school, neighborhood or both) because of family mobility in the time window between the years each sibling turns 15. Specifically, variation in community exposure is coming from families who move after the year the older sibling turns 15 and before the year the younger sibling turns 15, so they do not reside in the same community at age 15. This variation in community exposure due to family mobility within a specific time window allows separating community effects from family effects.[2]

However, it could be argued that this variation in community exposure may be subject to selection due to underlying unobserved differences between moving and staying families, where the latter are families for which siblings are in the same community at age 15. To support the validity of the identifying assumption we also exploit plausibly exogenous variation in community exposure coming from differences in the *timing of mobility* among only moving families. These are all families who move, so it is only the difference in the timing of mobility that exposes siblings to different environments. In this case, the comparison is between families who move after

---

[1] By defining community exposure at a specific age we consider closer neighbors and schoolmates since, all else equal, peers of the same age are more likely to interact than peers of different ages. In sensitivity checks we show robustness to defining neighborhood exposure at a specific age down to age 10.

[2] In the analysis we consider only males by sampling brother pairs, their male neighbors and male schoolmates. As we discuss in detail in Section 3, in our sample each male aged 15 has on average 19.5 male schoolmates born in the same year, and 14.6 male neighbors living in the same parish born in the same year.

the year the older sibling turns 15 – inducing within-family variation in community exposure – and families who move after the older sibling turns 10 but before turning 15 – so both siblings are in the same community at age 15. We can also separate neighborhood from school effects because in Denmark an important institutional feature is that public school catchment areas do not completely overlap with parishes. This implies that neighbors may be enrolled in different schools and schoolmates may come from different parishes.[3]

For estimation, we exploit the linked earnings records of siblings, neighbors and schoolmates and develop a model of individual earnings dynamics in the spirit of Baker and Solon (2003), which we extend to account for the joint earnings dynamics of multiple groups of individuals.[4] We consider time series of individual earnings from age 24 and up to age 51. Our model features life-cycle profiles, representing the long-term or permanent component of earnings, and transitory shocks. We allow the inequality of long-term earnings to depend on inequality in initial earnings, which reflects heterogeneous human capital investments, and inequality in earnings growth rates, capturing heterogeneous returns to investments. We allow these two sources of inequality in long-term earnings to depend on heterogeneity among families and heterogeneity among youth communities, consistent with the idea that, by affecting human capital investments and their returns, circumstances early in life may have long lasting effects. We also allow for purely idiosyncratic unit root shocks over the life cycle which permanently shift individual earnings trajectories. We use the estimated parameters of the model to decompose for the first time the sibling correlation of earnings into the three components of interest – family, neighborhood and school – allowing for life cycle dynamics.

---

[3] Among schools 89.5 percent have individuals from more than one parish, and amongst parishes 60.1 percent have individuals from more than one school. See Section 3 for details.

[4] Bingley and Cappellari (2013) extended Baker and Solon (2003) to account for dynamics within a three-person family, i.e. three pairwise relationships. With respect to Bingley and Cappellari (2013) we account for dynamics within three *groups* (family, school and neighborhood), i.e. three types of relationships, but each type of relationship could have arbitrarily many pairwise relationships, depending on how many individuals belong to each group.

We find that family is far more important than community in influencing earnings inequality over the life cycle. There is an almost equal positive influence of neighborhoods and schools on earnings only early in the working life, which falls rapidly becoming insignificant after age 30. This implies that measuring earnings at relatively young ages leads to overstate the long term relevance of community effects. In particular, the share of sibling correlation accounted for by community effects is 27 percent if incomes are measured up to age 25, it drops to 19 percent if incomes are measured up to age 30, but it is only 6 percent if we consider the entire available life cycle profile of earnings up to age 49. This implies on average over the life cycle a limited influence of community background because any community effects early in life 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 the working life. These results are robust to the definition of youth communities, the degree of exposure to communities and various sample selection choices.

Generally speaking, this paper contributes to the literature on the influence of family and community background on earnings. 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 as an omnibus measure of the influence of *both* family and community background (for reviews see Solon, 1999; Björklund and Jännti, 2009; Black and Devereux, 2011). To disentangle family from community effects, a common approach followed in this literature is to compare the correlation of sibling earnings with the correlation of earnings among unrelated neighbors (e.g. Page and Solon, 2003; Raaum, Sørensen and Salvanes, 2006). The idea is that while siblings share both the family and the neighborhood, unrelated neighbors share only the neighborhood but not the family. The findings from this approach suggest a substantial effect of neighborhoods on earnings. However, the estimated 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. Dealing with 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.

We contribute to this literature by developing a unified framework that allows us to measure the relative influence of family and community background on earnings over the life cycle, and also by distinguishing between the influence of neighborhoods and schools. The key difference between our research design and the one common in the sibling correlation literature is that, by defining communities on the basis of individual year of birth, we can separately identify sorting from neighborhood, school and family effects by exploiting arguably exogenous family mobility across communities and the partial overlap of schools and neighborhoods. We show that without exploiting the variation in community exposure that allows identifying sorting, the influence of neighborhoods and schools are found to be persistent throughout the life cycle and upward biased by a factor of three. This highlights the importance of sorting and also that, despite Denmark being generally considered a more equitable society, there are still important community differences that families tend to react to similarly to other countries.

Outside the sibling correlation literature, this paper relates to the large and growing literature on the impact of neighborhoods on children and adult outcomes. The evidence from social experiments such as the Moving to Opportunity program (MTO), in which families living in high poverty neighborhoods in five U.S. cities were randomly assigned vouchers to move to less impoverished neighborhoods, suggests that changes in neighborhood quality had on average little impact on economic outcomes (e.g. Ludwig et al., 2013). Recently, Chetty, Hendren and Katz (2016) complement MTO data with administrative information from tax returns and find that moving to a higher income neighborhood is associated with increased earnings in the mid-twenties for children who were below age 13 when their families moved. Using tax records for the general U.S. population, Chetty and Hendren (2015) exploit differences in the timing of mobility during childhood to show that moving to a better neighborhood has positive effects on earnings during early adulthood up to age 30, and the effect is greater with longer childhood exposure.

In our study, we exploit variation in community exposure for Denmark – due to differences in the timing of mobility – which is similar to the variation exploited by Chetty and Hendren (2015) for the U.S. We find similar positive youth neighborhood effects on earnings during early adulthood and up to age 30, but these effects are not persistent later in the working life. This evidence of no effects in the very long run is consistent with the evidence by Gould, Lavy and Paserman (2011) who use the airlift of Yemenite immigrants as a natural experiment and find that there are no very long run effects of early childhood environment on economic outcomes.[5]

Finally, this paper relates to the literature on the effect of school quality and peers on earnings. Early studies measuring school quality with the pupil/teacher ratio report a positive effect on the rate of return to schooling and earnings (Card and Krueger, 1992) whereas others find no impact at all (Dearden, Ferri and Meghir, 2002). More recently the focus has shifted to long run earnings. Black, Devereux and Salvanes (2013) correlate characteristics of 9th grade Norwegian schoolmates with outcomes at age 33-44 and find a strong relation between focal pupil earnings and peer fathers' earnings. Another branch of this literature exploits (quasi-) random variation of class size with most studies finding no effects (Leuven and Kokken, 2015; Falch, Strom and Sandsor, 2015 for Norway, Chetty et al., 2011 for the US), while Fredriksson, Öckert and Oosterbeek (2013) find a significant (at the 10 percent level) effect of smaller classes on long-term

[5] 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).

earnings. In summary, long run school effects on earnings are often found to be zero or borderline significant. In our study, we measure the long-term effects of any school attribute shared between schoolmates, finding – similar to neighborhood effects – a positive effect on earnings only in the beginning of the working life and not in the very long run.

The paper is structured as follows. In Section 2 we develop the econometric model based on the joint analysis of life-cycle earnings for brothers, neighbors and schoolmates. In Section 3 we describe the data and contrast our neighborhood definition with that used in comparable studies. In Section 4 we present descriptive statistics on earnings of siblings and peers over the life cycle. In Section 5 we present the main results comparing our findings to the previous literature, robustness to the identification assumption and detailed sensitivity analysis. We conclude in the last section.

## 2. Econometric Model

The aim of this paper is to identify the determinants of long-term earnings and, in particular, the extent to which earnings inequality can be explained by differences in family and social background. To estimate the contributions of family and community background on permanent earnings we exploit the linked earnings records 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.

We separate life cycle effects from calendar time trends by considering the distribution of earnings within two-year birth cohorts. In particular, the residual log earnings – after regressing log real gross annual earnings on year dummies and a quadratic age trend by birth cohort group – denoted by $w$, are assumed to be the sum of two components, which are orthogonal by definition: (i) a permanent component denoted by $y$ and (ii) a transitory component denoted by $v$,

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

where the indices *i, f, s, n* and *a* stand for individual, family, school, neighborhood and age, respectively.[6] Schools and neighborhoods are defined as the school attended at age 15 and the parish of residence at age 15. In Section 3 we provide details on sample construction and community definition.

The model extends the joint earnings dynamics model of Bingley and Cappellari (2013) for three persons (a father and two sons) to several multi-person groups. The model also tackles the two measurement error biases in the estimation of correlations in permanent earnings between persons which are 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 make 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; Bohlmark and Lindquist, 2006; Nybom and Stuhler, 2016).

**2.1 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 siblings, 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

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

also known as a *random growth* model. We augment this with a *restricted income profile* (RIP) process for individual-specific components, which is an idiosyncratic unit root (*random walk*) shock.

The heterogeneous income profile 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). Based on the analysis of Becker and Tomes (1979), parents can influence the human capital of their offspring by transmitting abilities, preferences and resources, and thereby affect offspring 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. Human capital models predict that heterogeneous investments in human capital induce an inverse relationship 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 initial earnings 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 motivate our specification choice for shared earnings determinants, which reflects the idea that the across persons resemblance of earnings stems from similarities in social background and human capital investments. We show in Section 4 that 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 restricted income profile, which is modeled as a unit root in age

($\omega_{ia}$) and 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:

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

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

where $c(i)$ is the birth cohort of person $i$, $\delta_c$ is a birth cohort effect and $\pi_t$ is a calendar time shifter allowing for the possibility of aggregate changes of the permanent earnings process over time. The intercept and the slope 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 $\mu_f, \mu_s, \mu_n$) and life-cycle *earnings growth* (denoted by $\gamma_f, \gamma_s, \gamma_n$).

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

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

$$\left( \mu_f, \gamma_f \right) \sim \left( 0,0; \sigma^2_{\mu\Phi}, \sigma^2_{\gamma\Phi}, \sigma_{\mu\gamma\Phi} \right); \tag{3.b}$$

$$(\mu_s, \gamma_s) \sim \left( 0,0; \sigma^2_{\mu\Sigma}, \sigma^2_{\gamma\Sigma}, \sigma_{\mu\gamma\Sigma} \right); \tag{3.c}$$

$$(\mu_n, \gamma_n) \sim \left( 0,0; \sigma^2_{\mu\mathrm{N}}, \sigma^2_{\gamma\mathrm{N}}, \sigma_{\mu\gamma\mathrm{N}} \right), \tag{3.d}$$

where the specific dimensions of heterogeneity of the variance-covariance parameters are denoted by $\Phi$ (for family), $\Sigma$ (for school) and $\mathrm{N}$(for neighborhood).

Assumption (3a) allows the idiosyncratic parameters to vary by siblings' birth order, which we denote by *b*. We include singletons among the first born. Assumptions (3.b-3.d) specify the distribution of the shared components and allow for unrestricted covariance (denoted by $\sigma_{\mu\gamma\Phi}, \sigma_{\mu\gamma\Sigma}, \sigma_{\mu\gamma\mathrm{N}}$) of initial heterogeneity and growth rate heterogeneity within each component.

We also model the sorting of families across communities by allowing for the covariance between family and community effects, as well as for the covariance of school and neighborhood effects between families that share the community:

$$cov(\mu_f, \mu_s) = \sigma_{\Phi\Sigma};\ cov(\mu_f, \mu_n) = \sigma_{\Phi N};\ cov(\mu_s, \mu_n) = \sigma_{\Sigma N}. \tag{3.e}$$

Correlation across family and community effects is allowed through the intercepts of the individual-specific profiles. This choice is made because empirically most of the community effects vanish after two or three years (see Figure 4), and in order not to overcrowd the parameter space.[7] The first two covariances ($\sigma_{\Phi\Sigma}$ and $\sigma_{\Phi N}$) are non-zero if families sort themselves across communities. The third parameter ($\sigma_{\Sigma N}$) is a contextual term capturing the similarity of family characteristics in the community, which induce a positive covariance between school and neighborhood effects.

## 2.2 Specification of transitory earnings

We model transitory earnings ($v$) in equation (1) to capture any serial correlation of transitory shocks using an AR(1) process. We allow siblings 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:

$$v_{ifsna} = \eta_t u_{ifsna};\ u_{ifsna} = \rho_b u_{ifsn(a-1)} + \varepsilon_{ifsna}; \tag{4}$$

$$\varepsilon_{ifsna} \sim \left(0, \sigma^2_{\varepsilon b} \exp\left(g_b(a)\right)\right), \qquad u_{ifsn24} \sim \left(0, \psi_c \sigma^2_{u_{24}b}\right),$$

where $\eta_t$ is a time loading factor and $u_{ifsna}$ is the birth-order-specific AR(1) process with birth order denoted by the index $b$. The autoregressive process begins at age 24 and we specify the

[7] Allowing for sorting effects also in the slopes of the HIP model resulted in some negatively estimated variances for the slopes, indicating lack of identification for that parameterization in our data. We could estimate sorting effects of both intercepts and slopes without identification issues in simpler models that consider one dimension of community, either school or neighborhood. Including sorting effects in the slopes of the HIP model did not alter our substantive conclusions.

variance of the initial condition, denoted by $\sigma^2_{u_{24}b}$, to be birth-cohort-specific with parameter $\psi_c$. The process evolves through the arrival of white noise shocks (denoted by $\varepsilon$) whose variance is age-and-birth-order-specific ($\sigma^2_{\varepsilon b}\exp\big(g_b(a)\big)$, 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 siblings, the use of birth-order-specific distributions of shocks enables identifying the contemporaneous correlation of AR(1) innovations. Let $i$ and $i'$ index two individuals; the sibling covariance of AR(1) innovations is specified as follows:

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

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 covary with parameter $\sigma_f$. This covariance of shocks between siblings does not depend on whether they attended the same school, or lived in the same parish and is transmitted to later time periods through the autoregressive structure of the model.

Due to the high dimensionality that would result from parameterizing the covariance of shocks between numerous community members belonging to different families ($f$ and $f'$), we follow a different approach to that used for pairs of siblings. We allow for catch-all “mass-point” covariances ($\lambda$) 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 $a'$, covariances of transitory shock across non-sibling peers are specified as follows:

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

$$E\big(u_{ifsna} u_{i'f's'na'}\big) = \lambda_{n}^{1+|t-t'|}, |\lambda_j| < 1, j = sn, s, n.$$

**2.3 Identification of permanent earnings components**

In this sub-section, we discuss the identification of the *permanent* earnings components, which besides time and cohort effects, includes three sets of parameters: i) family effects, ii) community effects and iii) sorting effects. Assumptions (3.a) – (3.e) fully specify the intertemporal and interpersonal distribution of permanent earnings. Identification of parameters is achieved by exploiting three different types of moment restrictions (earnings covariances) generated by the model: i) for siblings who share the community, ii) for siblings who do not share the community and iii) for non-sibling peers.

For a given individual, earnings covariances between 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:

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

$$\{\sigma^2_{\mu\Phi} + \sigma^2_{\mu\Sigma} + \sigma^2_{\mu\mathrm{N}} + \big(\sigma^2_{\gamma\Phi} + \sigma^2_{\gamma\Sigma} + \sigma^2_{\gamma\mathrm{N}}\big)aa' +$$

$$\big(\sigma_{\mu\gamma\Phi} + \sigma_{\mu\gamma\Sigma} + \sigma_{\mu\gamma\mathrm{N}}\big)(a + a') + 2\sigma_{\Phi\Sigma} + 2\sigma_{\Phi\mathrm{N}} + 2\sigma_{\Sigma\mathrm{N}} +$$

$$\sigma^2_{\omega_{24}b} + \sigma^2_{\xi b} min(a, a')\}\delta^2_c \pi_t \pi_{t'}.$$

Cross-person moments (between siblings, neighbors, or schoolmates) do not depend on idiosyncratic heterogeneity. Moment restrictions between siblings (different $i$ but same $f$) depend on the family effects and sorting effects. Moreover, they are also functions of school effects, neighborhood effects, both, or neither, depending on the extent to which siblings share schools and/or neighborhoods. Moment restrictions for siblings can be written as follows:

$$E\left(y_{ifsna}, y_{i'fs'n'a'}\right) = \tag{8}$$

$$\{\sigma^2_{\mu\Phi} + \sigma^2_{\gamma\Phi}aa' + \sigma_{\mu\gamma\Phi}(a + a') +$$

$$I(s = s')\left[\sigma^2_{\mu\Sigma} + \sigma^2_{\gamma\Sigma}aa' + \sigma_{\mu\gamma\Sigma}(a + a')\right] +$$

$$I(n = n')\left[\sigma^2_{\mu\mathrm{N}} + \sigma^2_{\gamma\mathrm{N}}aa' + \sigma_{\mu\gamma\mathrm{N}}(a + a')\right] +$$

$$2\sigma_{\Phi\Sigma} + 2\sigma_{\Phi\mathrm{N}} + 2\sigma_{\Sigma\mathrm{N}}\}\delta_c\delta_{c'}\pi_t\pi_{t'} \,,$$

where $I(.)$ is an indicator function.

Equation (8) nests moments restrictions for four types of siblings who: (1) share both the school and the neighborhood, i.e. $I(s = s') = 1$ and $I(n = n') = 1$; (2) share only the school, i.e. $I(s = s') = 1$ and $I(n = n') = 0$; (3) share only the neighborhood, i.e. $I(s = s') = 0$ and $I(n = n') = 1$; and (4) share only the family but neither the school nor the neighborhood, i.e. $I(s = s') = 0$ and $I(n = n') = 0$.

The above moment conditions permit the identification of community effects separately from family and sorting because not all sibling pairs share the community (school and neighborhood). Since communities are defined on the basis of individual year of birth, siblings can be exposed to different communities (school, neighborhood or both) because of family mobility in the time window between the years each sibling turned 15, where age 15 is the year we use to define communities. This variation in community exposure due to family mobility within a specific time window – after the year the older sibling turned 15 and before the year the younger sibling turned 15 – allows separating community effects from family effects. Within-family variation in communities represents the key source of variation for our identification strategy.

To see this in equation (8), for siblings that share both components of the community, i.e. $I(s = s') = 1$ *and* $I(n = n') = 1$, the covariance function depends on family, sorting, and community effects. For siblings that do not share the community, i.e. $I(s = s') = 0$ *and* $I(n = n') = 0$, the covariance function depends only on family and sorting effects. This is the case in which siblings attended different schools at age 15 and lived in different neighborhoods because the family moved in the time window between the years each sibling turned 15. The difference in covariance functions between these two types of siblings identifies the *sum* of school and neighborhood effects.

We can further disentangle school from neighborhood effects using siblings that share only one component of the community effect, i.e. $I(s = s') = 1$ *or* $I(n = n') = 1$. For them, the covariance function depends on family and sorting, plus the common community component. This is the case in which siblings attended different schools at age 15 but lived in the same neighborhood, or lived in different neighborhoods at age 15 because of family mobility between the years each sibling turned 15, but attended the same school. The difference in covariances between families sharing both community dimensions and those sharing only one dimension identifies heterogeneity in the non-shared dimension.

The within-family variation in community exposure that we exploit for identifying community effects rests on the assumption that families changing community in this specific time window are not selected and can be compared with families in which siblings reside in the same community in the year they turn 15. However, these two types of families may be different due to underlying unobserved characteristics that also affect earnings. To support the validity of the identifying assumption, we also exploit differences in the *timing of family mobility* and perform the analysis focusing only on families who moved.

Siblings who share the community in the year they turn 15 belong to two different types of families: i) families who never moved, and ii) families who moved before the older sibling turned 15. Therefore, instead of comparing moving to staying families, we can compare families who moved after the older sibling turned 15 with families who moved before the older sibling turned 15. These are all families who move, so it is only the difference in the timing of mobility that exposes siblings to different environments. This is plausibly exogenous variation which can be argued to be less prone to selection than contrasting movers and stayers. We show in Section 5.3 that our results are robust when variation is restricted only to movers exploiting differences in the timing of family mobility.

Although moment conditions for siblings can identify community effects, they are not sufficient to separate family effects from sorting. This is evident from the fact that the term $2\sigma_{\Phi\Sigma} + 2\sigma_{\Phi\mathrm{N}} + 2\sigma_{\Sigma\mathrm{N}}$ 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 siblings, and such covariance is not separable from the variance of family effects in equation (8). To separately identify the sorting parameters $\sigma_{\Phi\Sigma}$, $\sigma_{\Phi\mathrm{N}}$ and $\sigma_{\Sigma\mathrm{N}}$ from the family effects, we exploit moment restrictions for non-sibling peers (neighbors and schoolmates) that by definition *do not* share the family ($f \neq f'$):

$$E\left(y_{ifsna}, y_{i'f's'n'a'}\right) =$$

$$\{I(s = s')\left[\sigma^2_{\mu\Sigma} + \sigma^2_{\gamma\Sigma}aa' + \sigma_{\mu\gamma\Sigma}(a + a') + 2\sigma_{\Phi\Sigma}\right] + \tag{9}$$

$$I(n = n')\left[\sigma^2_{\mu\mathrm{N}} + \sigma^2_{\gamma\mathrm{N}}aa' + \sigma_{\mu\gamma\mathrm{N}}(a + a') + 2\sigma_{\Phi\mathrm{N}}\right] + 2\sigma_{\Sigma\mathrm{N}}\}\delta^2_c\pi_t\pi_{t'}.$$

For non-sibling peers, the covariance function depends on community and sorting effects. Because community effects are already identified by the moment restrictions for siblings, moment restrictions in (9) effectively identify sorting of families across communities. Equation (9) nests three moment restrictions (for peers sharing both school and neighborhood and for peers sharing either school or neighborhood) which provide identification for the three parameters $\sigma_{\Phi\Sigma}$, $\sigma_{\Phi\mathrm{N}}$ and $\sigma_{\Sigma\mathrm{N}}$. Identification is achieved thanks to the partial overlapping of peer groups, which depends on the design of school catchment areas across parishes (see Section 3).

Finally, combining the moment restriction of equation (8) – which provides both the total sibling effect and the identification of community effects – with those of equation (9) – which provides identification of sorting effects – we identify family effects.

The key difference between our research design and that of the sibling correlation literature is that by defining communities on the basis of individual year of birth we can separately identify sorting from neighborhood, school and family effects by exploiting arguably exogenous family mobility across communities based on differences in the timing of mobility and the partial overlap of schools and neighborhoods. Earlier research could not separate community from sorting effects because – due to the sampling design – siblings always shared the community. We are the first to estimate both the sorting and contextual covariances within sibling covariances.[8]

### 2.4 Estimation and decomposition of the sibling correlation

The model is estimated by Minimum Distance matching moment restrictions implied by the model to the empirical moments derived from the data.[9] As mentioned earlier, the empirical moments are based on the residuals, after regressing log real gross annual earnings on year 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 intertemporal

[8] Raaum, Sørensen and Salvanes (2006) use a linear projection of earnings on neighborhood characteristics and neighborhood fixed effects to derive an approximation for the contextual term.

[9] 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 (see, for example, Haider, 2001).

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 once in the estimation of sibling empirical moments, but families with only one son do not contribute. We estimate separate empirical moments for siblings depending on whether they shared the school, the neighborhood, both or neither, 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. In contrast to families, the number of peers varies within community clusters. 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-cluster 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.

Using parameter estimates from the model we can predict the contributions of family and community to the sibling correlation of permanent earnings over the life cycle. For this purpose, we can use moment conditions (8) and (9) and attribute sorting parameters, whose allocation to either factor is inherently ambiguous, in equal parts to family and to communities:

$$r^F(a) = \frac{E\left(y_{ifsna}, y_{i'fs'n'a}\right) - \delta_c \delta_{c'} \pi_t \pi_{t'} (\sigma_{\Phi\Sigma} + \sigma_{\Phi \mathrm{N}})}{E\left(y_{ifsna}, y_{ifsna}\right)} \; \forall\, s \neq s' and\; n \neq n',$$

$$r^S(a) = \frac{E\left(y_{ifsna}, y_{i'f'sn'a}\right) - \delta_c^2 \pi_t \pi_{t'} (\sigma_{\Phi\Sigma} + \sigma_{\Sigma \mathrm{N}})}{E\left(y_{ifsna}, y_{ifsna}\right)} \; \forall\, f \neq f' and\; n \neq n', \tag{10}$$

$$r^N(a) = \frac{E\left(y_{ifsna}, y_{i'f's'na}\right) - \delta_c^2 \pi_t \pi_{t'} (\sigma_{\Phi \mathrm{N}} + \sigma_{\Sigma \mathrm{N}})}{E\left(y_{ifsna}, y_{ifsna}\right)} \; \forall\, f \neq f' and\; s \neq s',$$

where $r$ denotes correlation coefficients of permanent earnings, while $F$, $S$ 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 ($r^B(a)$) for brothers sharing the community is the sum of the three components:

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

## 3. Data

### 3.1 Sample Selection

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 our analysis we consider only males and we construct our dataset as follows. First we create a 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 create our dataset of brothers we consider men born in the years 1960-1983 and select first and second sons who share both legal parents from registration at birth and are not adopted. The year of birth selection starts in 1960 because of prior incompleteness of registered parentage and stops in 1983 to allow observation of individual earnings up to the late 20s for younger cohorts (earnings data described later in the Section are available until 2011).[10] We exclude from the sample sons who are also observed as fathers, brothers born less than 12 months apart and brothers who are born more than 12 years apart. In this way we obtain 122,652 sibling pairs. Following the

[10] 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.

tradition in the sibling literature we keep singletons (men without a younger brother) in the sample; there are 380,948 singletons in the sample. The robustness analysis in Section 5 shows that excluding singletons does not affect our findings.

Next we link our sampled brothers (and singletons) to all other males who are born in the same year and share their youth communities (schools or neighborhoods). In contrast to our treatment of siblings, peers are included in the analysis irrespective of birth order and age spacing from their own brothers, which adds 69,708 individuals to the analysis. Using information from the educational register, 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.[11] Our sample contains 1,858 schools with males aged 15. They have on average 19.5 male schoolmates born in the same calendar year. Using information on individual addresses obtained from the central person register, we define neighborhoods as the parish of residence. Individuals are required to report changes of address to the municipality within five days. As with schools, our census point is 31 October of the calendar year a male turns 15. There are 2,123 parishes in our sample containing on average 14.6 males turning 15 in the same year.[12]

School enrolment rules were such that pupils should start in first grade in the August of the calendar year they turn 7. The national pupil database was established to monitor compliance with the 1972 school reform, which made 8th and 9th grade schooling compulsory in 1972/3 and 1973/4 respectively. Beginning in August 1973, the database links pupils to the schools they are

[11] In 1980, 95 percent of pupils began 9th grade during the year they turned 15. In recent years delays have been more common – in 2007, 13 percent of pupils delayed their school start by a year and 4 percent repeated the same grade the following year.
[12] Complete information on municipality of residence is available from 1971 and full addresses are complete from 1977 (see Pedersen, et al. 2006 for details). We use an intermediate aggregation of locality as our neighborhood indicator – parish of residence – which is available from 1973.

enrolled from 8th grade and above.[13] 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).[14]

During our sample period, pupils were assigned to public schools on a catchment area basis according to place of residence. Primary and lower secondary education usually takes place in the same school and most pupils attend the same school for all grades. From 2007 we can see that 92% 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. This institutional setting makes Denmark a good place to look for school effects, because of the coherence of the schoolmate group.

Also, an important Danish institutional feature is that parishes and public school catchment areas do not completely overlap. As a consequence, neighbors may attend different schools, and schoolmates may come from different neighborhoods. Amongst school-birth-cohort clusters, 89.5 percent have individuals from more than one parish, and amongst parish-birth-cohort clusters 60.1 percent have individuals from more than one school.

Because communities are defined on the basis of individual year of birth, not all brothers will share the community. Among our 122,652 brother pairs, 70 percent share both schools and neighborhoods at age 15, 5 percent share only the school, 16 percent share only the neighborhood, while the remaining 9 percent do not share either component of the youth community.

[13] In anticipation of grade K enrolment being made compulsory from August 2009, the national pupil database was extended to cover grades K-7 from August 2007. Hence, we are unable to match pupils to schoolmates in earlier grades to look at long run outcomes.

[14] We exclude from our estimation sample individuals enrolled at schools which are exclusively for pupils with special educational needs. Private schools are smaller on average than public schools, are primarily in urban areas and are heavily subsidized, with municipalities covering 85 percent of expenditure. In sensitivity analysis we show that results are robust to private school exclusion.

For both brothers and for peers we use pre-tax annual labor earnings between 1984 and 2011, measured in 2005 prices. We select all valid observations on earnings and exclude zero-earnings observations. The exclusion of zero-earnings observations is common with most of the earnings dynamics literature assuming that earnings are missing at random, and is also applied in the sibling correlation literature by Björklund et al (2009). We 'trim' a quarter of a percentile from each tail of the annual earnings distribution and require at least three consecutive earnings observations for an individual to be included in the sample, a selection rule that is intermediate between the one used by Baker and Solon (2003), i.e. continuous earnings strings for each individual within a cohort, and the approach of Haider (2001), who allows individuals to move in and out of the sample only requiring two positive but not necessarily consecutive valid observations on earnings.

Table 1 presents the cohorts we include in the sample, the years for which we observe earnings and sample sizes in various dimensions. Following Baker and Solon (2003), we group data in two-year birth cohorts, as shown in column 1, and we compute age by imputing each cohort with its first year of birth. The selection of birth cohorts and time window ensures that each cohort is observed starting at age 24 for at least 6 years (last cohort 1982) and for as long as 28 years (first cohort 1960). Columns 5 and 6 present the number of earnings observations and number of men used in estimation. The number of community cohorts into which men are grouped is shown in columns 7 and 8, totaling 33,907 school cohorts and 47,622 parish cohorts.

Half of earnings observations are from the first third of birth cohorts. Cohort size peaks in 1966-7 and falls by 44 percent by the last cohorts in 1982-3. Number of school cohorts and parish cohorts is quite stable, reflecting an absence of administrative unit reform. Population demographics dictate birth cohort sizes and the falling number of earnings observations by birth cohort is due to later cohorts having less time to accumulate earnings histories.

### 3.2 Sample comparison to other studies

It is informative to contrast our neighborhood definition with that used in comparable studies such as Page and Solon (2003), Raaum, Sørensen and Salvanes (2006) and Oreopoulos (2003).[15] Table 2 characterizes ours alongside these three other studies according to characteristics of the different types of neighborhoods and exposures considered, and outcomes observed. Neighborhood geography and exposure group – an area and an age range – together define the cluster of individuals within which later outcomes are correlated. Neighbors in study (3) have the closest proximity because of the medium-to-high density of housing projects, followed by study (1) because of the clustered PSID sampling frame. Interestingly, study (1) finds neighborhood effects only for urban areas, where neighbors are in closest proximity. Our Danish parishes cover a wider area than the neighborhoods used in studies (1) and (3), but are only about half the size of Norwegian census tracts used in study (2). For Denmark in the year 2000 we can calculate the distribution of distances between the different residences of 15 year olds within parish: 25 percent of distances are within 0.5 km, 50 percent within 1.1km, and 75 percent within 1.9km.

The other three studies pool neighbors together of different ages – with up to 9 and 11 years age differences – to form neighborhood clusters. In the main part of the analysis we consider neighbors at 15 years of age as belonging to the same cluster. Neighborhood exposure at age 15 is at the upper end of the 5-16 age range together considered in the other studies. All else equal, if neighbors of the same age are more likely to interact than neighbors of different ages, then we would expect to find stronger neighbor correlations with our definition. In sensitivity checks we show robustness of results to neighborhood definitions based on exposure down to age 10.

---

15 In what follows we refer to Page and Solon (2003) as study (1), Raaum, Sørensen and Salvanes (2006) as study (2) and Oreopoulos (2003) as study (3).

The number of men in each of our neighborhood clusters is in the middle of the range for the other studies. The estimation sample in studies (1) - (3) comprises a similar percentage of the total population of individuals in each cluster with 4.6, 3.1, and 4.8 percent respectively. Due to our narrower age range for clustering neighbors for Denmark the estimation sample covers only 0.6 percent of the cluster population. Although our neighbors are more homogeneous in terms of age, they represent only between one eighth and one fifth of the within-cluster sampling density of the other studies. However, this sparser sampling would reduce precision rather than introduce any bias.

The duration of neighborhood exposure observed for an individual differs between studies, with only study (3) using more than a single point in time. Although studies (1) and (2) consider a range of ages of exposure, attachment to a neighborhood is only considered at a single point in time, just as in our study. The persistence of neighborhood affiliation is likely to differ between contexts, and this will affect saliency of the neighbor groupings used, but it is difficult to establish how important this might be for explaining differences in findings between studies. In sensitivity analyses, we use alternative definitions for neighbors based on those living in the same parish at each age 14 to 15, at each age 10 to 12, or based on the more frequent parish of residence between ages 14 to 18.

**4. Descriptive statistics on earnings of siblings and non-sibling 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 siblings (brothers) and 2) between male non-sibling peers attending the same school and/or residing in the same neighborhood (parish) at age 15.

For brothers, we compute the covariance of their earnings from families with at least two male children. For male non-sibling peers, we group them in clusters depending on whether they share 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. 841), which gives more importance to more populated clusters, and makes inference person-representative.

We begin by describing the correlation of sibling earnings by age in Figure 1. The solid line labeled "At same age" reports the computed correlation when the brothers are at the same point in their life cycle, a comparison 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 shared between brothers are negatively correlated with heterogeneity in earnings growth. As discussed in Section 2.1, human capital models predict investments in education or training to induce such a negative correlation. The dotted line fixes the age of the older among the two brothers at age 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 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 discussed in 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 depicted in Figure 1 may also reflect the correlation of transitory shocks. It is well known that earnings instability is large in the beginning of the working life (see e.g. Baker and Solon, 2003). It is also plausible that siblings may be subject to common shocks; for example due to similar local economic conditions at labor market entry. As a way to assess if the relatively large sibling correlation at labor market entry is driven by permanent earnings differences or transitory fluctuations, in Figure 2 we present the earnings correlation for brothers born at least five, eight or ten years apart. The larger the age difference, the less likely it is for brothers to enter the labor market at the same time and share transitory shocks. Therefore, these samples are less likely to be influenced by transitory fluctuations compared with the samples underlying Figure 1. As Figure 2 shows, the declining pattern of the sibling correlation between the mid-20s and the early-30s persists even after excluding closely-spaced-brothers who are more likely to share transitory earnings fluctuations. This suggests that the source of the convex evolution of sibling correlations shown in Figures 1 and 2 is in the permanent earnings component.

In Figure 3, we plot the earnings correlations for male non-sibling 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 similarity, which include both those correlated with family effects and those independent of them. There are a few points worth highlighting in this figure. The first is related to the magnitude of the earnings correlation of non-sibling peers, 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, Figure 3 also reports the correlation of earnings for "Unrelated" individuals, who do not belong to the same family and who are neither schoolmates nor neighbors in the parish of residence. We compute this correlation by randomly matching each individual in the sample with 1000 unrelated individuals of the same age. We find this correlation to be equal to zero for all ages, which suggests that the evolution of sibling and non-sibling peer correlations over the life cycle is picking up some underlying forces due to families, schools and neighborhoods, and is not simply an artifact of age effects.

## 5. Results

We concentrate the discussion on estimates for the 'core' parameters of the permanent and transitory components, which we present in Section 5.1.[16] In Section 5.2, we discuss the decomposition of the sibling correlation and compare the main findings with the existing evidence in the literature. In Section 5.3, we provide evidence for the validity of our identification strategy comparing the results of the full sample to those exploiting differences in the timing of family mobility among moving families. Finally, in Section 5.4 we present a set of sensitivity checks related to the sample and the measurement of community membership.

### 5.1 Parameter estimates

#### 5.1.1 Permanent earnings

Based on equation (2), permanent earnings depend on *shared* and *idiosyncratic* components. The parameter estimates for the shared components reported in Panel A of Table 3 show that family is by far the most relevant factor for long-term earnings. This is true both for initial earnings

[16] Parameter estimates of the time and cohort effects on both components are reported in Tables A1 and A2 of the Appendix. Their identification requires setting one parameter for each component and each dimension (time and cohort) equal to one. Exceptions are the cohort shifters on the initial condition of the transitory component ($\psi_c$). Because the model includes not only the variance of shocks at age 24 ($\sigma^2_{u_{24}b}$, the initial condition) but also a specific parameter for the variance of innovations at age 25 ($\sigma^2_{\varepsilon b}$), representing the baseline of the age spline, there is an additional cohort shifter that needs to be constrained. Empirically, we find estimates of shifters on the two oldest cohorts to be imprecise and negative, and we constraint also those two shifters.

(intercept) and for earnings growth rates (slope). The other relevant source of permanent inequality in earnings is the individual idiosyncratic component reported in Panel B of Table 3.

Shared components of long-term earnings in Table 3 display the Mincerian cross-over property. This is indicated by the negative covariance between intercepts and slopes of earnings profiles ($\sigma_{\mu\gamma\Phi}, \sigma_{\mu\gamma\Sigma}, \sigma_{\mu\gamma\mathrm{N}}$), which suggests that families or communities associated with low earnings at age 24 are also associated with faster growth in life-cycle earnings.[17] A corollary of the negative covariance is that the variance of permanent earnings across families or communities is *U-shaped in age* because it falls in the years of catch-up and increases after the point of cross-over. The point of cross-over can be computed as the year in which the earnings variance is minimized, and it is located at age 31 for the between-family earnings distribution, at age 34 for the between-neighbors earnings distribution, and at age 39 for the between-schoolmates earnings distribution.

The covariances between the three components of shared earnings determinants ($\sigma_{\Phi\Sigma}$, $\sigma_{\Phi\mathrm{N}}$ and $\sigma_{\Sigma\mathrm{N}}$) capture sorting of families into schools and neighborhoods. The estimates in Panel A suggest that sorting is relevant, as the covariances of family effects with school and neighborhood effects ($\sigma_{\Phi\Sigma}$, $\sigma_{\Phi\mathrm{N}}$) are 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 ($\sigma_{\Sigma\mathrm{N}}$), which suggests that school and neighborhood effects are positively correlated because similar families choose similar schools and neighborhoods.

### 5.1.2 Transitory earnings

Parameter estimates of transitory earnings in Table 4 show a clear age pattern of transitory shocks whose variance decreases between the mid-20s and the mid-30s, while the decrease slows down

[17] Taking into account that sorting contributes to the variance of intercepts, the correlation between intercept and slope is equal to -0.725.

around age 35. This sharp decline followed by a leveling-off is consistent with the patterns reported by Baker and Solon (2003) who find that the variance of transitory shocks declines at decreasing rate between the ages of 25 and 45. The patterns that we find by age are very similar between brothers. Autoregressive coefficients are also very similar between brothers, and are of moderate size. Finally, Table 4 shows that the correlation of transitory shocks between siblings or non-sibling peers is generally not statistically significant, with the exception of the correlation among schoolmates.[18]

### 5.2 Decomposition of sibling correlation

We can assess the relative importance of family, community and sorting by considering the decomposition of the sibling correlation at age 24. This is because the model allows for sorting of families across communities only within the intercept of the model, and the initial point of the income profile is set at age 24. Using the average of brothers' random walk intercepts as a measure of initial idiosyncratic variance, the overall sibling correlation is equal to 0.75.[19] The part of the sibling correlation that is accounted for by family effects is equal to 0.48, while the part accounted for by community effects and sorting is equal to 0.12 and 0.15, respectively. This suggests that communities alone exert only a small effect on earnings inequality measured at age 24.

We now move to the decomposition of the sibling correlation throughout the life cycle. We use the estimates reported in Table 3 to generate predictions of the sibling correlation and its decomposition into the three factors of interest (family, school and neighborhood) based on the formulae provided in Section 2.4 (equations 10 and 11), imputing estimated sorting parameters to family and community in equal parts. In particular, we consider the case represented by equation

18 Background analysis shows that these parameters are still small but statistically significant if we exclude cohort effects from the model.

19 This is computed using the following formula: $r_{24}^{B} = \frac{\sigma_{\mu\Phi}^2+\sigma_{\mu\Sigma}^2+\sigma_{\mu N}^2+2\sigma_{\Phi\Sigma}+2\sigma_{\Phi N}+2\sigma_{\Sigma N}}{0.5(\sigma_{\omega 24,1}^2+\sigma_{\omega 24,2}^2)+\sigma_{\mu\Phi}^2+\sigma_{\mu\Sigma}^2+\sigma_{\mu N}^2+2\sigma_{\Phi\Sigma}+2\sigma_{\Phi N}+2\sigma_{\Sigma N}}$.

(11) of two brothers who attend the same school and live in the same neighborhood at age 15. 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 sibling correlation is equal to 0.57 (s.e. 0.11) at age 25, drops to 0.19 (s.e. 0.121) at age 35, and rises back to 0.43 (s.e. 0.014) by age 49, which is the last age for which we observe younger brothers. The average sibling correlation over the life cycle, reported in Column (1) of Table 5, is equal to 0.32 (s.e. 0.010), which is in line with previous estimates for Denmark.[20] As mentioned earlier, the U-shaped pattern is a symptom of the "Mincerian cross-overs" of earnings profiles. That is, the negative estimates of the covariance between the intercept and the slope of the earnings profiles implies that the distribution of shared components, and therefore the siblings and non-sibling peers correlation, first shrinks and then fans out over the life cycle. The same U-shaped pattern is also a feature of the raw cross-person correlations shown in Figures 1 to 3, and especially in Figure 2, which depicts the earnings correlation for brothers born only a few years apart.

Considering the decomposition of the sibling correlation, it is evident from Figure 4 that family is far more important than community in influencing the dispersion of permanent earnings over the life cycle. The community effects are limited and are only significant at the beginning of the working life, while by age 30 they become negligible and not significantly different from zero. In particular, the estimated community correlation of earnings (the sum of neighborhood and school effects) is equal to 0.21 (s.e. 0.021) at age 24, drops to 0.053 (s.e. 0.010) at age 27 and becomes zero at age 30. As reported in Column (1) of Table 5, on average over the life cycle the estimated correlation in permanent earnings is equal to 0.011 (s.e. 0.010) across schoolmates, and

[20] Using a model without community effects, Björklund et al. (2002, p. 765) report for men aged 25-42 a sibling correlation of 0.29. We obtain an average estimate of 0.28 if we limit our sample in the same age range.

0.010 (s.e. 0.010) across neighbors. In both cases, the estimates are close to zero and not statistically significant. The overall correlation for community effects is small (0.021) and marginally significant (s.e. 0.009). These results indicate that there is not much room for community effects in shaping the sibling correlation in the long run. The family is the only factor that generates a substantial correlation in permanent earnings between brothers throughout the life cycle.

One implication of the decomposition depicted in Figure 4 is that measuring earnings at relatively young ages exaggerates the long run relevance of community effects. For example, the share of sibling correlation accounted for by community effects is 27, 24 and 19 percent, if earnings are measured up to ages 25, 27 and 30, respectively. However, the corresponding figure is only 6 percent if we consider the entire available life-cycle profile up to age 49. This shows the importance of analyzing earnings beyond the early part of the working life.

Our findings compare with those of Oreopoulos (2003) who finds a zero correlation of earnings between neighbors after accounting for sorting of families into neighborhoods by using quasi-random assignment of neighbors. Without taking sorting into account, Page and Solon (2003) find a positive correlation of neighbors earnings in the PSID equal to 0.16, which is half of their estimated sibling correlation of 0.32. Raaum, Sørensen and Salvanes (2006) report a youth neighbors' earnings correlation of 0.06 and a sibling correlation of 0.2, implying an incidence of 30 percent of neighborhood effect over the total sibling correlation.

The key feature of our sample design is that we exploit variation of communities within families, which combined with the partial overlap of schools and neighborhoods, enables disentangling family, community and sorting effects. In addition, within our model we can compare our baseline results with the case in which we ignore sorting in the following two ways. First, we can restrict the sample of siblings only to those who share *both* the school and the neighborhood. In this case only two out of three sets of parameters (family, community, sorting) are identified, so we restrict the sorting effects to zero. Second, we can exclude sibling moments altogether and constrain the family-related model parameters to be zero. In both cases, by ignoring the sorting of families into communities, the community effects should capture not only the effects of communities but also pick up the influence of families.

We present the results from these restricted models in Columns (2) and (3) of Table 5. Column (2) shows that the incidence of the community effect over the overall sibling correlation grows from 6 percent in the baseline (=0.02/0.32) to 14 percent (=0.045/0.31) in the sample where siblings always share the community. In Column (3), we present the results when we estimate the model excluding family information altogether. The correlation between non-sibling peers, that includes community and sorting effects, increases to 0.07, or 21 percent of the baseline sibling correlation. These restricted models suggest that, when family is ignored, the community effects are upward biased by a factor of three and are closer to the estimates reported in studies which do not account for sorting.

### 5.3 Robustness of identification

In this sub-section, we check the validity of our identification strategy by exploiting differences in the *timing of family mobility* and focusing only on siblings whose families have moved across parishes. Restricting the analysis only to movers can be argued to be less prone to selection. Specifically, we use additional information on the residential location of parents when the siblings were younger than age 15, which is available only for younger cohorts. In particular, the more recent the cohorts we consider, the younger we can first observe parental residence. We focus on cohorts born after 1965 for whom we know the parental parish back to age 10. We use mother's parish and we fill in missing information with father's parish.[21]

Using the information on parental parish at age 10, within the selected birth cohorts, we can partition siblings into three groups. The first group consists of siblings who reside in different parishes at age 15. We label them as "*late movers*" because family mobility occurs after the 15th birthday of the older brother and before the 15th birthday of the younger brother. The second group consists of siblings who reside in the same parish at age 15 but for whom the parental parish of the older brother changed between ages 10 and 15. We label this group of siblings as "*early movers*" because family mobility occurs before the 15th birthday of the older brother. This second group of "early movers" differs from the group of "late movers" because of the timing of family mobility. Finally, there is the group in which parental parish of the older brother did not change from age 10 to 15. These are families that never moved, or moved before the older brother turned 10 and we label them as "*stayers*". Arguably, there is more similarity between "early movers" and "late movers" than there is between "late movers" and "stayers". To check the robustness of the identification to selectivity between movers and stayers we repeat the analysis by restricting the sample only to "early" and "late" movers.

We present the results of this robustness analysis in Columns (4) and (5) of Table 5. Column (4) replicates the baseline model but limits the sample to the cohorts born after 1965 to provide the correct benchmark for the robustness check. Column (5) reports results based on the same cohorts as in Column (4) but restricting the sample only to movers.

We find that more recent cohorts are characterized by larger community effects compared to the baseline, but they are still limited below 10 percent of the overall sibling correlation. When we exploit only the contrast between "early movers" and "late movers" we find that the part of sibling

[21] Parental parishes when the individual is age 15 coincide with individual parish at the same age for 95% of the cases.

correlation accounted for by community effects is equal to 14 percent (0.045/0.32), which is still much lower than previous studies of sibling correlations. This evidence suggests that including "stayer" families in the baseline sample lowers the estimated community effect. However, it does not alter our main result that family accounts for most the sibling correlation and that community effects do not persist in the long run. This conclusion is supported by looking at Figure 5, which reports the life-cycle decomposition of the sibling correlation for cohorts born after 1965, both for the full sample and for the sample of movers. The two graphs are very similar and, in particular, the long-term level of the community effect is essentially zero, no matter which sample one considers.

### 5.4 Sensitivity analysis

We subject our results to several sensitivity checks. We first estimate the model for different family sizes (up to 2 or up to 3 children) and we exclude singletons. We report in Table 6 the average over the life cycle sibling correlation and its decomposition for each of these sensitivity checks. Overall, these different sample selections do not alter the main conclusion from the baseline model.

We also check the sensitivity of our findings by varying the degree of exposure and the definition of youth communities. One potential concern with the baseline model may be related to the definition of community, which is based on membership only in one single year. By defining communities only at age 15 we might miss part of the community effects due to potentially limited exposure (see also Gibbons, Silva, Weinhardt, 2013, and Chetty, Hendren and Katz, 2016, for similar discussions). To address this we re-estimate the model using two alternative criteria for 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 those sharing schools and neighborhoods, respectively, during both ages 14 and 15. Second, we define the neighborhood as the prevalent parish of residence between ages 14 and 18.[22]

As we report in Table 6, none of these alternative definitions alters our finding that community effects account for only a limited share of the sibling correlation in earnings. Defining non-sibling peers as those sharing schools and neighborhoods both at age 14 and 15 yields an average correlation of permanent earnings between schoolmates equal to 0.009 (s.e. 0.009), and an average over the life cycle correlation between neighbors equal to 0.011 (s.e. 0.010). Similarly, defining youth neighborhoods as the parish in which individuals lived most frequently between the ages of 14 and 18, we find an average correlation of permanent earnings between neighbors equal to 0.006 (s.e. 0.010), and an average correlation between schoolmates equal to 0.013 (s.e. 0.009). We also report in Table 6 further robustness checks excluding private schools, or using the ZIP code of residence at age 15 to define neighborhoods instead of parishes. In all cases we find our baseline results to be very robust.

Another concern with the community definition might be that age 15 lies at the upper bound of the age ranges that have been used by other papers in the literature. To address this concern, we use the sample of cohorts born after 1965 for which we can observe parental parish back to the year in which the individual was age 10. We report these results in Table 7. Moving progressively back in time, the age in which we define parish affiliation does not affect the main conclusion that the share of sibling correlation accounted for by community effects is generally limited below 10 percent. Alternatively, we can use these cohorts in a model where community is defined as the neighborhood of residence at ages 10 to 12 (excluding cases changing residence in that age range) to see if there is any additional impact of prolonged early exposure. Even with this

[22] More than three quarters of individuals in our sample (76.5%) do not change parish of residence between ages 14 and 18. We cannot apply a similar definition to schools because of compulsory schooling ending typically at age 15.

definition we find a limited role of community, which accounts for 8% of the overall sibling correlation.

## 6. Conclusion

Exploiting population-based administrative data for Denmark, by virtue of which we can link earnings records of siblings, schoolmates and parish neighbors, we analyze the relative influence of family, schools and youth neighborhoods on earnings inequality. We develop and estimate a model of the joint earnings dynamics of multiple groups of individuals, which accounts for sorting and allows us to decompose for the first time the sibling correlation of earnings into family, neighborhood and school effects over the life cycle. To separate the influence of family and community attributes from sorting we exploit two sources of plausibly exogenous variation: i) variation in community exposure, which arises because of family mobility, and more specifically, due to differences in the timing of family mobility across siblings, and ii) variation due to the partial overlap of schools and neighborhoods, which implies that neighbors may be enrolled in different schools and schoolmates may come from different parishes.

This research design shows that, within the environment individuals grow up and live, family is the most important factor in accounting for the inequality of permanent earnings over the life cycle. Neighborhoods and schools influence earnings only early in the working life in an almost equal way, but this influence falls rapidly and becomes negligible after age 30. This implies that, when earnings can only be observed while relatively young, the influence of community on long run earnings inequality is overstated.

Our findings are based on data from Denmark, which, because of its welfare system, is typically considered to promote equality of opportunity. However, as highlighted recently by Landersø and Heckman (2016), there is much less educational mobility than income mobility in

Denmark. Low private financial returns to schooling fail to incentivize educational investments among the children of less educated parents. This is consistent with our finding that family is the most important determinant of long run earnings similarities across siblings. Communities seem to affect earnings early in the working life, for example through peers influencing educational choices or youth behaviors, but these influences determine only short term deviations from an earnings profile that – apart from idiosyncratic abilities – mainly reflects characteristics and choices of the family. As administrative datasets and cohort studies mature in other countries, our approach could be applied to measure family and community effects on long run outcomes in other contexts.

## 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.

Björklund, Anders, Tor Eriksson, Markus Jännti, Oddbjrön Raaum and Eva Österbacka. 2002. "Brother Correlations in Earnings in Denmark, Finland, Norway and Sweden Compared to the United States." *Journal of Population Economics*, 15:757–72.

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.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2013. "Under Pressure? The Effect of Peers on Outcomes of Young Adults." *Journal of Labor Economics*, 31 (1): 119-153.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2008. "Consumption Inequality and Partial Insurance." *American Economic Review*, 98 (5): 1887-1921.

Bohlmark, Anders, and Matthew J. Lindquist. 2006. "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden." *Journal of Labor Economics*, 24 (4): 879-900.

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, and Nathalien Hendren. 2015. "The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County-Level Estimates." Harvard University, mimeo.

Chetty, Raj, Nathalien Hendren, and Lawrence F. Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review*, 106(4): 855-902.

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.

Falch, T., B. Strom, and A.M.J. Sandsor (2015), "Do Smaller Classes Always Improve Students' Long Run Outcomes?" Norwegian University of Science and Technology Working Paper 3.

Francesconi, Marco, and James J. Heckman. 2016. "Symposium on Child Development and Parental Investment: Introduction." Institute for the Study of Labor Discussion Paper No. 9977.

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, ed. 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.

Landersø, Rasmus and James J. Heckman. 2016. "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the U.S." Forthcoming, *Scandinavian Journal of Economics*.

Leuven, Edwin and Sturla A. Løkken. 2015. "Long Term Impacts of Class Size in Compulsory Schooling." University of Oslo mimeo.

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 from 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.

Nybom, Martin, and Jan Stuhler. 2016. "Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation." *Journal of Human Resources*, 51 (1): 239-68.

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.

Carsten Bøcker Pedersen, Heine Gøtzsche, Jørgen Østrup Møller and Preben Bo Mortensen. 2006. "The Danish Civil Registration System. A cohort of eight million persons." *Danish Medical Bulletin*, 53 (4): 441-9.

Raaum, Oddbjrön, 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*, ed. 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**

| (1) Birth cohorts | (2) First year observed | (3) # years observed | (4) Last age observed | (5) Earnings Observations | (6) Persons | (7) School cohorts | (8) Parish cohorts |
|---|---|---|---|---|---|---|---|
| 1960-61 | 1984 | 28 | 51 | 1,468,021 | 61,366 | 2,402 | 3,762 |
| 1962-63 | 1986 | 26 | 49 | 1,438,443 | 64,167 | 2,514 | 3,950 |
| 1964-65 | 1988 | 24 | 47 | 1,432,075 | 68,440 | 2,724 | 4,081 |
| 1966-67 | 1990 | 22 | 45 | 1,326,020 | 69,213 | 2,873 | 4,078 |
| 1968-69 | 1992 | 20 | 43 | 1,077,084 | 61,503 | 2,896 | 4,052 |
| 1970-71 | 1994 | 18 | 41 | 981,639 | 61,724 | 2,933 | 4,033 |
| 1972-73 | 1996 | 16 | 39 | 879,885 | 61,720 | 2,964 | 4,005 |
| 1974-75 | 1998 | 14 | 37 | 749,119 | 59,731 | 2,958 | 3,991 |
| 1976-77 | 2000 | 12 | 35 | 579,591 | 53,530 | 2,909 | 3,958 |
| 1978-79 | 2002 | 10 | 33 | 452,792 | 49,850 | 2,903 | 3,954 |
| 1980-81 | 2004 | 8 | 31 | 327,221 | 44,426 | 2,913 | 3,900 |
| 1982-83 | 2006 | 6 | 29 | 227,969 | 40,320 | 2,918 | 3,858 |
| 1960-83 | 1984-2006 | 6-28 | 29-51 | 10,930,859 | 695,960 | 33,907 | 47,622 |

The table reports sample characteristics by birth cohort for men born 1960-1983. Schools are defined using school of enrollment at age 15; neighborhoods are defined using the parish of residence at age 15. Column 7 reports the number of year-of-birth-by-school clusters within each birth cohort; Column 8 reports the number of year-of-birth-by-parish clusters within each birth cohort.

**Table 2**
**Neighborhood and long run earnings – key study characteristics.**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Page and Solon (2003) | Raaum, et.al. (2006) | Oreopoulos (2003) | Our study |
| Location | United States | Norway | Toronto, Canada | Denmark |
| Neighborhood | PSID cluster | Census tract | Housing project | Parish |
| Proximity | 20-30 dwellings | 44 km$^2$ | 20 buildings | 20 km$^2$ |
| #Clusters | 120 | 7,996 and 8,818 | 81 | 47,622 |
| #Men observed | 443 | 228,700 | 4,060 | 695,960 |
| Men/cluster | 4 | 14 | 50 | 15 |
| Others/cluster | 86 | 450 | 1,036 | 2,443 |
| Exposures | | | | |
| Birth cohorts | 1952-62 | 1946-65 | 1963-70 | 1960-84 |
| Years | 1968 | 1960 and 1970 | 1978-86 | 1975-99 |
| Ages | 6-16 | 5-15 | 8-16 | 15 |
| Duration | snapshot | snapshot | 1-9 years | snapshot |
| Outcomes | | | | |
| Measure | Earnings | Residual earnings | Income | Residual earnings |
| Duration (years) | 5 | 6 | 3 | 6-28 (mean 16) |
| Transformation | total mean | total mean | total mean | untransformed |
| Years observed | 1987-91 | 1990-95 | 1997-99 | 1984-2011 |
| Ages observed | 25-39 | 25-50 | 27-36 | 24-51 |

**Table 3**
**Parameter estimates of permanent earnings**

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

| | Coeff. | s.e. |
|---|---|---|
| | Variance of intercepts | |
| Family ($\sigma^2_{\mu\Phi}$) | 0.1482 | 0.0241 |
| School ($\sigma^2_{\mu\Sigma}$) | 0.0174 | 0.0083 |
| Neighborhood ($\sigma^2_{\mu\mathrm{N}}$) | 0.0203 | 0.0095 |
| | Variance of slopes | |
| Family ($\sigma^2_{\gamma\Phi}$) | 0.0026 | 0.0004 |
| School ($\sigma^2_{\gamma\Sigma}$) | 0.0002 | 0.0001 |
| Neighborhood ($\sigma^2_{\gamma\mathrm{N}}$) | 0.0005 | 0.0002 |
| | Covariance intercepts-slopes | |
| Family ($\sigma_{\mu\gamma\Phi}$) | -0.0117 | 0.0020 |
| School ($\sigma_{\mu\gamma\Sigma}$) | -0.0030 | 0.0009 |
| Neighborhood ($\sigma_{\mu\gamma\mathrm{N}}$) | -0.0053 | 0.0012 |
| | Covariance between components | |
| Family-School ($\sigma_{\Phi\Sigma}$) | 0.0074 | 0.0034 |
| Family-Neighborhood ($\sigma_{\Phi\mathrm{N}}$) | 0.0101 | 0.0039 |
| School- Neighborhood ($\sigma_{\Sigma\mathrm{N}}$) | 0.0050 | 0.0009 |

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

| | Coef. | s.e. |
|---|---|---|
| | Initial condition (age 24) | |
| Brother 1 ($\sigma^2_{\omega 24,1}$) | 0.0822 | 0.0134 |
| Brother2 ($\sigma^2_{\omega 24,2}$) | 0.0731 | 0.0129 |
| | Variance of innovations | |
| Brother 1 ($\sigma^2_{\xi 1}$) | 0.0730 | 0.0127 |
| Brother 2 ($\sigma^2_{\xi 2}$) | 0.0593 | 0.0101 |

The table reports Equally-Weighted Minimum Distance estimates for the parameters of the permanent component of the earnings process. Panel A reports parameter estimates for the earnings components shared by siblings, whereas Panel B reports parameter estimates for sibling-specific components. Estimates are derived using 65,199 empirical variances and covariances.

**Table 4**
**Parameter estimates of transitory earnings**

| | Coeff. | s.e. |
|---|---|---|
| | **Initial condition (age 24)** | |
| Brother 1 ($\sigma^2_{24,1}$) | 0.6385 | 0.0298 |
| Brother 2 ($\sigma^2_{24,2}$) | 0.6062 | 0.0293 |
| | **Variance of innovations at 25** | |
| Brother 1 ($\sigma^2_{\varepsilon 1}$) | 0.4786 | 0.0246 |
| Brother 2 ($\sigma^2_{\varepsilon 2}$) | 0.4655 | 0.0248 |
| | **Age splines in variance of innovations** | |
| Brother 1 | | |
| 26-28 | -0.1595 | 0.0047 |
| 29-33 | -0.1318 | 0.0035 |
| 34-38 | 0.0447 | 0.0060 |
| 39-43 | 0.0507 | 0.0045 |
| 44-51 | 0.0255 | 0.0041 |
| Brother 2 | | |
| 26-28 | -0.1575 | 0.0067 |
| 29-33 | -0.1290 | 0.0057 |
| 34-38 | 0.0294 | 0.0084 |
| 39-43 | 0.0580 | 0.0084 |
| 44-51 | 0.0314 | 0.0117 |
| | **Autoregressive coefficient** | |
| Brother 1 ($\rho_1$) | 0.4418 | 0.0035 |
| Brother 2 ($\rho_2$) | 0.4511 | 0.0045 |
| | **Cross-person associations in transitory earnings** | |
| Sibling covariance of innovations ($\sigma_f$) | 0.0009 | 0.0014 |
| Peers covariance of transitory earnings (catch-all components) | | |
| Sharing both school and neighborhood ($\lambda_{sn}$) | -0.0010 | 0.0008 |
| Sharing only school ($\lambda_s$ ) | 0.0019 | 0.0009 |
| Sharing only neighborhood ($\lambda_n$ ) | -0.0010 | 0.0010 |

The table reports Equally-Weighted Minimum Distance estimates for the parameters of the permanent component of the earnings process. Estimates are derived using 65,199 empirical variances and covariances.

**Table 5**
**Decomposition of average sibling correlation: baseline, ignoring sorting and robustness to identification**

| | (1) Baseline | | (2) Only Siblings Sharing Community | | (3) Without Siblings | | (4) Cohorts 1966-1983 | | (5) Cohorts 1966-1983 With only Movers | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Corr. | s.e. | Corr. | s.e. | Corr. | s.e. | Corr. | s.e. | Corr. | s.e. |
| Siblings | 0.319 | 0.010 | 0.313 | 0.011 | | | 0.299 | 0.008 | 0.316 | 0.013 |
| Family | 0.298 | 0.012 | 0.268 | 0.011 | | | 0.272 | 0.011 | 0.271 | 0.013 |
| Neighborhood (N) | 0.010 | 0.010 | 0.015 | 0.001 | 0.026 | 0.001 | 0.004 | 0.011 | 0.011 | 0.012 |
| School (S) | 0.011 | 0.010 | 0.029 | 0.001 | 0.044 | 0.001 | 0.023 | 0.010 | 0.034 | 0.011 |
| Community (N+S) | 0.021 | 0.009 | 0.045 | 0.001 | 0.071 | 0.001 | 0.027 | 0.010 | 0.045 | 0.012 |

The table reports the predicted sibling correlation for the case of siblings sharing the community and its decomposition into family and community effects, for the main estimating sample in Column 1 and in various subsamples in Columns 2 to 5. Predictions are generated using the formulae provided in Section 2.4.

**Table 6**
**Decomposition of average sibling correlation: robustness to sample selection and definition of community**

| | (1) Siblings | (2) Family | (3) Neighborhood | (4) School | (5) Community |
|---|---|---|---|---|---|
| Baseline | 0.319 | 0.298 | 0.01 | 0.011 | 0.021 |
| | (0.010) | (0.012) | (0.010) | (0.010) | (0.009) |
| Up to 2 children | 0.344 | 0.318 | 0.017 | 0.009 | 0.026 |
| | (0.016) | (0.021) | (0.015) | (0.012) | (0.015) |
| Up to 3 children | 0.296 | 0.286 | 0.020 | -0.010 | 0.010 |
| | (0.013) | (0.016) | (0.015) | (0.012) | (0.012) |
| Excluding singletons | 0.329 | 0.308 | 0.010 | 0.011 | 0.021 |
| | (0.011) | (0.013) | (0.010) | (0.009) | (0.009) |
| Excluding private schools | 0.311 | 0.291 | 0.005 | 0.015 | 0.020 |
| | (0.011) | (0.013) | (0.011) | (0.010) | (0.010) |
| Peers at 14 and 15 | 0.319 | 0.299 | 0.011 | 0.009 | 0.021 |
| | (0.010) | (0.012) | (0.010) | (0.009) | (0.009) |
| Main parish of residence 14-18 | 0.319 | 0.300 | 0.006 | 0.013 | 0.019 |
| | (0.011) | (0.013) | (0.010) | (0.009) | (0.009) |
| ZIP codes | 0.322 | 0.313 | -0.015 | 0.024 | 0.008 |
| | (0.011) | (0.014) | (0.012) | (0.010) | (0.010) |

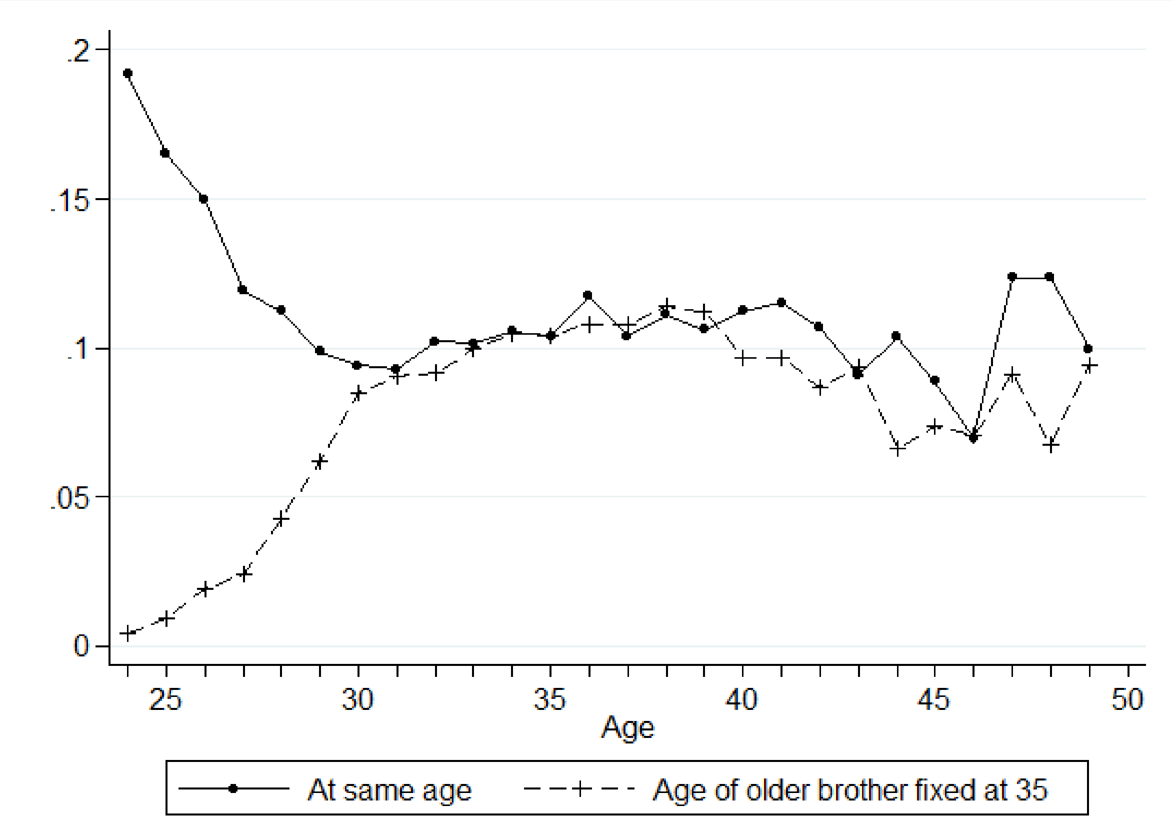
The table reports the predicted sibling correlation for the case of siblings sharing the community and its decomposition into family and community effects, providing a sensitivity analysis of the main decomposition to various sample selections and community definitions. Predictions are generated using the formulae provided in Section 2.4. Standard errors in parentheses.

**Table 7**
**Decomposition of average sibling correlation: robustness to age used to define neighborhoods**

| | (1) Siblings | (2) Family | (3) Neighborhood | (4) School | (5) Community |
|---|---|---|---|---|---|
| Baseline on cohorts | 0.299 | 0.272 | 0.004 | 0.023 | 0.027 |
| 1966-1983 | (0.008) | (0.011) | (0.011) | (0.010) | (0.010) |
| Parental parish | 0.329 | 0.307 | 0.004 | 0.018 | 0.022 |
| when 15 | (0.012) | (0.016) | (0.013) | (0.012) | (0.013) |
| Parental parish | 0.356 | 0.325 | 0.011 | 0.020 | 0.031 |
| when 14 | (0.016) | (0.019) | (0.013) | (0.012) | (0.014) |
| Parental parish | 0.356 | 0.325 | 0.011 | 0.020 | 0.030 |
| when 13 | (0.016) | (0.019) | (0.013) | (0.012) | (0.014) |
| Parental parish | 0.349 | 0.312 | 0.019 | 0.018 | 0.037 |
| when 12 | (0.015) | (0.018) | (0.011) | (0.011) | (0.013) |
| Parental parish | 0.289 | 0.261 | 0.011 | 0.017 | 0.028 |
| when 11 | (0.009) | (0.015) | (0.012) | (0.012) | (0.015) |
| Parental parish | 0.294 | 0.267 | 0.009 | 0.018 | 0.028 |
| when 10 | (0.010) | (0.015) | (0.013) | (0.012) | (0.015) |

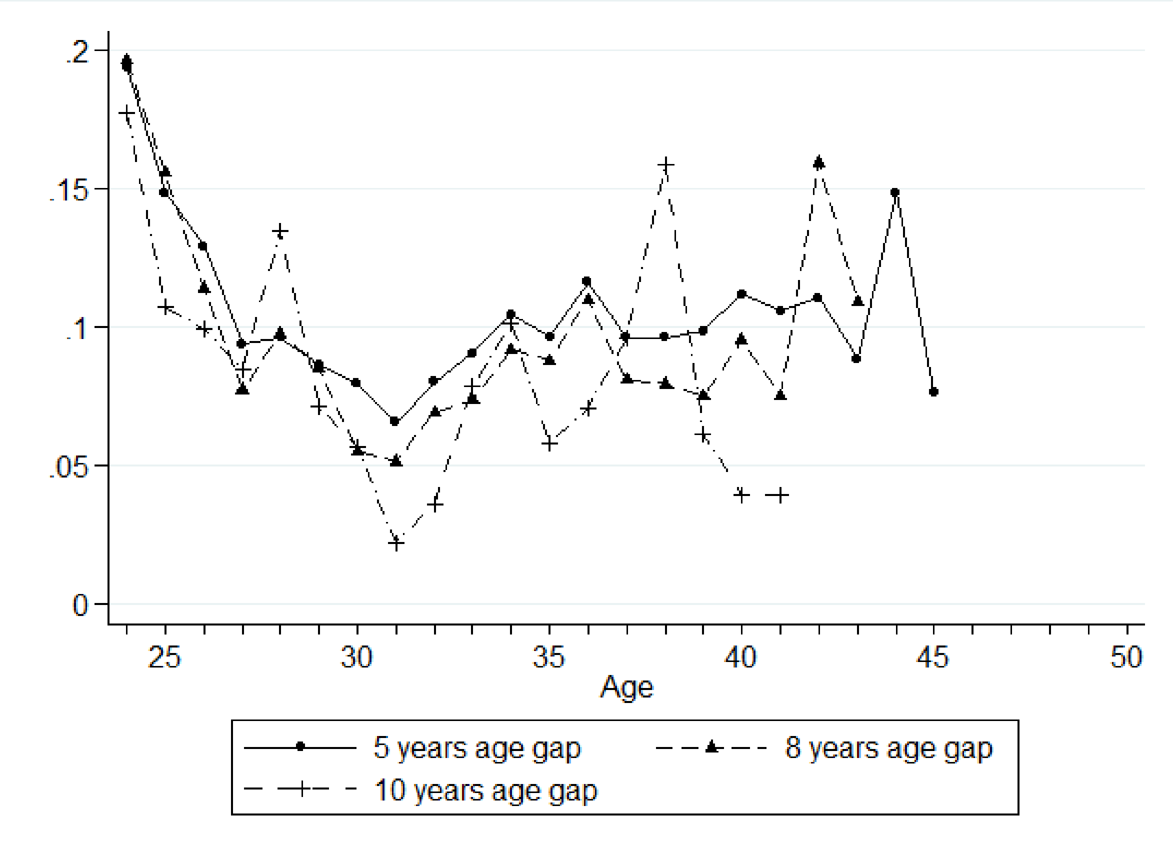
The table reports the predicted sibling correlation for the case of siblings sharing the community and its decomposition into family and community effects, providing a sensitivity analysis of the main decomposition to the age used to define neighborhoods. Predictions are generated using the formulae provided in Section 2.4. Standard errors in parentheses.

**Figure 1**
**Sibling correlation of annual earnings**



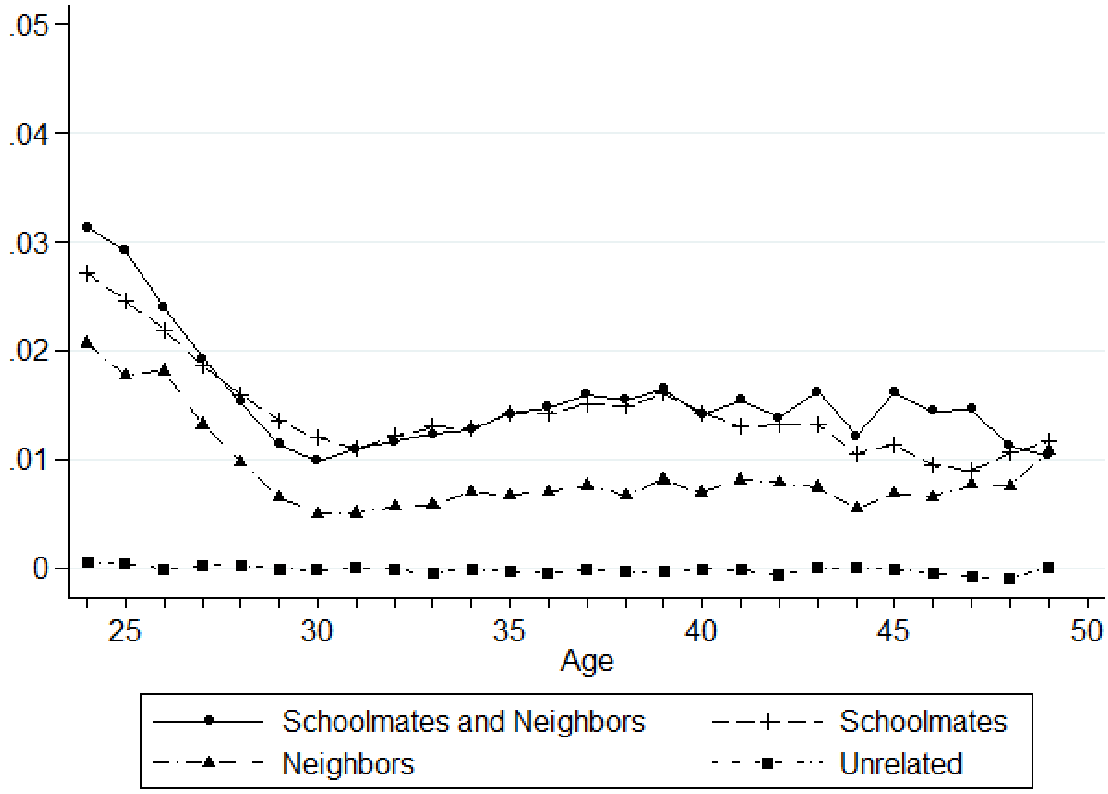

The figure shows raw sibling correlations of earnings over the life cycle. The line labelled "At same age" is obtained by computing the sibling correlation when the brothers are at the same point in their life cycle, while the line labelled "Age of older brother fixed at 35" is obtained by computing the sibling correlation when the older brother is 35.

**Figure 2**
**Sibling correlation of annual earnings by siblings' age gap**



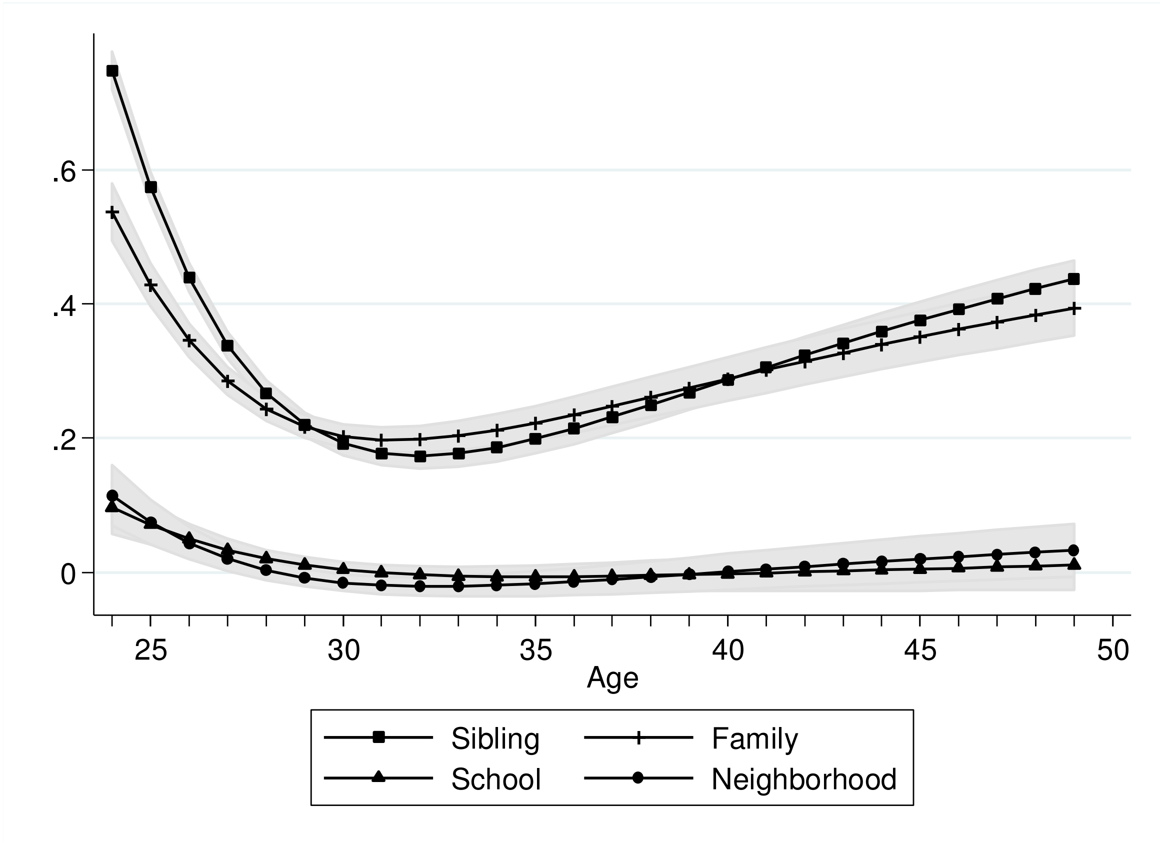

The figure shows raw sibling correlations of earnings over the life cycle. The lines are obtained by computing the sibling correlation when the brothers are at the same point in their life cycle, and refer to sibling pairs with an age gap of 5, 8 and 10 years.

**Figure 3**
**Correlation of annual earnings for members of youth communities**



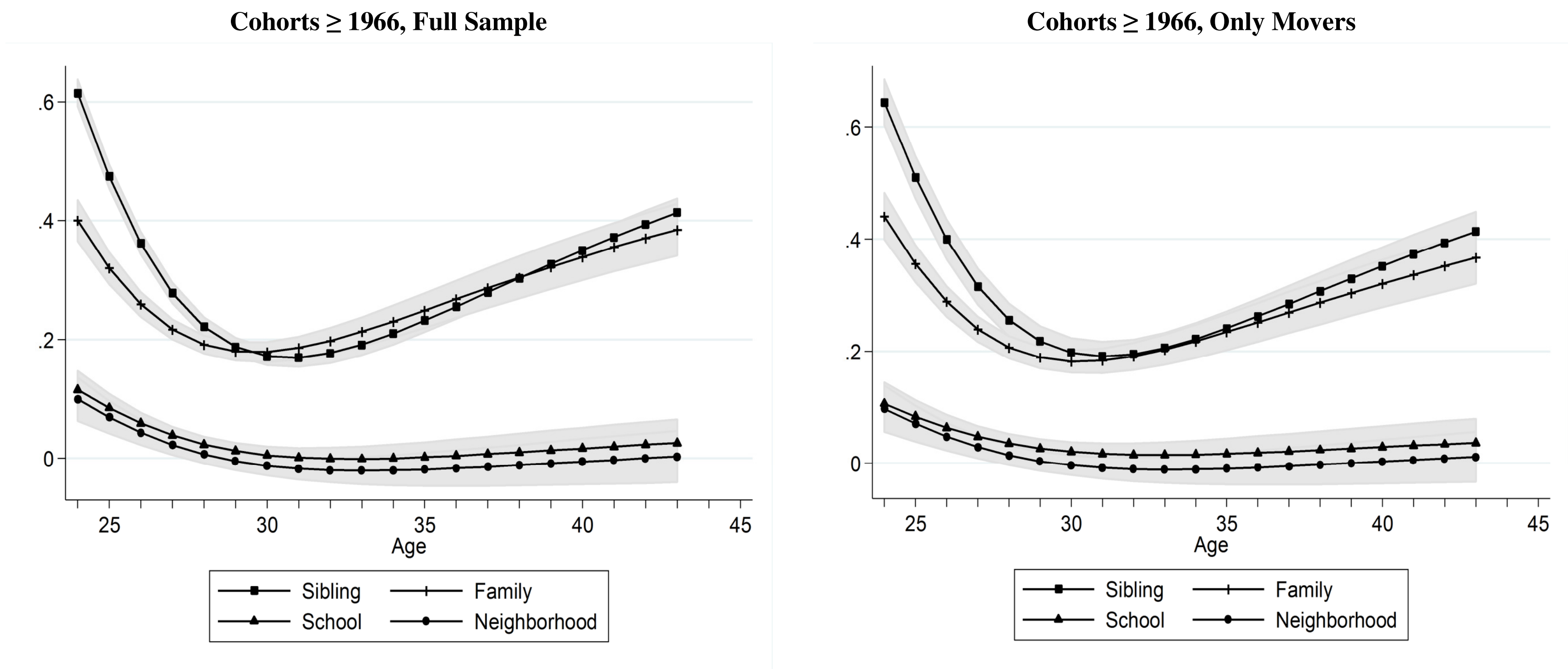

The figure shows raw correlations of earnings for non-sibling peers over the life cycle. Schoolmates are defined as individuals attending the same school at age 15. Neighbors are defined as individuals residing in the same parish at age 15. Unrelated are defined as individuals sharing neither the family nor the school at age 15; in this case correlations are computed after generating 1000 random matches from the sub-population satisfying that definition.

**Figure 4**
**Predicted sibling correlation of permanent earnings and factor decomposition**




The figure shows the predicted sibling correlation over the life cycle for the case of siblings sharing the community and its decomposition into family and community effects. Predictions are generated using the formulae provided in Section 2.4.

**Figure 5**
**Predicted sibling correlation of permanent earnings and factor decomposition: full sample vs. sample of movers**




The figure shows the predicted sibling correlation over the life cycle for the case of siblings sharing the community and its decomposition into family and community effects. Predictions are generated using the formulae provided in Section 2.4. Predictions are obtained for the full sample of cohorts born 1966 or later (left panel), or for the subsample among those cohorts whose family moved parish after the older brother turned 10 (right panel).

## 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:

$$E(v_{ifsna}v_{ifsna'}) = [I(a = a' = 24)\psi_c\sigma^2_{24b} + \\ I(a = a' > 24)(\exp(g_b(a)) + var(u_{ifsn(a-1)})\rho^2_b) + \\ I(a \neq a')(E(u_{ifsn(a-1)}u_{ifsna'})\rho_b)]\eta_t\eta_{t'} \,. \tag{A.1}$$

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 $a'$, covariances of transitory earnings across non-sibling peers are as follows:

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

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

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 2.3 of the paper. In general, these restrictions are a non-linear function of a parameter vector $\theta$. We estimate $\theta$ 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 ($\pi_t$) | | Transitory Component ($\eta_t$) | |
|---|---|---|---|---|
| | Coeff. | s.e. | Coeff. | s.e. |
| $t$= | | | | |
| 1985 | 0.9038 | 0.0712 | 0.9666 | 0.0255 |
| 1986 | 0.8603 | 0.0700 | 0.9769 | 0.0258 |
| 1987 | 0.8695 | 0.0725 | 0.9699 | 0.0292 |
| 1988 | 0.8500 | 0.0705 | 1.0053 | 0.0281 |
| 1989 | 0.8395 | 0.0707 | 1.0463 | 0.0313 |
| 1990 | 0.8479 | 0.0700 | 1.0680 | 0.0300 |
| 1991 | 0.8736 | 0.0713 | 1.0505 | 0.0326 |
| 1992 | 0.7734 | 0.0628 | 1.1302 | 0.0346 |
| 1993 | 0.7512 | 0.0615 | 1.1441 | 0.0348 |
| 1994 | 0.7178 | 0.0586 | 1.1453 | 0.0353 |
| 1995 | 0.6622 | 0.0545 | 1.0580 | 0.0330 |
| 1996 | 0.6406 | 0.0524 | 1.0678 | 0.0337 |
| 1997 | 0.5888 | 0.0482 | 1.0533 | 0.0317 |
| 1998 | 0.5636 | 0.0461 | 1.0386 | 0.0325 |
| 1999 | 0.5342 | 0.0440 | 1.0665 | 0.0321 |
| 2000 | 0.5115 | 0.0424 | 1.0972 | 0.0335 |
| 2001 | 0.4779 | 0.0398 | 1.1174 | 0.0326 |
| 2002 | 0.4643 | 0.0390 | 1.1588 | 0.0341 |
| 2003 | 0.4558 | 0.0384 | 1.1859 | 0.0349 |
| 2004 | 0.4230 | 0.0358 | 1.1488 | 0.0350 |
| 2005 | 0.3917 | 0.0333 | 1.1061 | 0.0337 |
| 2006 | 0.3481 | 0.0297 | 1.0931 | 0.0338 |
| 2007 | 0.3147 | 0.0269 | 1.0610 | 0.0329 |
| 2008 | 0.2914 | 0.0252 | 1.0676 | 0.0345 |
| 2009 | 0.2896 | 0.0250 | 1.1795 | 0.0386 |
| 2010 | 0.2764 | 0.0239 | 1.2178 | 0.0410 |
| 2011 | 0.2667 | 0.0232 | 1.1668 | 0.0401 |

The table reports Equally-Weighted Minimum Distance estimates for the time shifters. Estimates are derived using 65,199 empirical variances and covariances.

**Table A2: Parameter estimates of cohort effects (1966-67=1)**

| | Permanent Component ($\delta_c$) | | Transitory Component ($\psi_c$) | |
|---|---|---|---|---|
| | Coeff. | s.e. | Coeff. | s.e. |
| $c$= | | | | |
| 1960-61 | 0.7380 | 0.0218 | 1 | |
| 1962-63 | 0.8253 | 0.0208 | 1 | |
| 1964-65 | 0.9457 | 0.0216 | 1 | |
| 1968-69 | 1.1437 | 0.0305 | 1.0041 | 0.0399 |
| 1970-71 | 1.3462 | 0.0361 | 0.9208 | 0.0383 |
| 1972-73 | 1.5193 | 0.0408 | 0.9118 | 0.0418 |
| 1974-75 | 1.6202 | 0.0494 | 1.0180 | 0.0474 |
| 1976-77 | 1.8552 | 0.0622 | 0.9538 | 0.0454 |
| 1978-79 | 2.2424 | 0.0793 | 0.8450 | 0.0426 |
| 1980-81 | 2.5622 | 0.1052 | 0.8693 | 0.0497 |
| 1982-83 | 2.801 | 0.1486 | 0.9127 | 0.0579 |

The table reports Equally-Weighted Minimum Distance estimates for the time shifters. Estimates are derived using 65,199 empirical variances and covariances.