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

August 2014

(Work in Progress – Preliminary Version)

Abstract

Correlations between the earnings of siblings reflect shared family and community background, but evidence is mixed on the relative magnitudes of these influences. We estimate long run earnings correlations between brothers, school mates and teenage neighbors jointly in a unified framework. Using administrative data on the Danish population we find that: (1) family is by far the most relevant factor that shapes long-term earnings; (2) the contribution of neighborhood and school quality on long-term earnings is overestimated if the family component is ignored, and becomes negligible and not significantly different from zero by age 30; and (3) the importance of family declines over the life-cycle.

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

JEL codes: D31, J62

[§]The research for this paper was supported by a British Academy Research Grant (award SG113192).

1. Introduction

Understanding the determinants of long-term labor market outcomes, such as earnings, is important for identifying the driving forces of existing inequalities and for interventions that aim to reduce them. Besides idiosyncratic individual-specific abilities, the family and the community, or the environment in which individuals grow up and live, are the two main forces that shape the earnings potential of individuals (see for example the reviews by Solon, 1999; Björklund and Jännti, 2009; Black and Devereux, 2011). The family is considered to affect outcomes through the transfer of abilities, preferences and resources (Becker and Tomes, 1986). The community can determine individual outcomes through institutions such as the school and its quality (e.g. Hanushek, 2006), or through peer influences, social norms and role models in the neighborhood (e.g. Wilson, 1987).

While there is a large literature on the influence of neighborhood *or* school quality on labor market outcomes, the evidence is limited and mixed on the relative magnitude of these influences on long-term earnings. In this paper, we estimate long run earnings correlations between brothers, school mates and teenage neighbors jointly in a unified framework. Disentangling the influence of family and community on earnings dynamics is crucial for assessing the potential of community-based policies in reducing inequalities in the long-run.

There are several approaches used in the literature for the estimation of neighborhood effects. One common approach is to compare the correlation of sibling outcomes with the correlation of outcomes among unrelated neighbors (e.g. Solon et al., 2000 on educational attainment, or Page and Solon, 2003a and 2003b on earnings). An alternative method is to compare siblings who spend time in different neighborhoods because of family mobility (e.g. Aaronson, 1998 on educational outcomes). The findings in both cases suggest a substantial impact of neighborhoods on educational outcomes or earnings, but it is also recognized that these estimates are upper bounds because of non-random sorting of families into neighborhoods, which lead to positive correlation between the two factors.[1]

In order to address the sorting of families into neighborhoods a number of studies have exploited quasi-experimental variation (e.g. Gibbons et al., 2013; Jacob, 2004; Oreopoulos, 2003). The evidence from these studies suggests a small impact of neighborhoods, while most have focused on educational achievement (e.g. Gibbons et al. 2013; Jacob 2004). Oreopoulos (2003) is a rare study on the effect of neighborhood quality on earnings (among other outcomes), which combines quasi-experimental variation with the approach of comparing sibling with neighbors' earnings correlations. Exploiting quasi-random assignment of families to public housing projects in Toronto, Oreopoulos (2013) finds a zero influence of neighborhood quality in the total variance of income and wages. Evidence from the "Moving to Opportunity" project (MTO) - a randomized social experiment in the U.S. - also suggests that changes in neighborhood quality had little impact on economic outcomes (e.g. Ludwig et al., 2013).[2]

The literature on the effect of school quality on children outcomes is vast (see e.g. Hanushek, 2006 for a review). For the purpose of this study we focus on the evidence for the effect of school quality on earnings, which is generally mixed. Card and Krueger (1992) exploit variation of school quality across cohorts within U.S. regions and find that higher quality (a lower pupil/teacher ratio) increases the rate of return to schooling and earnings. The evidence of the effect of pupil/teacher ratio in the UK, however, is found to be insignificant (Dearden et al. 2002). Linking the data from the Tennessee STAR experiment with tax return data, Chetty et al. (2011) examine the effect of class size on earnings at age 27 finding no

[1] If families are randomly assigned then the correlation of sibling outcomes will provide the proportion of variance due to both neighbourhood effects and to family factors that are common between the two siblings. The neighbors' outcome correlation, instead, will give the proportion of variance due only to neighborhood effects. The comparison of the two correlations identifies separately the two components. However, without random assignment this would not be the case because of the correlation between the two components.

[2] The experiment randomly assigned eligible families living in high poverty neighborhoods to a treatment group, who could use a voucher to move to only a better neighborhood or to any neighborhood, and to a control group which were given no treatment.

effects, but they find a positive effect of teacher quality.[3] More recently, using Swedish data and exploiting a maximum class size rule, Fredriksson et al. (2013) find a positive effect of smaller class size on adult earnings at ages 27 to 42.

Although some of the previous studies have used a credible identification strategy to identify the effect of either the neighborhood or the school on earnings, there is no existing evidence which decomposes the effects of family, neighborhoods and schools on earnings over the life-cycle. Using Danish registers we link brothers to their school mates and teenage neighbors at age 15 and follow them over their life-cycle. Following the method of direct decomposition of Bingley and Cappellari (2014), we develop a model of multi-person earnings dynamics in which we distinguish permanent from transitory earnings and allow for heterogeneous earnings growth. The model specification is motivated by the human capital model of Mincer (1958) and Ben-Porath (1967) in which heterogeneity of initial earnings and heterogeneous earnings growth is generated by differential investments. We use this specification to model the components of permanent earnings that are shared between siblings, neighbors and schoolmates. Given the richness of the data and the proposed model we can jointly decompose the sibling correlation of earnings over the life-cycle into three components: the family, the school and the neighborhood.

We find that the family is the most important factor that shapes long-term earnings. The correlation of earnings between neighbors and schoolmates are measured around zero on average over the life-cycle. Ignoring the family component and estimating the model only considering the community factor leads to an overestimation of the influence of neighborhoods and schools on long-run earnings. This suggest that by jointly estimating the family and community factors within our model we are able to take into account the sorting that would otherwise lead to biased estimates of the community influence. Although the

[3] The STAR experiment randomly assigned students and their teachers to classrooms of different size in grade K-3.

community effects are measured around zero on average over the life-cycle, there is some contribution from community effects at the beginning of the life cycle, but this becomes negligible and not significantly different from zero by age 30. Finally, we find that the importance of family declines over the life-cycle.

The paper is structured as follows. In Section 2 we discuss the data and the way we identify neighbors and schoolmates, while in Section 3 we present descriptive statistics on earnings of siblings and peers. In Section 4 we develop the econometric model for assessing the relative importance of families, schools and neighborhoods within the sibling correlation, based on the joint analysis of life-cycle earnings for brothers, schoolmates and neighbors. The main results are presented in Section 5. In Section 6 we discuss robustness and in Section 7 we present additional results by birth order, class size and by urban/rural residency. We conclude in the last section.

2. Data

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

In order to establish our dataset of brothers, we sample fathers born from 1940 and consider only sons born to first father-mother pair, conditional on father's age at first birth being 18 or older. We drop sons who were adopted before age 17, and sons who are themselves observed as fathers. First sons and second sons are included, and subsequent sons (4 percent) are ignored. Brothers born less than 12 months apart are dropped. Second sons are dropped if they are born more than 12 years after the first. Finally we derive a sample of father/first-son/second-son triads and father/first-son couples. Women play no role in the main analysis after determining full brotherhood, and son birth order is determined irrespective of daughters. We select first sons born 1960-1982 and second sons born 1962-1982. This is because of completeness of registered parentage and the small number of first sons observed born before 1960.

Next we link our sampled brothers to their teenage communities. School attendance rules were such that pupils should start in first grade in the August of the calendar year they turn seven. Attendance through to ninth grade is compulsory, and early or late school start and grade retention were uncommon (less than 10 percent), meaning most pupils begin the final year of compulsory schooling in the calendar year they turn fifteen. We define school mates as all males attending the same school on 31 October of the calendar year they turn fifteen. Neighbors are defined as all males who are resident in the same parish on 31 October of the calendar year they turn 15. [4]

We use pre-tax annual labor earnings. In order to model life-cycle dynamics we require observation of individual earnings strings over time. We group the data in 2-years birth cohorts; as in Baker and Solon (2003) we compute age by imputing each cohort with its earlier year of birth. We select birth cohorts so that each cohort is observed for at least 5 years. In practice, the shortest span of observation is the one of cohort 1982, corresponding to the five years between 2007 and 2011. We require availability of at least 3 consecutive earnings observations at the individual level.

We provide a description of the structure of peer clusters in Table 1, which reports the number of clusters and within-cluster density by cluster type, for selected cohorts. Clusters

[4] We assess the robustness of our results to this specific choice by defining peers as either those born in the same year and sharing schools and/or neighborhoods at age 14, or as those born in the same year and sharing schools and/or neighborhoods at both age 14 and age 15. Results were not affected by defining peers in these alternative ways.

defined by the intersection of school and neighborhood are, not surprisingly, more numerous and have smaller density. The smallest number of clusters is observed for individuals sharing only the school. However these clusters do not have the largest density, which is a characteristic of clusters of individuals sharing only the neighborhood. On average, clusters defined on both school and neighborhood contain about 12 individuals, with a maximum of 64 individuals; corresponding numbers are 25 and 111 for clusters of individuals sharing only the school, and 32 and 134 for clusters of individuals sharing only the neighborhood. The minimum number of individuals per cluster is 1 in any case. Irrespective of cluster type, younger cohorts have lower density, which reflects the fact that individuals have to appear in the earnings registry to be counted in Table 1, and this is more likely for the older individuals.

3. Descriptive statistics on earnings of siblings and peers

In this section we provide a description of the interpersonal covariance structure of earnings. There are two types of cross-person relationships that are of interest to our analysis: between members of the same family (brothers) and between members of the same youth community (peers), the community being defined as either the school one attended at age 15 or the neighborhood (parish) one lived in at the same age. Brothers' earnings covariances are estimated from families with two male children, which as mentioned above represent 96% of the population of families with male children. We use the same sample to group non-sibling peers in clusters depending on whether they shared the school and the neighborhood, only the school, or only the neighborhood. As we saw in the last section, the number of clusters varies with the type of association.[5] We obtain the between-peers covariance of earnings (at each relevant age) by first computing the within-cluster covariance and then averaging covariances between clusters using the weighting scheme of Page and Solon (2003, pp. 840), which gives

[5] Changing the sample by using also information on first and second brothers coming from families with more than two male children, or by including any individuals irrespective of birth order in the sample of community peers had hardly any impact on the estimated earnings covariances.

more importance to more populated clusters.[6] We obtain empirical moments by cohorts and use moments for all cohorts jointly in estimation so as to separate time and cohort effects. As in Baker and Solon (2003) birth cohorts are defined with 2-year groupings on the birth year, each cohort being conventionally imputed the larger age in each year.

We begin by describing the sibling earnings covariance by age in Figure 1. The plot labeled "Same age" reports the estimated covariance when the brothers are at the same point in their life cycle, a counterfactual that is available in our data. The earnings covariance declines between age 24 and 30, and remains stable after age 30. The decline suggests that sources of initial earnings heterogeneity that are shared between brothers are negatively correlated with heterogeneity in earnings growth. Human capital models predict investments in education or training to induce such a negative covariance: the graph suggests that whatever the source of the negative relationship, they might be shared between brothers. The second plot labeled "Age B1=35" fixes the age of the older among the two brothers at 35 and reports the sibling covariance by age of the younger brother. This time, the earnings covariance in the second plot, is relatively low at age 24 (actually close to zero) and increases sharply so that by the early-30s it matches "Same age" figures. This pattern illustrates that the earnings covariance estimated between siblings of different ages is an underestimate of the covariance one would obtain observing siblings at the same point in the life-cycle, a form of life-cycle bias as discussed in Jenkins (1987) and Haider and Solon (2006). The figure shows that our data allow observing the bias, 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 lifecycles in Figure 1 may also reflect the correlation of transitory shocks. It is well known that earnings instability is large for young cohorts (see e.g. Baker and Solon,

[6] We apply the same scheme also to earnings fourth moments that we exploit to correct the variance of the GMM estimator in the next section.

2003) and it is plausible that siblings are subject to common shocks, for example because of similar local economic conditions at labor market entry. As a way to assess if the relatively large sibling correlation at young ages is driven by permanent earnings differences or transitory fluctuations, we also computed sibling covariances for brothers born at least five, eight or ten years apart, which are shown in Figure 2. The larger the age difference, the less likely it is that brothers entered the same labor market and shared transitory shocks at entry, so that these samples are less likely to be influenced by transitory fluctuations compared with the samples underlying Figure 1. A declining pattern of the sibling covariance between the mid-20s and the early-30s persist even after excluding closely spaced brothers that most likely share transitory earnings fluctuations. This suggests that the source of the convex evolution of sibling covariances is in the permanent earnings component.

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

4. Econometric model

We develop an econometric model for assessing the relative importance of families, schools and neighborhoods within the sibling correlation, based on the joint analysis of life-cycle earnings for brothers, schoolmates and neighbors. We define schoolmates as individuals attending the same school at the age of 15, which corresponds to the end of compulsory education in Denmark. Neighborhood is defined as the parish of residence (still at age 15), which is consistently recorded during our sample period.

We tackle the estimating challenges highlighted in the literature (transitory shocks and life-cycle biases) with a model of multi-person earnings dynamics distinguishing permanent from transitory earnings and allowing for heterogeneous earnings growth. We specify individual earnings as:

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

where the indices *i, f, s, n* and *a* stand for individual, family, school, neighborhood and age, respectively.[7] The log of age- and time-adjusted gross annual earnings, denoted by *w* are assumed to be the sum of two components: a permanent one denoted by (*y*) and a transitory one denoted by (*v*), which are orthogonal by definition. Separate identification of permanent and transitory earnings is granted by the availability of individual level panel data and ensures that we estimate correlations in permanent earnings, avoiding measurement error biases due to transitory shocks.

4.1 Specification of permanent earnings

We allow permanent earnings to depend on both shared and idiosyncratic components. Shared components capture those determinants of permanent earnings that are common between brothers, schoolmates and neighbors. The idiosyncratic component represents individual-specific sources of variation in permanent earnings. We model life-cycle dynamics of shared

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

components using a heterogeneous income profile (HIP) specification, 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.

HIP specifications are inspired by human capital models a-là Mincer (1958) and Ben-Porath (1967) in which heterogeneity of initial earnings and heterogeneous earnings growth are generated by differential investments. More specifically, these models predict that heterogeneous investments in human capital induce a negative correlation between initial earnings and earnings growth rates, because investors trade off initial earnings against earnings growth throughout the life cycle, and the HIP specification allows for these features of life-cycle earnings. Combining these observations with insights from the Becker and Tomes (1986) model of parental preferences for child human capital, motivates our specification choice for shared earnings determinants, reflecting the idea that earnings similarities across individuals stem from similarities in social background and human capital investments. We saw in Section 3 that life-cycle patterns of earnings covariances between siblings and peers are consistent with these mechanisms.

Besides the earnings profile shared by siblings and peers, we assume permanent earnings to follow a unit root, capturing long-term individual deviations from the shared profile. This represents idiosyncratic ability which is revealed over time, either to the labor market or to individuals themselves. Overall, our permanent earnings model is specified as follows:

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

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

where $c(i)$ is the birth cohort of person i and π_t is a calendar time shifter, allowing for the possibility of aggregate changes of the permanent earnings process over time. We separate time and age effects exploiting earnings variances and covariances computed within 2-year birth cohorts, and use them jointly in estimation, where each cohort is imputed its earlier year of birth.

The parameters of the individual-specific linear profile are factored into three zero-mean components. Their variances capture family, school and neighborhood heterogeneity in initial earnings (the μs) and life-cycle earnings growth (the γs). We assume the shared earnings components to be correlated both within each dimension of heterogeneity and across dimensions. While previous studies of neighborhood and sibling correlations have acknowledged the importance of family-community covariances, we are the first to actually estimate them (see Page and Solon, 2003, and Oreopoulos, 2003). By allowing for these covariances across the shared components in the model, we explicitly take into account the sorting of families across communities. On the other hand, by allowing for correlation of initial and growth rate heterogeneity within components, we allow for the possibility that human capital investments lower initial earnings and raise their life-cycle growth rates, as predicted by human capital models. In that case, the resulting negative covariance of intercepts and growth rates would generate a u-shaped evolution of earnings dispersion by age due to the 'Mincerian cross-overs' of earnings profiles. The empirical profiles of earnings cross-person covariances shown in Figures 1, 2 and 3 provide support for tis choice of specification.

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

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

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

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

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

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

where *b* indexes brothers and idiosyncratic parameters are allowed to be different between brothers. Correlation across family and community effects is allowed through the intercepts of the individual-specific profiles (assumption (3.d)), both because empirically most of the community effects vanish after two or three years (see Figure 4), and so as not to overcrowd the parameter space.

These assumptions fully specify the intertemporal and interpersonal distribution of permanent earnings. Identification of parameters is achieved through exploiting different types of moment restrictions generated by the model. For a given individual, moment restrictions for two time periods are a function of all sources of earnings heterogeneity: idiosyncratic, family, school and neighborhood. Cross-persons moments do not depend on idiosyncratic heterogeneity. Sibling covariances depend on family effects. Moreover, they are also functions of school effects, neighborhood effects, both, or none, depending on the extent to which siblings shared schools and neighborhoods. Similarly, covariances for non-sibling peers are functions of school effects, neighborhood effects, both, or none, depending on the extent to which peers shared schools and neighborhoods.[8]

4.2 Specification of transitory earnings

We model transitory earnings using an AR(1) process in order to capture serial correlation of transitory shocks. As for idiosyncratic parameters of permanent earnings, we allow the distribution of transitory earnings to be brother-specific. We account for age effects in transitory shocks through an exponential spline. We allow for correlation of shocks across persons. In particular, for brothers we model the correlation of AR(1) innovations. We cannot follow the same approach for non-sibling peers because it would require distinguishing idiosyncratic components of transitory earnings for each member of school or neighborhood clusters, generating dimensionality issues. To counter these issues we only exploit

[8] We discuss identification in detail in the Appendix of the paper.

information on between-peers covariances in transitory earnings. This is modeled through a specific "catch-all" component collapsing all the parameters of the underlying stochastic process. Our model for transitory earnings is as follows:

$$
\begin{gathered}
v_{ifsna} = \eta_t u_{ifsna};\ u_{ifsna} = \rho_b u_{ifsn(a-1)} + \varepsilon_{ifsna}; \\
\varepsilon_{ifsna} \sim \left(0, \sigma^2_{\varepsilon b} \exp\left(g_b(a)\right)\right), \qquad u_{ifsn24} \sim \left(0, \sigma^2_{24b}\right), \\
E\left(\varepsilon_{ifsna}\varepsilon_{i'f\tilde{s}\tilde{n}a}\right) = \sigma_f, \qquad E\left(u_{ifsna}\, u_{i'f'sn\tilde{a}}\right) = \lambda_{sn}^{1+|a-\tilde{a}|} \\
E\left(u_{ifsna} u_{i'f's\tilde{n}\tilde{a}}\right) = \lambda_s^{1+|a-\tilde{a}|}, \qquad E\left(u_{ifsna} u_{i'f'\tilde{s}n\tilde{a}}\right) = \lambda_n^{1+|a-\tilde{a}|}
\end{gathered} \tag{4}
$$

where η_t is a time loading factor, and $g_b(\ )$ is a brother-specific linear spline in age with knots at 28, 33, 38 and 43, "primes" denote different indices, whereas "tildes" denote indices that can be either different or equal.

5. Results

We report estimates of model parameters in Table 2 (permanent component), Table 3 (transitory component) and Table 4 (time effects on both components). We concentrate our discussion on estimates of the 'core' parameters of the HIP, RIP and AR(1) processes.

5.1 Permanent earnings correlation between siblings, schoolmates and neighbors

Parameters estimates of the shared components of permanent earnings indicate that the family is by far the most relevant factor that shapes long-term earnings (Table 2, top panel). This is true both for initial earnings and for earnings growth rates. In particular, there is no statistically significant heterogeneity in initial earnings related to schools and neighborhood on top of the sorting effects captured by the covariances between family and community components (see the discussion below). The other relevant source of permanent inequality in earnings is the individual idiosyncratic component (Table 2, lower panel).

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

The top panel of Table 2 also reports estimates of the covariances across the three components of shared earnings determinants, families, schools and neighborhoods. As already noted, these parameters absorb all the earnings heterogeneity that is related to community effects.[9] These parameters are important because they capture the sorting of families into communities; they feature in the specification of variance decomposition models in both Page and Solon (2003) and Oreopoulos (2003), but none of these papers has actually estimated them. Our results indicate that these sorting effects are relevant, as the covariance of family effects with both school and neighborhood effects is positive, sizeable and statistically significant, indicating 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. There is also a positive covariance among community effects.

We use these parameter estimates to generate predictions of the sibling correlation and its decomposition into the three factors of interest, families, schools and neighbors. The exercise is carried out in Figure 4 using the formulae provided in the Appendix. In particular

[9] Estimating the model imposing zero cross-component covariances yielded estimates of the variances of community effects that were statistically significant and of about the same size as the covariances in the unconstrained model.

we consider the case of two brothers who attended the same school and lived in the same neighborhood when they were 15, so that the resulting sibling correlation is the sum of family, school and neighborhood effects. The life-cycle pattern of the sibling correlation is u-shaped in age. In particular, the estimated correlation is about 0.6 at age 24, drops to 0.15 at age 37, and rises back to 0.34 by age 49, the last age at which we observe younger brothers. The average sibling correlation is 0.29, which is in line with previous estimates for Denmark.[10] The u-shaped pattern is a symptom of "Mincerian cross-overs" of earnings profiles: the negative intercept-slope covariances that we estimated for all the shared factors of earnings profiles implies that the distribution of shared components (and therefore the siblings and peers correlations) first shrink and then fan out over the life cycle. The u-shaped pattern was also a feature of raw cross-person covariances in Figures 1 to 3, and in particular Figure 2, which depicted siblings' earnings covariances for brothers with few years of age spacing.

Looking at the different factors behind the overall sibling correlation, it is clear that family accounts for the most of it. There is some contribution from community effects at the beginning of the life cycle, but this becomes negligible and not significantly different from zero (even insignificantly negative) by age 30. Overall, we estimate the correlation in permanent earnings across schoolmates to be 0.0075, while the equivalent figure for neighbors is 0.0007. These results clearly indicate that there is not much room for community effects in shaping the sibling correlation; the only factor that generates a substantial correlation in permanent earnings between brothers is the family.

Our findings are in line with those of Oreopoulos (2003) who used quasi-random assignment of neighbors and showed that the neighborhood correlation in adult earnings was virtually zero in a variance decomposition exercise similar in spirit to ours. Page and Solon (2003), instead, find in the PSID that the neighborhood correlation was about half of the

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

sibling correlation (0.16 versus 0.34). We can replicate the approach of Page and Solon (2003) by estimating community effects on a sample that exclude siblings and constraining family-related model parameters to be zero. The results of this exercise are reported in Figure 5 in which we plot community effects (the sum of schools and neighbors effects) from the model that ignores family effects alongside community estimated from our full model. The comparison is striking. When family effects are ignored we find a sizeable correlation among members of the youth community, which is concentrated in the initial part of the life cycle. The average correlation in this model is .068, which amounts at 23% of the sibling correlation (the ratio between neighbor and sibling correlations was 0.47 in Page and Solon, 2003). As we have already seen, the model that controls for family effects tells a radically different story about the relevance of community effects, with a correlation of permanent earnings between members of the same youth community of just 0.0082, with a factor of 10 smaller than the model that ignores family effects. The comparison depicted in Figure 5 suggests that including family effects in the model of life-cycle earnings allows controlling for the sorting of families into neighborhoods and to obtain results that are close to those from randomized studies that are specifically meant to control for that type of selection.

5.2 Transitory earnings

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

6. Robustness

We subject our results to several sensitivity checks. We begin by assessing results robustness to alternative definitions of youth communities. In particular, one concern is that our definition of youth communities based on membership at age 15 might miss part of the effects of communities due to limited exposure to peers (see also Gibbons et al., 2013, for similar discussions). We therefore re-estimate the model using two alternative criteria to define community membership, characterized by greater exposure to peers relative to the one-year definition used in the main model. First, we define peers as individual sharing schools and neighborhoods for two years, at ages 14 and 15. Second, we define neighborhoods as the prevalent parish between ages 14 and 18.[11] None of these alternative definitions alters our finding that community effects account for only a limited share of the sibling correlation in earnings. Defining peers as those sharing schools and neighborhoods both at age 14 and 15, yields an average correlation of permanent earnings between schoolmates equal to 0.007, while the neighbors' correlation is 0.002. Similarly, when we use the parish in which individuals lived most frequently between the ages of 14 and 18 as identifier of youth neighborhoods we find the average earnings correlation among neighbors to be 0.007, and the correlation among schoolmates to be 0.01. Based on this evidence it seems plausible to conclude that our finding of negligible community effect is not driven by the specific community definition that we adopt.

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

Our second concern with the estimated community effects is that their estimation is based also on siblings who grew up in different communities, who maybe a selected group in the population due to omitted family characteristics associated with mobility decisions. To address these concerns we therefore re-estimate the model without using information on families where the two brothers attended different schools or grew up in different neighborhoods. Using this strategy we obtain slightly larger community effects (the earnings correlations were 0.03 and 0.01 on average for schoolmates and neighbors), but these are still much lower than the ones we obtain ignoring family effect altogether, and it is safe to say that results in the main model are not driven by those families whose children were exposed to different communities when they were aged 15.

As a last robustness check in this Section, we re-estimate the sibling correlation on hourly wages rather than on annual earnings. Our concern is that the steep decline of sibling correlations that we observed over the initial part of the life-cycle might be driven by differences in the labor market participation of brothers between families. If this was the case, the convergence of life-cycle earnings profiles would be the symptom of changes in labor market participation rather than the effect of heterogeneous human capital investments as we argued when motivating our empirical model. We find that the life-cycle profile of the sibling correlation on hourly wages is rather similar to the one estimated on annual earnings: the correlation is relatively pronounced at the beginning of the life-cycle (0.43 at age 24), then declines towards the early 30s (to a minimum of 0.13 at age 31) and increases again afterwards (reaching the level of 0.35 in the late 40s). The average sibling correlation was also similar to the one estimated on annual earnings, 0.23. Only some of the sibling correlation of annual earnings appears to reflect correlation in working time between brothers.

7. Evidence on birth order, class size and urbanicity

There is a vast literature showing that first-born children do better than later-born on a number of outcomes ranging from education to the labor market (see, e.g., Black et al. 2005). The usual explanation for this finding is that families invest more in the human capital of the first-born. For example, Hotz and Pantano (2013) argue that such differential investment is driven by reputational considerations: societies use first-borns to make inference on unobservable family quality, and families respond to this information problem by concentrating resources on them.

In this section we ask whether birth order plays a role in mediating the impacts of communities on long run earnings. One possibility is that second-born children react to differential (unfavorable) treatment of parents by investing in social relationships outside the family more than their elder brothers do. This would mean that, for them, community effects are a more relevant source of income heterogeneity. A similar result would emerge if communities indeed use the first-born as a signal for unobserved family quality, so that for the second-born community effects would be reinforced by reputational effects. While at the school level differential parental investment on the one hand, and second born reactions and reputational effects on the other hand, might offset each other, neighborhood effects would be univocally amplified by either second born reaction to differential parental investments or reputational effects.

We investigate birth order effects by estimating our model separately on the samples of first and second born sons, and excluding singletons. Clearly, such a strategy cannot identify family effects because this requires the joint modeling of brothers' earnings trajectories: we focus therefore on community effects only. Our previous analysis shows that ignoring family effects leads to upwardly biased estimates of community effects and in this section we assume that such bias is invariant with birth order.

Results on birth order effects are reported in the first two graphs of Figure 6. The overall community effect in permanent earnings correlations is approximately the same for the two brothers, on average 0.077 and 0.073 for first and second born, respectively. However, the relative contributions of schools and neighborhoods differ markedly between brothers: while school effects are the predominant factor for first born, earnings correlations for second born receive a greater contribution from neighborhoods. Note that confidence intervals in these graphs are larger than in Figure 5 and differences between brothers do not appear to be statistically significant.

There are two additional sample splits in Figure 6 that are of interest. The first is by class size, depicted in the third and fourth graph. In particular we split observations depending upon whether school enrollment in the grade attended at age 15 was above or below the threshold used in Denmark for splitting classes, which is 24 pupils. Observations for whom enrollment at age 15 was below 24, between 37 and 48, between 61 and 72, and so on where grouped in the "Below threshold" group, whereas remaining cases (with enrollment between 25 and 36, between 49 and 60 and so on) where grouped in the "Above threshold" sample. Individuals in this latter group will have attended schools with a lower average class size compared with individuals in the "Below threshold" sample. The most striking difference between the two cases is the family effect, which appears to be much stronger for children attending small classes (0.33 versus 0.25 for children in large classes on average). Community effects, on the other hand, are very similar in the two groups.

The last dimension of heterogeneity that we take into account is urbanicity. To do this we exploit information on population density (measured in 1976) in the parishes individuals lived in when they were 15. Specifically, we cut the density distribution across parishes at the upper third, and consider urban individuals living in parishes that are above this threshold, and rural all remaining individuals. Urbanicity is relevant also because it is the factor upon which Page and Solon (2003a and 2003b) concentrated in their analysis of siblings and neighbors earnings correlations, finding that it accounts for much of the neighbors' correlation. We also find similar results. In urban areas, the overall community effect is 0.03, to be contrasted with a sibling correlation of 0.33 in this group. In rural areas, the average community effect is actually negative (-0.02).

8. Conclusion

We develop a unified framework which enables disentangling the contribution of the family, the school and the neighborhood in labor earnings over the life-cycle. This is achieved within a model of multi-person earnings dynamics distinguishing permanent from transitory earnings and allowing for heterogeneous earnings growth. The analysis is based on administrative registers from the Danish population which we use to link brothers to school mates and teenage neighbors.

We find that family is by far the most relevant factor that shapes long-term earnings. The contribution of neighborhood and school quality on long-term earnings is overestimated if the family component is ignored, and becomes negligible and not significantly different from zero by age 30. Finally, the importance of family declines over the life-cycle. These results shed light to previous mixed evidence on the relative magnitude of community and family influences and suggest that the family exerts the predominant role in shaping lifetime earnings inequality. This has important policy implications for the design of policies aiming at reducing inequalities in the long-run.

Appendix

The model of Section 4 fully specifies the intertemporal and interpersonal distribution of permanent earnings. This Appendix discusses identification of its parameters, which is achieved by exploiting different types of moment restrictions generated by the model, and, correspondingly, different types of empirical moments.

Staring with permanent earnings, for a given individual moment restrictions for two non-necessarily different age levels a and $\tilde{a}$ are functions of all sources of earnings heterogeneity:

$$E\left(y_{ifsna}, y_{ifsn\tilde{a}}\right) = \quad \text{(A.1)}$$

$$[\sigma^2_{\mu\Phi} + \sigma^2_{\mu\Sigma} + \sigma^2_{\mu\mathrm{N}} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \left(\sigma^2_{\gamma\Phi} + \sigma^2_{\gamma\Sigma} + \sigma^2_{\gamma\mathrm{N}}\right)a\tilde{a}$$

$$+ \left(\sigma_{\mu\gamma\Phi} + \sigma_{\mu\gamma\Sigma} + \sigma_{\mu\gamma\mathrm{N}}\right)(a + \tilde{a}) + \ \sigma^2_{\omega 24,b} + \sigma^2_{\xi b}\min(a, \tilde{a})]\pi_t\pi_{\tilde{t}}\ , b = 1{,}2$$

Interpersonal moment restrictions do not depend on individual heterogeneity. Moment restrictions between siblings (different *i* but same *f*) depend on the family effect. Moreover, depending upon siblings sharing schools and neighbors, moment restrictions will also depend on school and neighbor effects.[12] Therefore, we distinguish among four types of between-sibling moments: sharing both school and neighbors, sharing school only, sharing neighbors only, not sharing school or neighbor. Moment restrictions are as follows:

- Brothers sharing all community effects

$$E\left(y_{ifsna}, y_{i'fsn\tilde{a}}\right) = \quad \text{(A.2)}$$

$$[\sigma^2_{\mu\Phi} + \sigma^2_{\mu\Sigma} + \sigma^2_{\mu\mathrm{N}} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \left(\sigma^2_{\gamma\Phi} + \sigma^2_{\gamma\Sigma} + \sigma^2_{\gamma\mathrm{N}}\right)a\tilde{a}$$

$$+ \left(\sigma^2_{\mu\gamma\Phi} + \sigma^2_{\mu\gamma\Sigma} + \sigma^2_{\mu\gamma\mathrm{N}}\right)(a + \tilde{a})]\ \pi_t\pi_{\tilde{t}}$$

- Brothers sharing only schools

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

$$E\left(y_{ifsna}, y_{i'fsn'\tilde{a}}\right) = \tag{A.3}$$

$$[\sigma^2_{\mu\Phi} + \sigma^2_{\mu\Sigma} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \left(\sigma^2_{\gamma\Phi} + \sigma^2_{\gamma\Sigma}\right)a\tilde{a}$$

$$+ \left(\sigma^2_{\mu\gamma\Phi} + \sigma^2_{\mu\gamma\Sigma}\right)(a + \tilde{a}) +]\ \pi_t\pi_{\tilde{t}}$$

- Brothers sharing only neighborhoods

$$E\left(y_{ifsna}, y_{i'fs'n\tilde{a}}\right) = \tag{A.4}$$

$$[\sigma^2_{\mu\Phi} + \sigma^2_{\mu\mathrm{N}} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \left(\sigma^2_{\gamma\Phi} + \sigma^2_{\gamma\mathrm{N}}\right)a\tilde{a}$$

$$+ \left(\sigma^2_{\mu\gamma\Phi} + \sigma^2_{\mu\gamma\mathrm{N}}\right)(a + \tilde{a}) +]\ \pi_t\pi_{\tilde{t}}$$

- Brothers sharing neither schools nor neighborhoods

$$E\left(y_{ifsna}, y_{i'fs'n'\tilde{a}}\right) = \quad [\sigma^2_{\mu\Phi} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \sigma^2_{\gamma\Phi}a\tilde{a} + \sigma^2_{\delta\Phi}a^2\tilde{a}^2 + \tag{A.5}$$

$$\sigma^2_{\mu\gamma\Phi}(a + \tilde{a}) + \sigma^2_{\mu\delta\Phi}(a^2 + \tilde{a}^2) + \sigma^2_{\gamma\delta\Phi}(a^2\tilde{a} + a\tilde{a}^2)]\ \pi_t\pi_{\tilde{t}}$$

The above moment conditions are sufficient for identifying family, school, and neighborhoods effects. In particular, identification of school and neighbor effects is ensured by the presence of siblings that went to different schools or grew up in different neighborhoods. However, in order to avoid relying exclusively on these specific groups of siblings for the identification of school and neighborhoods effects, we exploit population data to recover inter-personal moment restrictions linking the two brothers to their peers in schools and neighborhoods at age 15. There are three relevant sets of peers: those attending the same school and living in the same neighborhood; those attending the same school but not living in the same neighbor; and those living in the same neighbor but attending different schools. Empirically, we estimate intertemporal earnings covariances between all individuals belonging to a given cluster of peers using the weighting scheme proposed by Page and Solon

(2003, pp. 841). First we estimate within-cluster covariances; next we take the between-clusters weighted average of within-cluster covariances using weights that are proportional to the density of the cluster. Moment restrictions are as follows:

- Peers sharing all community effects:

$$E\left(y_{ifsna}, y_{i'f'sn\tilde{a}}\right) = [\sigma^2_{\mu\Sigma} + \sigma^2_{\mu\mathrm{N}} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \left(\sigma^2_{\gamma\Sigma} + \sigma^2_{\gamma\mathrm{N}}\right)a\tilde{a} + \left(\sigma_{\mu\gamma\Sigma} + \sigma_{\mu\gamma\mathrm{N}}\right)(a+\tilde{a})]\ \pi_t\pi_{\tilde{t}} \tag{A.6}$$

- Peers sharing only schools:

$$E\left(y_{ifsna}, y_{i'f'sn'\tilde{a}}\right) = [\sigma^2_{\mu\Sigma} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}}\sigma^2_{\gamma\Sigma}a\tilde{a} + \sigma_{\mu\gamma\Sigma}(a+\tilde{a})]\ \pi_t\pi_{\tilde{t}} \tag{A.7}$$

- Peers sharing only neighborhoods:

$$E\left(y_{ifsna}, y_{i'f's'n\tilde{a}}\right) = [\sigma^2_{\mu\mathrm{N}} + 2\sigma_{\mu\Phi\Sigma} + 2\sigma_{\mu\Phi\mathrm{N}} + 2\sigma_{\mu\Sigma\mathrm{N}} + \sigma^2_{\gamma\mathrm{N}}a\tilde{a} + \sigma_{\mu\gamma\mathrm{N}}(a+\tilde{a})]\ \pi_t\pi_{\tilde{t}} \tag{A.8}$$

Using parameter estimates we can decompose the total sibling correlation of permanent earnings into its family and community components, which is shown in the graphs of Section 6. In particular, consider two brothers who attended the same school and lived in the same neighborhood when they were aged 15. The correlation coefficient of their earnings is:

$$\rho^B(a) = \rho^F(a) + \rho^S(a) + \rho^N(a) \tag{A.9}$$

where

$$\rho^F(a) = \frac{E\left(y_{ifsna}, y_{i'fs'n'a}\right)}{E\left(y_{ifsna}, y_{ifsna}\right)}; \tag{A.10}$$

$$\rho^S(a) = \frac{E\left(y_{ifsna}, y_{i'f'sn'a}\right)}{E\left(y_{ifsna}, y_{ifsna}\right)};$$

$$\rho^N(a) = \frac{E\left(y_{ifsna}, y_{i'f's'na}\right)}{E\left(y_{ifsna}, y_{ifsna}\right)}$$

Within person moment restrictions for the member-specific AR(1) model are as follows:

$$E\left(v_{ifsna}v_{ifsn\tilde{a}}\right) = [I(a = \tilde{a} = 24)\sigma^2_{24b} + I(a = \tilde{a} > 24)\left(\exp\left(g_b(a)\right) + var\left(u_{ifsn(a-1)}\right)\rho_b^2\right) + I(a \neq \tilde{a})\left(E\left(u_{ifsn(a-1)}u_{ifsn\tilde{a}}\right)\rho_b\right)]\eta_t\eta_{\tilde{t}}. \quad \text{(A.11)}$$

where $I(\quad)$ denotes indicator functions.

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

$$E\left(v_{ifsna}v_{i'f\tilde{s}\tilde{n}\tilde{a}}\right) = \sigma_f\left(\frac{\left(1-\left(\rho_1\rho_2^{|a-\tilde{a}|}\right)^P\right)}{1-\rho_1\rho_2^{|a-\tilde{a}|}}\right)^{I(a\leq\tilde{a})}\left(\frac{\left(1-\left(\rho_2\rho_1^{|a-\tilde{a}|}\right)^P\right)}{1-\rho_2\rho_1^{|a-\tilde{a}|}}\right)^{I(a>\tilde{a})}\eta_t\eta_{\tilde{t}}. \quad \text{(A.12)}$$

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

We also model the correlation of transitory earnings across peers. Differently from the brothers, in this case we do not model the correlation of AR(1) innovations, 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" factors absorbing all the parameters of the underlying stochastic process:

$$E\left(v_{ifsna}, v_{i'f'sn\tilde{a}}\right) = \lambda_{sn}^{1+|a-\tilde{a}|}\eta_t\eta_{\tilde{t}} \quad \text{(A.13)}$$

$$E\left(v_{ifsna}, v_{i'f's\tilde{n}\tilde{a}}\right) = \lambda_{s}^{1+|a-\tilde{a}|}\eta_t\eta_{\tilde{t}}$$

$$E\left(v_{ifsna}, v_{i'f'\tilde{s}n\tilde{a}}\right) = \lambda_{n}^{1+|a-\tilde{a}|}\eta_t\eta_{\tilde{t}}$$

The moments restrictions above characterize the inter-temporal distribution of permanent and transitory earnings for each individual and between siblings and peers. In general, they are a non-linear function of a parameter vector θ. We estimate θ by Minimum

Distance (see Chamberlain, 1984; Haider, 2001). We use Equally Weighted Minimum Distance (EWMD) and a robust variance estimator $\mathrm{Var}(\theta)=(G'G)^{-1}G'VG(G'G)^{-1}$, where V is the fourth moments matrix and G is the gradient matrix evaluated at the solution of the minimization problem.

References

Aaronson, Daniel. 1998. "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33(4): 915-46.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Hotz, Joseph V., and Juan Pantano. 2013. “Strategic Parenting, Birth Order and School Performance”, IZA Discussion Paper No. 7680.

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

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

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

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.

Page, Marianne, and Gary Solon. 2003. “Correlations between Sisters and Neighbouring Girls in their Subsequent Income as Adults.” *Journal of Applied Econometrics* 18 (5): 545-62.

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

Solon, Gary, Marriange Page, and Greg Duncan. 2000. “Correlations between Neighboring Children in their Subsequent Educational Attainment.” *Review of Economics and Statistics* 82: 383-92.

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

Figure 1: Sibling covariances of annual earnings

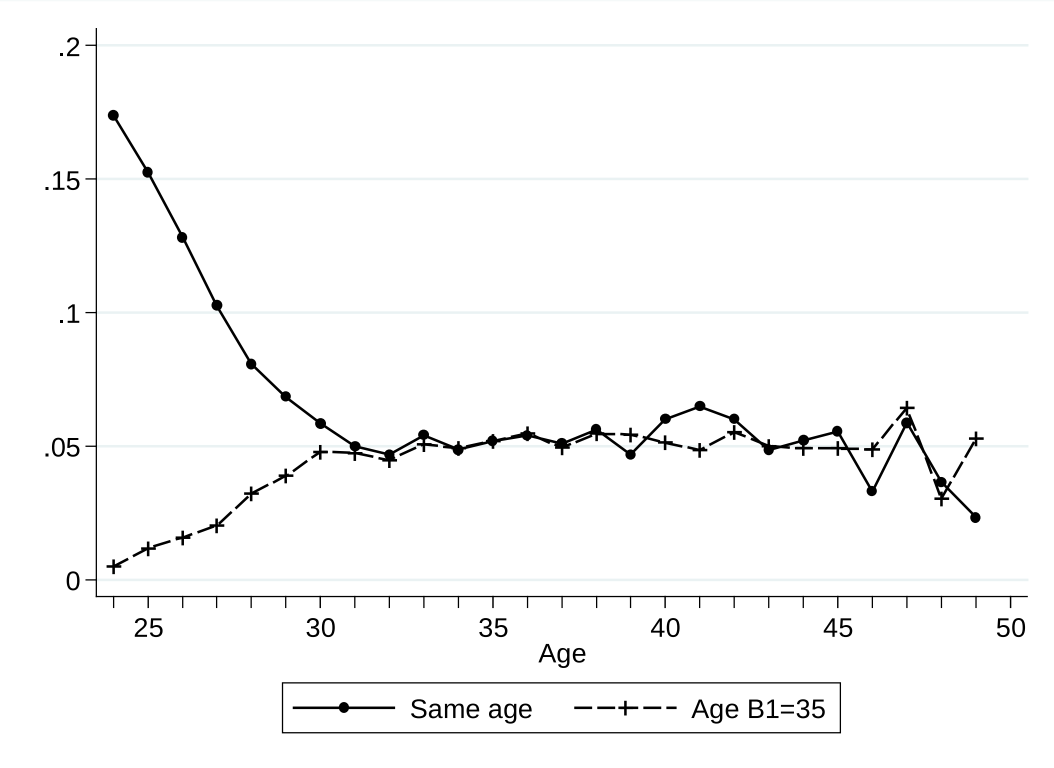

Figure 2: Sibling covariances of annual earnings – Siblings born years apart

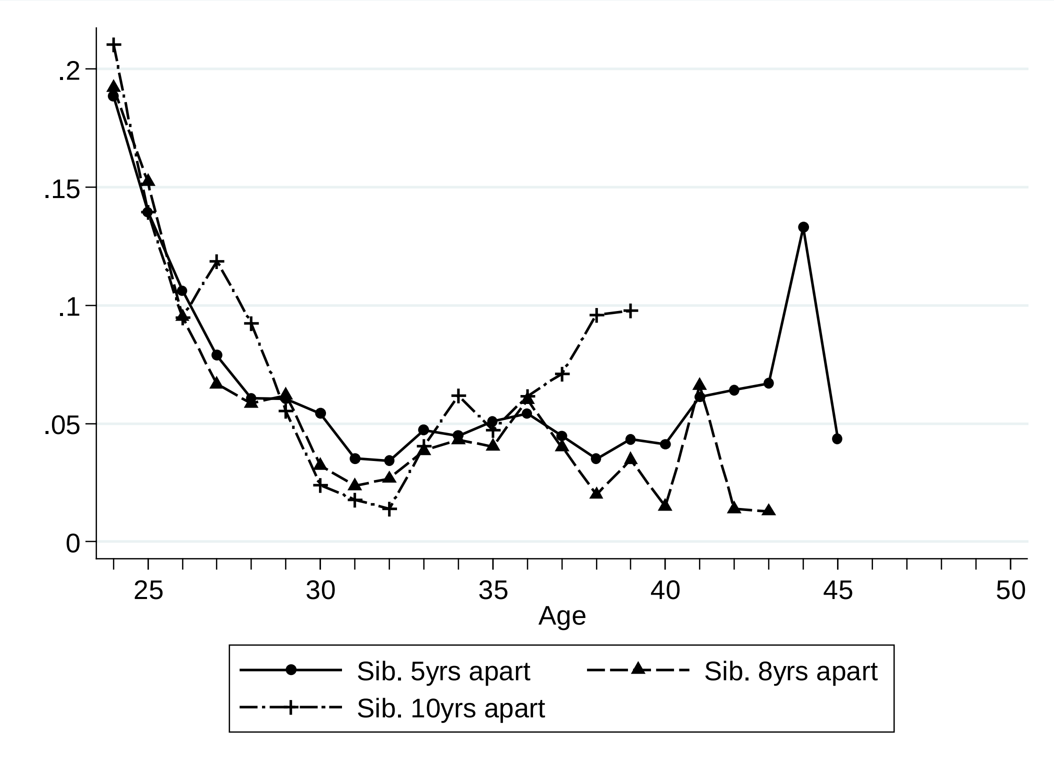

Figure 3: Covariances of annual earnings for members of youth communities

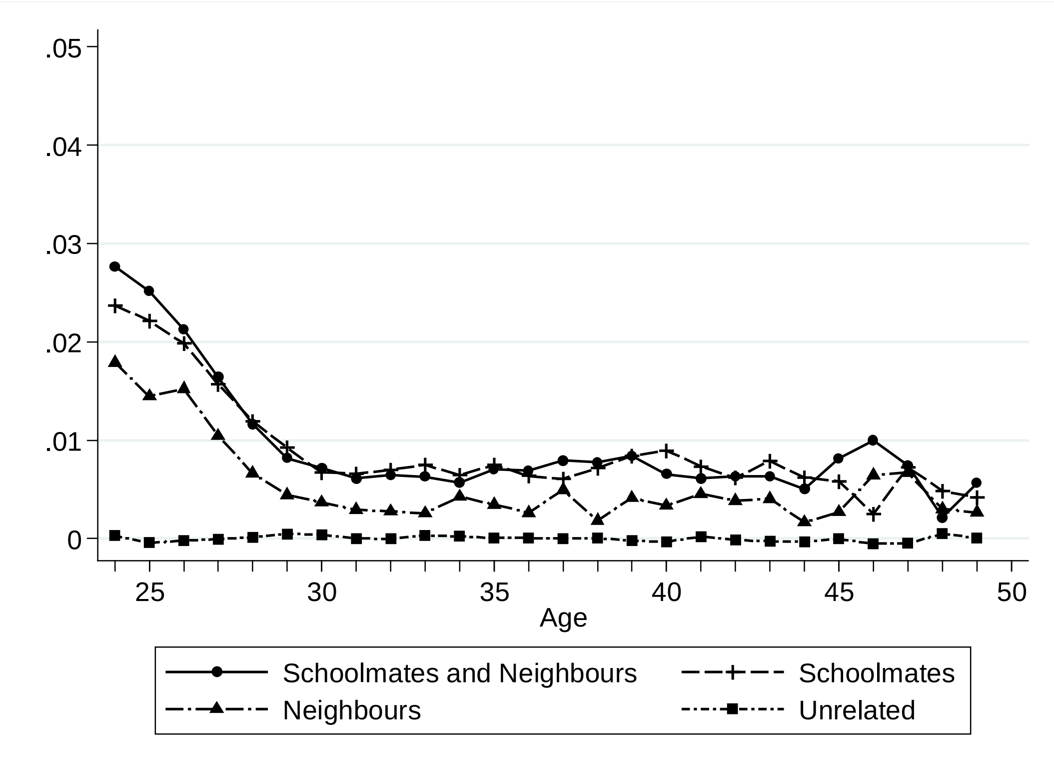

Figure 4: Predicted correlations of permanent earnings

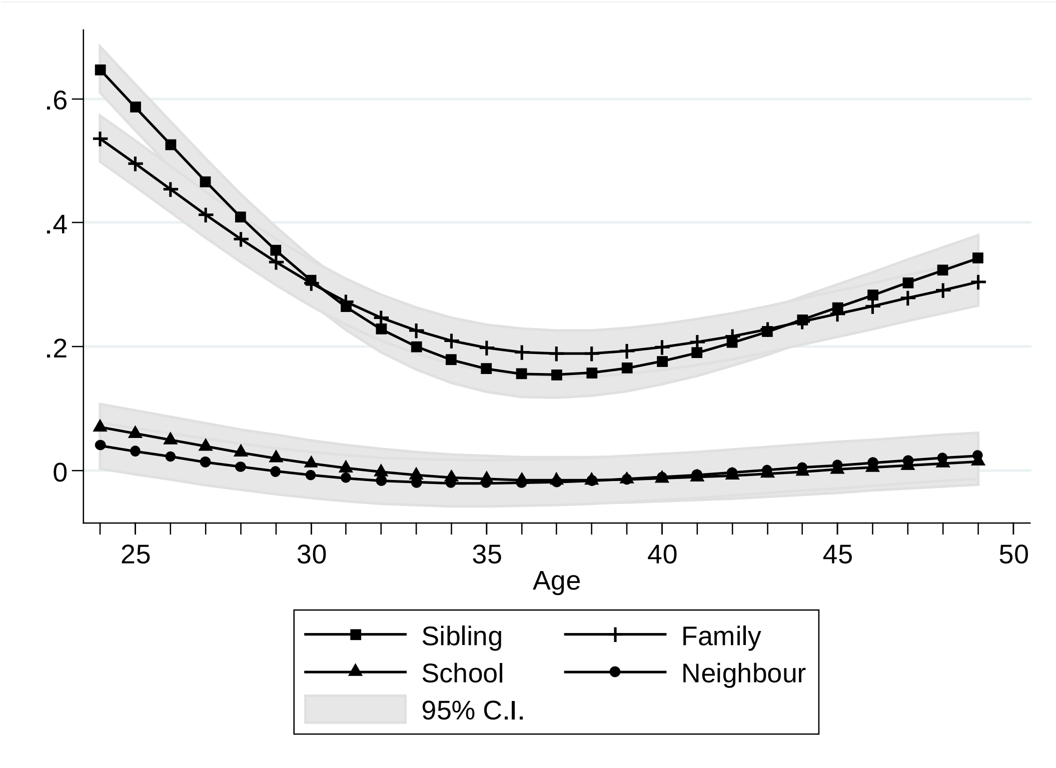

Figure 5: Predicted correlations of permanent earnings between members of youth communities Comparison of models with and without family effects

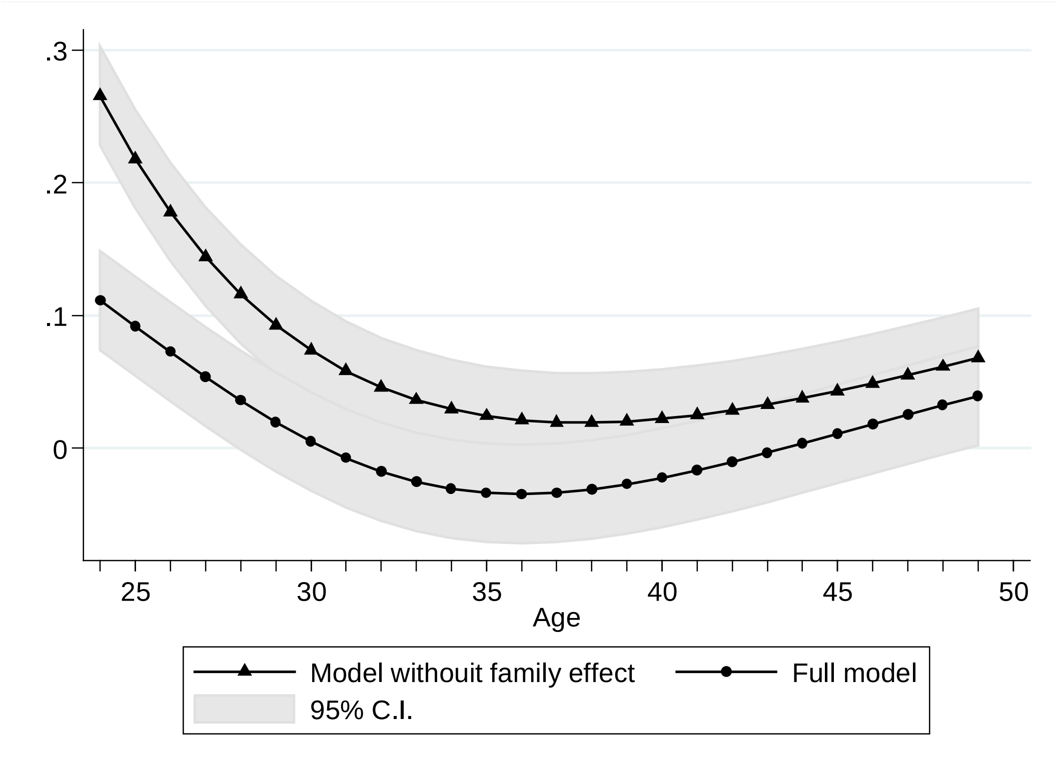

Figure 6: Predicted correlations of permanent earnings by birth order, class size and urbanicity

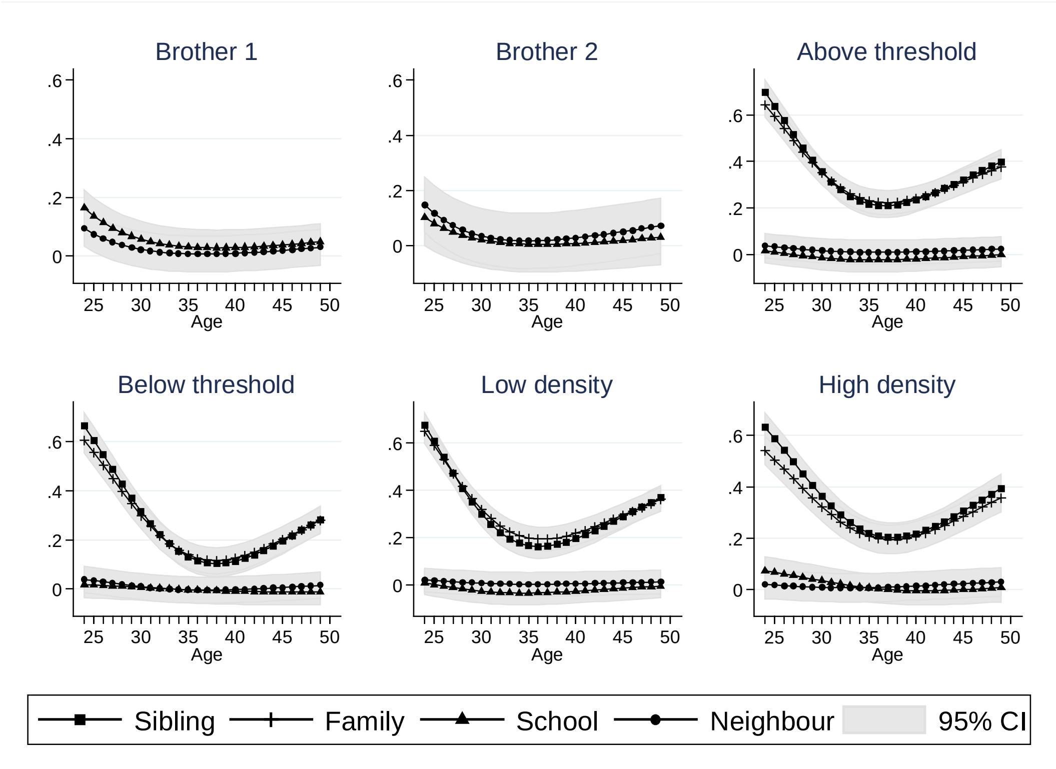

Table 1: Number and density of peer clusters, selected cohorts

Cohort	# clusters	Individuals per cluster			
		mean	*median*	*SD*	*max*
	Sharing N and S				
1960	5229	14.37	13	10.87	59
1965	6064	14.35	12	10.87	55
1970	6217	11.99	10	9.36	49
1975	6413	12.6	11	9.57	43
1980	5888	9.91	8	7.88	38
1983	5624	8.41	7	6.73	42
	Sharing only S				
1960	1202	28.08	27	11.9	90
1965	1459	27.24	26	11.77	107
1970	1505	23.2	22	10.63	96
1975	1533	24.7	24	11.84	111
1980	1487	20.47	19	10.36	90
1983	1500	17.53	17	8.67	58
	Sharing only N				
1960	1849	34.82	29	26.16	114
1965	1994	38.12	29	31.24	142
1970	1985	33.16	27	26.57	134
1975	1974	34.43	29	26.51	134
1980	1934	26.71	22	21.47	124
1983	1922	22.71	19	17.26	91

Table 2: Parameter estimates of permanent earnings

a) Shared components (HIP-RG)

	coeff	se
	Variance of intercepts	
Family ($\sigma^2_{\mu\Phi}$)	0.0524	0.0144
School ($\sigma^2_{\mu\Sigma}$)	0.0028	0.0030
Neighbor ($\sigma^2_{\mu\mathrm{N}}$)	0.0008	0.0033
	Variance of slopes	
Family ($\sigma^2_{\gamma\Phi}$)	0.0003	0.0001
School ($\sigma^2_{\gamma\Sigma}$)	0.00005	0.00002
Neighbor ($\sigma^2_{\gamma\mathrm{N}}$)	0.00005	0.00002
	Covariance intercepts-slopes	
Family ($\sigma_{\mu\gamma\Phi}$)	-0.0032	0.0009
School ($\sigma_{\mu\gamma\Sigma}$)	-0.0007	0.0003
Neighbor ($\sigma_{\mu\gamma\mathrm{N}}$)	-0.0006	0.0003
	Covariance between components	
Family-School ($\sigma_{\mu\Phi\Sigma}$)	0.0037	0.0015
Family-Neighbor ($\sigma_{\mu\Phi\mathrm{N}}$)	0.0025	0.0015
School- Neighbor ($\sigma_{\mu\Sigma\mathrm{N}}$)	0.0011	0.0003

b) Idiosyncratic components (RIP-RG)

	coeff	se
	Initial condition (age 24)	
Brother 1 ($\sigma^2_{\omega 24,1}$)	0.0465	0.0123
Brother2 ($\sigma^2_{\omega 24,2}$)	0.0309	0.0086
	Variance of innovations	
Brother 1 ($\sigma^2_{\xi 1}$)	0.0047	0.0013
Brother 2 ($\sigma^2_{\xi 2}$)	0.0056	0.0016

Table 3: Parameter estimates of transitory earnings

	coeff	se
Initial condition (age 24)		
Brother 1 ($\sigma^2_{24,1}$)	0.6196	0.0407
Brother 2 ($\sigma^2_{24,2}$)	0.6031	0.0406
Variance of innovations at 25		
Brother 1 ($\sigma^2_{\varepsilon 1}$)	0.4587	0.0307
Brother 2 ($\sigma^2_{\varepsilon 2}$)	0.4437	0.0305
Age splines in variance of innovations		
Brother 1		
26-28	-0.1374	0.0033
29-33	-0.1017	0.0027
34-38	-0.0208	0.0036
39-43	-0.0313	0.0057
44-51	-0.0333	0.0087
Brother 2		
26-28	-0.1497	0.0066
29-33	-0.1197	0.0066
34-38	-0.0340	0.0088
39-43	-0.0125	0.0149
44-51	0.0082	0.0294
Autoregressive coefficient		
Brother 1 (ρ_1)	0.5036	0.0036
Brother 2 (ρ_2)	0.5133	0.0051
Cross-person associations in transitory earnings		
Sibling covariance of innovations (σ_f)	0.0058	0.0012
Peers covariance of transitory earnings (catch-all components)		
Sharing both school and neighbor (λ_{sn})	-0.0002	0.0007
Sharing only school (λ_s)	0.0012	0.0006
Sharing only neighbor (λ_n)	0.0006	0.0007

Table 4: Parameter estimates of time effects (1984=1)

	Permanent component (π_t)		Transitory component (η_t)	
	coeff	se	coeff	se
t=				
1985	0.9241	0.1169	0.9544	0.0307
1986	0.9117	0.1308	0.9777	0.0361
1987	1.0808	0.1531	0.9394	0.0384
1988	1.0921	0.1538	1.0147	0.0372
1989	1.0956	0.1588	1.0668	0.0400
1990	1.2064	0.1683	1.0968	0.0394
1991	1.3695	0.1911	1.0610	0.0399
1992	1.2301	0.1703	1.1585	0.0408
1993	1.2520	0.1764	1.1798	0.0424
1994	1.2868	0.1789	1.1750	0.0416
1995	1.2835	0.1777	1.0857	0.0391
1996	1.3206	0.1818	1.0978	0.0388
1997	1.2768	0.1749	1.0815	0.0383
1998	1.2905	0.1780	1.0819	0.0384
1999	1.2737	0.1764	1.1086	0.0394
2000	1.3118	0.1809	1.1342	0.0399
2001	1.2684	0.1752	1.1557	0.0407
2002	1.3392	0.1853	1.1885	0.0418
2003	1.3512	0.1874	1.2596	0.0444
2004	1.3384	0.1863	1.2179	0.0430
2005	1.2693	0.1766	1.2072	0.0427
2006	1.2054	0.1678	1.1740	0.0413
2007	1.1053	0.1539	1.1599	0.0406
2008	1.0391	0.1453	1.1917	0.0419
2009	1.0339	0.1447	1.3624	0.0475
2010	1.0353	0.1451	1.4121	0.0494
2011	1.0212	0.1427	1.4126	0.0493