DISCUSSION PAPER SERIES

IZA DP No. 2608

Intergenerational Education Transmission: Neighborhood Quality and/or Parents' Involvement?

Eleonora Patacchini
Yves Zenou

February 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Intergenerational Education Transmission: Neighborhood Quality and/or Parents' Involvement?

Eleonora Patacchini
University of Rome "La Sapienza"

Yves Zenou
Research Institute of Industrial Economics, GAINS, CEPR and IZA

Discussion Paper No. 2608
February 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 2608
February 2007

ABSTRACT

Intergenerational Education Transmission: Neighborhood Quality and/or Parents' Involvement?*

We develop a model that analyzes the impact of residential neighborhood and parents' involvement in education on children's educational attainment and test it using the UK National Child Development Study. We find that the better the quality of the neighborhood, the higher the parents' involvement in children's education, indicating cultural complementarity. For high-educated parents, the child's educational attainment is more affected by the parents' involvement than by the neighborhood quality while, for low-educated parents, the neighborhood quality seems to play the major role.

JEL Classification: I21, J13, J24

Keywords: education, cultural transmission, cultural substitution, peer effects

Corresponding author:

Yves Zenou
Research Institute of Industrial Economics
Box 55665
102 15 Stockholm
Sweden
E-mail: yvesz@riie.se

* Yves Zenou thanks the Marianne and Marcus Wallenberg Foundation for financial support.

1 Introduction

Explaining children educational outcomes is one of the most challenging questions faced by economists. Most studies have found that school quality (e.g., Card and Krueger, 1992, 1996, and Hanushek, 2002) and family background (e.g., Ermisch and Francesconi, 2001, Sacerdote, 2002, Plug and Vijverberg, 2003) have a significant and positive impact on the level of education of children.[1] However, the effect of neighborhood quality seems to be less clear (see e.g., Durlauf, 2004). The empirical literature on neighborhood effects is far from having reached unanimous agreement that neighborhood effects are important. For example, Solon et al. (2000), Oreopolous (2003), and the papers using the Moving to Opportunity (MTO) programs (like e.g. Katz et al., 2001) find little evidence of these effects. The general consensus seems to be that the neighborhood where individuals grow up matters, although the effects are not large after controlling for individual and family characteristics and parental selection of residential neighborhood.

In the present paper, we are interested in factors that are important for the education transmission and that are affected by the quality of the neighborhood where children live. We focus on one of them, namely parents' involvement in their children's education and investigate how it is related to neighborhood quality. In our framework, neighborhood quality influences parents' involvement in education and this, in turn, affects children' education attainment.

To be more precise, we develop a theoretical framework in which parents' involvement in education as well as the neighborhood where children live are the key ingredients.[2] Indeed, based on some works on anthropology and sociology (see in particular Boyd and Richerson, 1985, Cavalli-Sforza and Feldman, 1981), there is a recent literature initiated by Bisin and Verdier (2000, 2001) arguing that the transmission of a particular trait (religion, ethnicity, social status, etc.) is the outcome of a socialization inside and outside the family (like e.g. peers and role models). These two types of socialization are cultural substitutes (complements) if parents have less (more) incentive to socialize their children the more widely dominant are their values in the population.

We use this idea to explain children's educational attainment. Altruistic parents, who can

[1] See also the literature survey by Haveman and Wolfe (1995) who compare the sociological and economic approaches.

[2] There are in fact theoretical papers that analyze either the effects of parents' input (see e.g. Becker and Tomes, 1979, Leibowitz, 1974) or neighborhood's quality (see e.g. De Bartolome, 1990, Benabou, 1993) on children's educational attainement. Our model links these two approaches and, as a result, gives a mechanism through which both effects affect children's outcomes.

either be educated or not, have to decide how much time they spend educating their children. This is costly since parents have to give up leisure, but also rewarding since it positively influences the chance for their children to be educated. Contrary to the cultural transmission literature cited above where each parent wants his/her children to be like him/her, here only educated parents conform to this behavior since uneducated parents spend time with their children trying to help them becoming different, that is educated. This is because education is not, like for example religion or ethnicity, a trait that is horizontally differentiated (so that it is a matter of taste that one trait is considered better than another) but a trait, or more exactly a characteristic, that it is vertically differentiated (so that everybody agrees that more education is better than less). As in the cultural transmission model, children can become educated either because parents have been successful in educating them (socialization inside the family) or, if this is not the case, because the neighborhood where they live is of sufficiently high-quality in terms of human capital (socialization outside the family). Based on our descriptive evidence, we assume that low-educated parents prioritize less education than high-educated parents. In this respect, there are two aspects of education: the time spent with children and the quality of this time. Given the quality of the neighborhood, the crucial decision for each parent is how much time to spend educating their children.

We test the predictions of the model by merging data on parents and children from the UK National Child Development Study (NCDS) with data on neighborhood characteristics from the 1971 UK Census (when children were 13 years old). Because of their longitudinal aspect (cohort's members are followed from age 7 to 33), the NCDS data are ideal for our purpose. Importantly, by monitoring any household's residential change prior to the child's birth and during the child's childhood and early teenage, this dataset allows us to deal with the problem of an endogenous sorting of families into neighborhoods.

Our empirical analysis points to the following qualitative evidence.

First, we find a significant and positive effect of neighborhood quality on the parents' effort in their children's education, suggesting cultural complementarity.[3] For high-educated parents, we find that both parents' involvement in education and neighborhood's quality significantly affect the intergenerational transmission of education, the former being more potent than the latter. On the other hand, we show that low-educated parents do not spend much time educating their offsprings and only the quality of the neighborhood has a sizeable

[3]To the best of our knowledge, few papers have tested whether cultural substitution or cultural complementarity prevails. A rare exception is Bisin et al. (2004) for the transmission of religion and, contrary to this paper, they find cultural substitution. This is quite intuitive since they are dealing with religion, and thus the more isolated is a religion in an area, the higher parents'effort in transmitting it. For education, we find the contrary because education and the quality of the neighborhood are more complement in nature.

impact on children's educational attainment.

These empirical evidence for the UK corroborate the results obtained for the US. As stated above, the latter indicate a limited role for neighborhood factors in accounting for inequality in educational attainment (Solon et al., 2000, Durlauf, 2004), which may be interpreted as the fact that family matters more than neighborhood. In the present paper, we go further by differentiating between high- and low-educated families. Family seems to play the most important role for children of high-educated parents while it is the neighborhood that shows the most (direct) influential effect on children's educational attainment for low-skilled parents.

The remainder of the paper is organized as follows. In the next section, we survey the literature on the impact of neighborhood effects on the intergenerational transmission of human capital. The theoretical model and its main predictions are exposed in section 3. Section 4 deals with the empirical model and estimation issues while section 5 is devoted to the description of the data and the definition of the variables. Section 6 presents the estimation results. Finally, section 7 concludes.

2 Neighborhood effects in the intergenerational transmission of human capital

There are different papers that have studied the relationship between neighborhood and the intergenerational transmission of human capital. We review here three important contributions to this literature. Kremer (1997) proposes the following equation:

$$H_{it+1} = a_0 + \frac{\alpha}{2}\left(h_{it} + h'_{it}\right) + \beta h_{\nu(i)t} + \varepsilon_i \tag{1}$$

where H_{it+1} denotes the human capital of a member of the ith dynasty in generation $t+1$ (i.e. the child), h_{it} and h'_{it} are the human capital levels of members of the same dynasty in generation t (i.e. the parents) with i designing the father and i' the mother, and $h_{\nu(i)t}$ is the human capital level of the neighborhood $\nu(i)$ where the individual lives at time t. If n_i is the size of this neighborhood, then $h_{\nu(i)t} = \frac{1}{n_i}\sum_{j\in\nu(i)} h_{jt}$. As usual, ε_i is a stochastic shock. Estimating equation (1) using the Panel Study of Income Dynamics (PSID), Kremer (1997) found that the estimate of β, the neighborhood effect, is large when it is compared to the effect of parents' education (α).

Borjas (1992) explores in more details equation (1) by focussing on the ethnic group g. In his model, parents value their own consumption as well as the human capital of their

offsprings. The crucial equation is however the one that relates the human capital obtained by a child belonging to ethnic group g, H_{igt+1}, as a function of that of his/her parents, h_{igt} (which could be his/her father or mother), and of the average human capital of his/her ethnic group, $\overline{h}_{gt} = \frac{1}{n_g}\sum_{j\in g} h_{jt}$, where n_g is the size of the ethnic group. Formally,

$$H_{igt+1} = \gamma_1 h_{igt} + \gamma_2 \overline{h}_{gt} + \xi_{igt} \tag{2}$$

Of course, the variables "neighborhood" and "ethnicity" are highly correlated since minorities tend to live together. The main findings of Borjas (1992) is to show that, using the General Social Surveys and the National Longitudinal Surveys of Youth, ethnic capital, as measured by h_{igt}, the average human capital level of the ethnic group in the parents' generation, plays a crucial role in intergenerational mobility, and slows down the convergence in the average skills of ethnic groups across generations.

Finally, Ioannides (2002, 2003) deepens the analysis of intergenerational transmission of human capital by explicitly developing a dynamic model of human capital formation with a neighborhood selection and estimating it using the Panel Study of Income Dynamics (PSID). He generalizes the two previous papers (Kremer and Borjas) by focussing on non-linear dynamic models. Following Borjas (1992), Ioannides assumes that parents value their consumption and the human capital of their children. The key human capital transmission equation to be estimated is nonlinear and given by:

$$\ln H_{it+1} = a_0 + a_p \ln\left[D h_{it}^{1-1/\phi} + (1-D)\left(h'_{it}\right)^{1-1/\phi}\right] + \ln\left(\sum_{j\in\nu(i)} \mu_{\nu(i)j} h_{jt}^{1-1/\psi}\right)^{\frac{\psi}{\psi-1}} \tag{3}$$

where $\mu_{\nu(i)j}$ denotes the frequency of the value of h_{jt} within the distribution of educational attainment of population in neighborhood $\nu(i)$ at time t, and $\mu(.)$ is the entire distribution. The idea here is to study the impact of parental education and of the distribution of educational attainment within a relevant neighborhood on the child educational attainment. From a theoretical viewpoint, Ioannides obtains a complete characterization of the properties of the intertemporal evolution of human capital. From an empirical viewpoint, he finds that there are strong neighboring effects in the transmission of human capital and that parents' education and neighbors' education have nonlinear effects that are consistent with the theory.

As we will see below, our model is different, since we focus on the interaction between cultural transmission and neighboring effects and their impact on the intergenerational transmission of human capital. To be more precise, the key distinguishing features of our model is twofold: (i) the suggestion of both direct and indirect mechanisms through which neighborhoods might affect human capital accumulation; (ii) the differences across households

of different education levels. This is why we do not test the same equations as the ones mentioned above but rather a model (see equations (14)-(15) below) where: (i) we allow for an interaction between the direct and indirect mechanisms of transmission of education; (ii) different levels of education for both the parents and the children are explicitly taken into account.

3 Theoretical model

In this section, we analyze the intergenerational transmission of education. The key question we would like to study is how much parents are influenced by the local environment when deciding the level of effort they put in educating their children. As in Bisin and Verdier (2000, 2001), the transmission of education is modeled as a mechanism that interacts socialization *inside* the family (*vertical* socialization) with socialization *outside* the family (*oblique* socialization) via imitation and learning from peers and role models.

There are two types of parents/workers: high-educated, $i = h$, and low-educated parents, $i = l$. There is a continuum of each of them. The instantaneous utility of a parent of type $i = h, l$ is given by:[4]

$$z^i + U(\lambda^i, e^i)$$

where z^i is the quantity of a consumption good (taken as the numeraire) consumed by parents, λ^i is the time spent on leisure and e^i is time (effort) they spend with their children trying to educate them. $U(.)$ is assumed to be increasing in λ^i and decreasing in effort e^i, and concave in both arguments. This choice of the utility function aims at capturing the fact that the time spent with children and on leisure are not independent activities for parents.

The budget constraint of a parent $i = h, l$ can be written as follows:

$$w^i T = z^i \tag{4}$$

where w^i is the per-hour wage (with $w^h > w^l$) and T denotes the amount of working hours. T is assumed to be the same and constant across workers, an assumption that agrees with most jobs in the vast majority of developed countries.[5]

Each parent provides a fixed amount of labor time T and spends some time on leisure and with the children. Thus, the time constraint of a parent $i = h, l$ can be written as:

$$1 - T = \lambda^i + e^i \tag{5}$$

[4] None of our results is affected by the fact that the utility function $U(.)$ is separable.

[5] We could have assumed that more educated parents work more hours than less educated workers. This would not affect any of our results.

where the total amount of time is normalized to 1 without loss of generality.

By plugging (4) and (5) into the utility function, we obtain the following instantaneous indirect utility for parents of type $i = h, l$:

$$V^i(e^i) = z^i + U(\lambda^i, e^i) = w^i T + U\left(1 - T - e^i, e^i\right) \tag{6}$$

Let us now focus on the parent's choice of effort $e^i \in [0, 1]$. In the empirical analysis, Table 2 shows that high-educated parents read more to their children and are more involved in the education of their children (for example by meeting more often any member of the teaching staff). As a result, we assume here that low-educated parents prioritize less education than high-educated parents. Therefore, the probability that their kids will be educated is e and δe, respectively, with $0 \leq \delta \leq 1$. For the educated parent, with a probability equal to the education effort e^h, education will be successful and the child will be like the parent (highly educated). For the uneducated parents, education will be successful with probability δe^l. Thus, here, $\delta = 1$ means that education is a top priority for the parents, so the time they spend with their kids is of high quality, while a δ close to zero implies that parents are not prioritizing education and the interaction with their kids while educating them is not of high quality (for example, they talk to their kids about education while watching television). In this respect, there are two aspects of education: e^i, the time spent with children, and δ^i ($\delta^h = 1$ and $\delta^l = \delta$), the quality of this time.

In both cases, if education is not successful, the child remains without education and gets randomly matched with someone else whose education he/she will adopt. It is at this second stage, after the parents' unsuccessful education, that children are influenced by their peers or teachers (role models).

We denote by π^{ij} the probability that a child of type-i parent ($i \in \{h, l\}$) obtains education $j \in \{h, l\}$. Since there is a continuum of agents, by the Law of Large Numbers, π^{ij} also denotes the fraction of children with a parent i who has education j. Denoting by q the proportion of *high-educated* individuals in the economy (or the neighborhood), we have the following transition probabilities:

$$\pi^{hh} = e^h + (1 - e^h)q \tag{7}$$

$$\pi^{hl} = (1 - e^h)(1 - q) \tag{8}$$

$$\pi^{ll} = (1 - \delta e^l)(1 - q) \tag{9}$$

$$\pi^{lh} = \delta e^l + (1 - \delta e^l)q \tag{10}$$

where $0 \leq \delta \leq 1$. Let us interpret equation (7). The child of a high-educated parent will also be highly educated if either his/her parent's education is successful (probability e^h) or

if the parent fails to transmit his/her trait (probability $1-e^h$) and the child picks up the education trait from the society (probability q). Equation (8) gives the probability that a child of educated parents is not educated: it is because both the parents and the society were unsuccessful in educating the child. For low-educated parents (equations (9) and (10)), we have a similar interpretation, with the difference that parents assign a lower priority to the education of their kids.

We are now able to write the expected utility function of all parents. We assume that all parents (educated or not) are altruist and thus do care of the future job situation of their children. We denote by V^{ij}, $i=h,l$, $j=h,l$, the future utility of a child j whose parent is of type i. Note that this utility is evaluated by the parents and thus take their point of view (imperfect empathy). The simplest interpretation of these utilities is in terms of the child's future income, given that $w^h > w^l$. In other words, all parents (educated or not) will be better off if their children achieve high education and thus make more money. For simplicity and without loss of generality, we have:

$$V^{hh} = V^{lh} = w^h$$

$$V^{hl} = V^{ll} = w^l$$

As a result, the expected utility of educated and non-educated parents are respectively given by:[6]

$$\begin{aligned} EV^h &= V^h(e^h) + a\left[\pi^{hh}V^{hh} + \pi^{hl}V^{hl}\right] \\ &= w^hT + U\left(1-T-e^h, e^h\right) + a\,e^h\,(1-q)\left(w^h - w^l\right) + a\left[qw^h + (1-q)w^l\right] \end{aligned}$$

$$\begin{aligned} EV^l &= V^l(e^l) + a\left[\pi^{ll}V^{ll} + \pi^{lh}V^{lh}\right] \\ &= w^lT + U\left(1-T-e^l, e^l\right) + a\,\delta\,e^l\,(1-q)\left(w^h - w^l\right) + a\left[qw^h + (1-q)w^l\right] \end{aligned}$$

where $0 < a < 1$ is the degree of altruism that is common to both educated and uneducated parents. Let us now determine e^i, the effort's choice of parents $i = h,l$. If we use the following notations $U_{\lambda^i} \equiv \frac{\partial U}{\partial \lambda^i}$ and $U_{e^i} \equiv \frac{\partial U}{\partial e^i}$, then the first order conditions for educated and uneducated parents are respectively given by (we only focus on interior solutions):[7]

$$-U_{\lambda^i} + U_{e^i} + a\,\delta^i\,(1-q)\left(w^h - w^l\right) = 0 \tag{11}$$

[6] The altruistic model was made famous by Becker (1974, 1991). For a recent survey on these types of models, see Laferrere and Wolff (2004).

[7] For each parent $i = h,l$, the second order condition is given by:

$$U_{\lambda^i\lambda^i} + U_{e^ie^i} - 2U_{\lambda^ie^i}$$

and is assumed to be negative.

where $\delta^h = 1$ and $\delta^l = \delta$. The solution of (11) is denoted by e^{i*}, which is equal to $e^{h*}(q)$ and $e^{l*}(q, \delta)$ for high- and low-educated parents, respectively. If we further adopt the following notations, $U_{\lambda^i \lambda^i} \equiv \frac{\partial^2 U}{\partial \lambda^i \partial \lambda^i}$ and $U_{\lambda^i e^i} \equiv \frac{\partial^2 U}{\partial \lambda^i \partial e^i}$, we have the following proposition:[8]

Proposition 1

(*i*) *High-educated parents spend more time in educating their offspring than low-educated parents, and this difference increases with δ, the relative inefficiency of success for low-educated parents.*

(*ii*) *Assume either $U_{\lambda e} > 0$ or $U_{\lambda\lambda} < U_{\lambda e} < 0$. Then, for both educated and uneducated parents, the higher the proportion of high-educated people in the area, the lower the effort parents put in educating their children, that is:*

$$\frac{\partial e^{i*}}{\partial q} = \frac{a\,\delta^i \left(w^h - w^l\right)}{U_{\lambda^i \lambda^i} - U_{\lambda^i e^i}} < 0 \tag{12}$$

This is referred to as cultural substitution.

(*iii*) *Assume $U_{\lambda e} < U_{\lambda\lambda} < 0$. Then, for both educated and uneducated parents, the higher the proportion of high-educated people in the area, the higher the effort parents put in educating their children, that is:*

$$\frac{\partial e^{i*}}{\partial q} = \frac{a\,\delta^i \left(w^h - w^l\right)}{U_{\lambda^i \lambda^i} - U_{\lambda^i e^i}} > 0 \tag{13}$$

This is referred to as cultural complementarity.

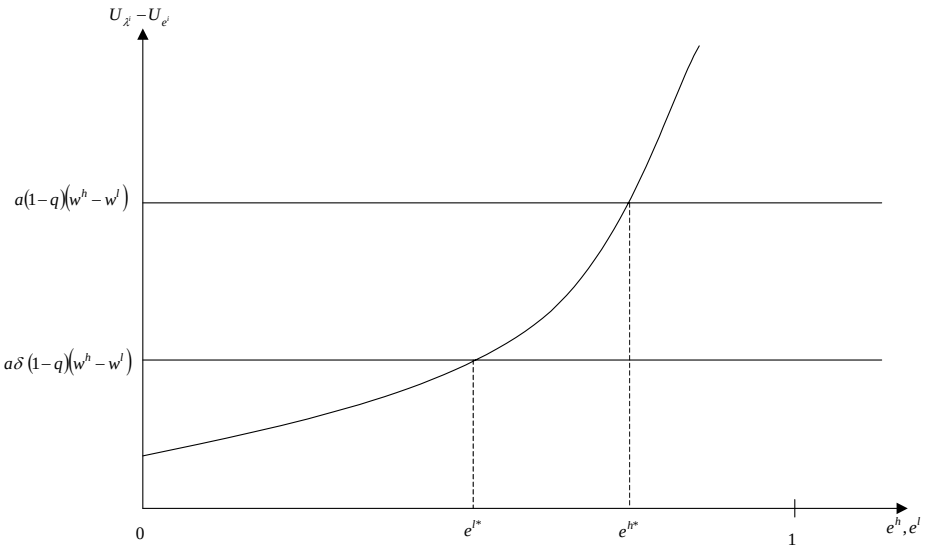
The first order condition (11) shows that the choice of e^* involves a trade off between the short-run costs of spending time with children (the resulting forgone leisure) and the long-run expected benefits, which consist in a better chance of having an educated child with a higher wage. When $U_{\lambda e} > 0$, which means that the higher the effort e, the higher the marginal utility of leisure, the costs are higher and the parents put less effort the higher the level of education in the economy (cultural substitution). On the contrary when $U_{\lambda e} < 0$ and $U_{\lambda e} < U_{\lambda\lambda}$, which means that the loss in the marginal utility of leisure following an increase in e is low, the parents put more effort the higher the level of education in the economy (cultural complementarity). Observe that Proposition 1 implies that for low educated parents, if $\delta \approx 0$

[8] Observe that exactly the same results would have been obtained in Proposition 1 (*ii*) and (*iii*) if we had expressed the assumptions on the technology rather than directly from preferences (see Bisin and Verdier, 2001).

the quality of the neighborhood should not affect parental effort in educating the children ($\frac{\partial e^{i*}}{\partial q} \approx 0$). This results is emphasized in Proposition 3.

Figure 1 helps us to understand result (i) in Proposition 1. Because it is costly to spend time with children, the higher the returns, the higher parents' effort e^i. So, because the returns to the investment e^i is lower for the low-educated parents, they spend less time educating their kids. As discussed above, this may be because low-educated parents have different priorities and thus may not prioritize education as high-educated parents do. Also, the lower δ, the return to low-educated parents' investment, the higher the difference between e^h and e^l.

Figure 1: Effort differences between educated and uneducated parents



We can now calculate the expected school achievement of each individual by focusing on the different transition probabilities.

Proposition 2

(i) *For educated parents whose effort is $e^*(q)$, the probability that their child will be educated is:*

$$\pi^{hh} = e^{h*}(q) + \left[1 - e^{h*}(q)\right] q$$

while the probability that their child will not be educated is:

$$\pi^{hl} = \left[1 - e^{h*}(q)\right] (1 - q)$$

(*ii*) *For low-educated parents, the probabilities that their child will be educated and non-educated are respectively given by:*

$$\pi^{ll} = \left[1 - \delta e^{l*}(q, \delta)\right](1 - q)$$

$$\pi^{lh} = \delta e^{l*}(q, \delta) + \left[1 - \delta e^{l*}(q, \delta)\right] q$$

(*iii*) *For both parents (educated or not), if there is cultural complementarity, a better-quality neighborhood increases the probability to be educated and decreases the probability to be uneducated, that is* $\frac{\partial \pi^{hh}}{\partial q} > 0$, $\frac{\partial \pi^{lh}}{\partial q} > 0$ *and* $\frac{\partial \pi^{hl}}{\partial q} < 0$, $\frac{\partial \pi^{ll}}{\partial q} < 0$. *If there is cultural substitution, all these effects are undetermined. Finally, for a given neighborhood quality* q*, parents' effort always increases the chance for their offspring to be educated, that is* $\frac{\partial \pi^{hh}}{\partial e^{h}} > 0$, $\frac{\partial \pi^{lh}}{\partial e^{l}} > 0$ *and* $\frac{\partial \pi^{hl}}{\partial e^{h}} < 0$, $\frac{\partial \pi^{ll}}{\partial e^{l}} < 0$.

Results (*i*) and (*ii*) just express the transition probabilities (7)-(10) in terms of optimal parents' effort. The interesting result is (*iii*) since it shows the impact of both the quality of the neighborhood and parents' involvement on children's education attainment. Since the education process is in two stages (first the parents' involvement e^i and then the neighborhood's quality q) and since both stages are influenced by q, there are two effects: an *indirect* one, in which e^i hinges on q, and a *direct* one, because if e^i fails, then only q affects children's educational attainment. So, when there is cultural complementarity, these two effects reinforce each other since a higher q implies a higher indirect (the higher the quality of the neighborhood, the higher parents' effort) and direct effects. If, on the contrary, there is cultural substitution, then a better quality neighborhood reduces the chance to be educated by parents (since parents spend less time with their kids) but increases the chance to be educated by peers (since q is higher the chance to meet a high-educated peer is higher). The net effect is thus ambiguous.

We have finally the following result, which is a consequence of the two propositions above:

Proposition 3 *For low-educated parents,*

(*i*) *the less prioritized education is for parents (i.e. the lower* δ*), the lower the time spent with their children* e^{l*}*, i.e.* $\frac{\partial e^{l*}}{\partial \delta} > 0$*;*

(*ii*) *the less prioritized education is for parents (i.e. the lower* δ*), the higher the probability to be uneducated and the lower the probability to be educated, i.e.* $\frac{\partial \pi^{ll}}{\partial \delta} < 0$ *and* $\frac{\partial \pi^{lh}}{\partial \delta} > 0$*;*

(*iii*) *When δ is low enough δ, i.e. $\delta \to 0$, the effort e^{l*} provided by uneducated parents is negligible and thus the quality of the neighborhood q has no impact on e^{l*}. In that case, the probability to be educated or not only depends on the quality of the neighborhood, that is $\pi^{ll} = 1 - q$ and $\pi^{lh} = q$.*

This last proposition focuses on low-educated parents. If education is not a priority at all ($\delta \to 0$), then obviously only the environment where children live (i.e. peers and role models) will affect children's educational attainment.

To summarize, the key feature of this model is that both socialization *inside* the family (the role of parents) and socialization *outside* the family (the role of peers, schools and role models) play an important role in the education process of children. If they live in a "good" environment with educated parents who take care of them, then the chance to reach a high education level is very high. If, on the contrary, they live in a rundown area with low-quality schools and negative peer pressures and if on top of that their parents are not educated and do not spend time with them, then the probability to be educated is quite low for these children.

4 Empirical strategy

4.1 Tests of the model

We would like to test propositions 1, 2 and 3, that is the influence of the local environment on the parents' decision in spending time with their children and the impact of both parents' investment and local environment quality on the education attainment of the children.

Let us begin by considering Proposition 1 (and Proposition 3).

The exact empirical counterpart of e^i is the share of time spent caring for a child's education. It is not reported in our data sets. Thus, we model the underlying parent's propensity in investing in their children education as a linear function of parental, child, household and neighborhood characteristics using as indicator a variable based on qualitative information on the parent's effort in his/her child's education. A probit specification is employed where the dependent variable is equal to one if the parents spend substantial time for their child's education and zero otherwise. From Proposition 1, the following model is considered:

$$e^i_{n,k,t} = \alpha q_{k,t} + \sum_{m=1}^{M} \beta_m x_{m,t} + \varepsilon_{n,t} \qquad n = 1, ..., N \tag{14}$$

where $e^i_{n,k,t}$ is the (unobservable) time spent by the parents of child n of type $i = h, l$ who resides in area $k = 1, ..., K$ at time t for educating the child; $q_{k,t}$ is the average quality in terms of education of area k at time t; $x_{m,t}$ (for $m = 1, ..., M$) is a set of M control variables at the parental, child, household and area level at time t accounting for differences in socio-economic characteristics between parents, children, families and areas; $\varepsilon_{n,t}$ is a white noise error term. A test of this equation will give important information regarding the prediction of the theoretical model. First, for high-educated parents ($i = h$), an α significantly different from zero will indicate either either cultural substitution (if negative) or cultural complementarity (if positive). Also, even if this coefficient is not statistically significant, it will provide information on the form of the trade-off between the neighborhood composition and parental investment in education. For low-educated parents, following Proposition 3, depending of the value of δ (which is unobservable), α could be significant (if δ is high enough) or not significant (if δ is low enough). There is an additional prediction of the model worth emphasizing. Indeed, as δ goes to zero, not only the effort for low educated parents should decline and thus the neighborhood composition on outcomes should be more important, but, by comparing (7) and (10), the effect of similar neighborhood composition q should be higher for low-educated households.

Turning our attention to the test of Proposition 2 (and Proposition 3), we model the likelihood of a successful or unsuccessful intergenerational transmission of education (transition probabilities in Proposition 2) as follows:

$$\pi^{ij}_{n,t+1} = \phi q_{k,t} + \gamma e^i_{n,k,t} + \rho(q_{k,t}.e^i_{n,k,t}) + \sum_{v=1}^{V} \theta_v x_{v,t} + \eta_{n,t+1} \qquad n = 1, ..., N \tag{15}$$

where $\pi^{ij}_{n,t+1}$ is the probability that an adult n at time $t+1$ (who was a child n at t) whose parents are of type $i = h, l$ attains the level of education $j = h, l$; $q_{k,t}$ is the quality of the neighborhood k when the adult was a child at time t; $\eta_{n,t+1}$ is a white noise error term. The set of controls $x_{v,t}$ (for $v = 1, ..., V$) includes some (but not all) of the variables that entered in the set $x_{m,t}$ from equation (14), allowing these variables to have a different impact on $\pi^{ij}_{n,t+1}$ than they had on $e^i_{n,k,t}$. This is of particular interest for our target variable $q_{k,t}$, which has been separated out from the set of control variables for ease of clarity.

The probabilities π^{ij}_{t+1} are analyzed using probit models, each of them having the dependent variable equal to one if the (observed) implied child's educational attainment is achieved and zero otherwise. A successful test of Proposition 2 would imply that for children of high-educated parents, both the effect of $q_{k,t}$ and $e^i_{k,t}$ have to be significant, whereas for children of low-educated parents, both effects are expected to be significant only if δ is

high enough, otherwise only the impact of $q_{k,t}$ should matter. Finally, if there is cultural complementarity, one would expect both the impacts of $q_{k,t}$ and $e^i_{n,k,t}$ to be positive for π^{hh} and π^{lh} (and negative for π^{hl} and π^{ll}) while, with cultural substitution, their signs would be undetermined.

Equations (14)-(15) are estimated jointly by Maximum Likelihood. This strategy accommodates the possibility that some unobservable factors might affect both parental engagement in their child's schooling and the child's ultimate educational attainment.

Because the use of cross-equation restrictions is always rejected by our data, model (14)-(15) is estimated separately for children of high-educated and low-educated parents.

Our bivariate probit model is properly identified if there is at least one variable that is correlated with parental interest (equation (14)) but is uncorrelated with the child's future educational achievement (equation (15)). In our analysis the regressors in equation (14) excluded from equation (15) are two indicators of the mother's behavior during pregnancy, namely whether the mother who declares being a smoker prior to pregnancy quits smoking after four months of pregnancy and the mother's total number of antenatal doctor visits (without abnormality during pregnancy). Table 1 reports the correlations of these variables with our indicator of parental effort in the child's education (i.e. the frequency of reading to the child, see Section 5.2), with other parental background variables and with the child's final education attainment. The evidence is presented for high-educated and low-educated parents separately. It reveals that, irrespective of the parents' education, both variables have no direct effect on the child's education attainment, while being (positively) significantly correlated with our measure of parental effort. Interestingly, in our case, these indicators of the mother's behavior during pregnancy do not appear to be significantly correlated with other parental background variables. Therefore these variables seem to be suitable instruments: they are predictors of parental care that do not affect directly the probability that a child will attain a certain education level (other than through their effect on parental interest in the child's education). Indeed, likelihood ratio tests do not reject the null hypothesis that the instruments either individually or jointly do not have a direct impact on the child's education attainment. Furthermore, these instruments also avoid the possibility of reverse causality, as the mother's attitude towards pregnancy in the antenatal period is clearly not determined by the child's future education attainment.

[*Insert Table* 1 *here*]

4.2 Endogeneity issues

There are two main issues that might hamper our empirical exercise: (*i*) a possible endogenous sorting of families into neighborhood, and (*ii*) the possible presence of (unobserved) inherited characteristics, which affect a child's education attainment. The longitudinal structure and richness of information of our dataset allows us to address both concerns.

The problem related to the first issue is that households with unobservable characteristics that make them highly concerned about their children's education are likely to locate in more educated neighborhoods, causing the average human capital of the neighborhood (q) to be positively correlated with the error term. Because in our dataset individuals are followed over time (for more than thirty years), we can compare the estimation results of our empirical exercise run on the sub-sample of families that changed (at least one) neighborhood during the child's childhood or early teenage (*movers*) with those obtained for the families that did not change residential address (*stayers*). If the impact of the average human capital of the neighborhood on our target variables is *mainly* driven by sorting, i.e. by the fact that more educated parents are more likely to look for more educated communities, we should find a sizeable difference between the estimated coefficients of q for *movers* and *stayers* (with the one for *movers* being upward biased). If the effect is *only* due to sorting, the estimated coefficient of q for *stayers* should not be statistically different from 0. Any statistical significant effect for this group is a signal that selection issues are *not the only* responsible of the uncovered effect. We are also able to control for any residential change previous to child's birth by using a retrospective question about the number of years the household has stayed at a given address. We include in our sub-sample *movers* the families that changed neighborhood up to three years before the child's birth as this move might be related to the plan of having a child. As a consequence, the *stayers* in our analysis are the families that did not change address both after the child's birth and up to three years before child's birth. It is reasonable to assume that residential changes beyond this time span, which are not monitored in our data, are not strongly correlated with the expectation of having a child in a faraway future. Also, because we use the Census 1971 data to describe the child's residential neighborhood (i.e. when the child is at age 13), we consider only residential changes that occurred before the child is 13 years old. Beyond this age, however, any residential change is definitely less related to the households' concerns about the children's education. In other words, in our sample of *movers*, we aim at capturing households who relocate because of parents' concerns about their children's education. These decisions should reasonably take place before the child is 13 years old.

The problem related to the second issue is that cognitive ability is a heritable trait so

that the better schooling performance of children of high-educated (high-ability) parents may be the result of the transmission of genes for high ability rather than of parental investment. We tackle this issue by considering in our analysis only children who have been *adopted* by parents. This exercise should eliminate the impact of any unobservable inherited characteristic.

5 Data and descriptive evidence

5.1 Data

Our empirical analysis is based on data from the National Child Development Study (NCDS). It is a longitudinal survey that follows all British persons who were born between the 3rd and 9th of March 1958, with follow-up surveys in 1965 at age 7 (NCDS sweep one), in 1969 at age 11 (NCDS sweep two), in 1974 at age 16 (NCDS sweep three), in 1981 at age 23 (NCDS sweep four), and in 1991 at age 33 (NCDS sweep five). This dataset is ideal for the purpose of this paper as it contains detailed parental and child information, as well as data on family background, school quality and area of residence identifiers for cohort's members residential addresses. Good family background information is essential when trying to find evidence of neighborhood effects since neighborhood characteristics may proxy for unobservable family characteristics. The information on the residential location allows us to match NCDS data with the 1971 Census data, obtaining a detailed picture of the residential neighborhood community when cohort's members were teenagers (age 13). Census information is taken from the Small Area Statistics (SAS) datasets. In particular, data on education, economic activity and occupation of each area residential community are only available for the 10% sample survey. This implies extremely small sample sizes per area if the most basic census spatial unit, i.e. enumeration district (with an average of 300-400 residents), is used as a neighborhood measure. Therefore, we are forced to choose the next available level of spatial disaggregation: we use *ward* level data, providing 17,500 areas in UK in 1971 with an average of 3,000-4,000 residents. A Census ward contains roughly ten enumeration districts.

Our empirical analysis matches information on individuals' education attainment at age 33 from the NCDS fifth sweep with the information on parental characteristics, quality of the school attended at age 16, ability in the childhood from earlier NCDS sweeps and residential neighborhood information from the 1971 Census (when individuals are 13 years old).

Considering only individuals without missing values in our target variables, our final sample is of 2,723 adopted children (and 5,015 parents) from families that changed neigh-

borhood at least once and of 1,981 adopted children (and 3,868 parents) from families that never changed residential address in the observation period.[9]

5.2 Definition of variables

The key variables in the theoretical model are parent's effort in child's education (e^i), the transition probabilities (π^{ij}) and the average neighborhood human capital (q).

Let us first discuss our empirical proxy for parents' effort in child's education. The ideal variable to measure parent's effort would have been the number of hours spent investing in children's education (reading to the child, meeting teachers, etc...). Unfortunately, this variable is not directly available in the NCDS. However, the NCDS provides rich qualitative information on parental interest in the child's education at different ages of the child (see Table 2). The closest proxy of "effort" is the frequency of reading to the child. It is taken when the child is of age 7. Specifically, in the NCDS sweep one (parental questionnaire) it is asked: "Does the mother/father read to, or read with, the child?", and the possible answers suggested are "Yes, at least every week; Yes occasionally; Never or hardly never; Don't know or inapplicable". We exclude the mothers and fathers who answer "Don't know or inapplicable" and we rank the parents according to the highest frequency declared by each couple.[10] This variable is coded as a dichotomous variable, taking value one if the parents read at least every week to the child and zero otherwise (i.e., if they read to the child only occasionally or never or hardly never).[11]

[*Insert Table 2 here*]

Let us now turn to the empirical counterparts of the other key variables of the theoretical model, that is π^{ij} and q. The NCDS sweep five (the child is now a 33 years old adult in 1991) provides information on the highest qualification obtained by the cohort's members. We define high-educated individuals the ones with A-levels[12] or above qualifications and low-educated individuals otherwise. As a proxy for the average neighborhood education quality

[9] As with most longitudinal surveys, attrition and incomplete information is an issue in the NCDS. In addition, the sample selection requirements for the purpose of this paper are quite stringent. However, comparing descriptive statistics of the whole sample and our selected one, it does not appear that we lose representativeness.

[10] Non-response is not an issue here. Missing values are reported in less than 5 % of the observations.

[11] A different coding of this variable allowing more than two levels that leads to the estimation of ordered probit models does not change qualitatively our results.

[12] The A-level in the UK is equivalent to the SAT in the US or the baccalaureat in France.

in an area k at time t, i.e. $q_{k,t}$, we use the percentage of persons over 18 years old holding a A-level or more in a ward k from the Census 1971 data, i.e. when the child is 13 years old.

Finally, the information on the parents' education is derived from the age the parents left school, which is reported in 1974 (NCDS sweep three). The parents' education is measured using completed years of schooling. Consistently with the aggregation used for cohort's member education, implying that high-educated individuals are those that left school at an age greater than 18 years (and low-skilled otherwise), we define parents of type $i = h$ if the mother and the father's average years of schooling is greater than 12 years and parents of type $i = l$ otherwise. So, for example, $\pi^{hh}_{n,t+1}$ is the probability that an adult n of 33 years old in $t + 1 = 1991$ has at least an A-level degree given that his/her parents have (on average) more than 12 completed years of schooling.

To summarize, we evaluate parents' effort when the child is 7 (in 1965), the quality of the neighborhood and parents' level of education when the child is 13 (in 1971) and the child's education attainment when he/she is 33 years old (in 1991).

Table A1 in the Appendix contains the summary statistics for our target variables, which are reported for the *movers* and *stayers* sub-samples separately. The control variables used in our regression analysis are described in Table A2. This table also gives the corresponding descriptive statistics on the two sub-samples separately.[13]

Not surprisingly, Tables A1 and A2 reveal that families who change neighborhood appear to have a higher socio-economic status (higher income, education and social status) than families that do not move, and to live on average in better quality neighborhoods (higher average human capital, lower unemployment rate, higher proportion of employees in managerial and professional occupations, lower proportions of disadvantaged households). This is a clear evidence of sorting of families into neighborhoods based on observables, which suggests a potential important sorting on unobservable as well. Our estimation strategy (comparing *movers* and *stayers*) will allow us to get some insights about the importance of these issues in our analysis.[14]

5.3 Descriptive evidence

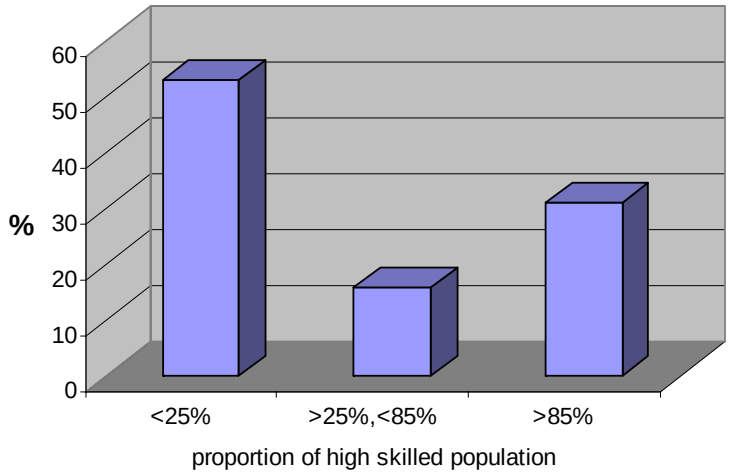
In this section, using our entire data set, we provide some descriptive evidence on the link between children's education attainment and the quality of neighborhood where they live.

[13] All data can be obtained from the UK Data Archive. We acknowledge the original data creators and depositors. They bear no responsibility for the analyses and interpretations presented here.

[14] We also report in Tables A3, A4, A5 and A6 in the Appendix the descriptive statistics separately for high educated and low educated both for the movers and stayers sub-samples.

First, looking at the distribution of UK Census wards by average human capital of the residential community (i.e., the proportion of high-skilled workers per ward), we find that most neighborhood residential communities are highly homogeneous with respect to their educational attainment. In Figure 2, we consider the percentage of high-skilled (that is the percentage of persons over 18 years old holding a A-level or higher qualification) in a ward (neighborhood) and the percentage of wards having a certain level of average human capital. It can be seen that roughly 50 percent and nearly 30 percent of wards have respectively less than 25 percent and more than 85 percent of high-skilled workers. This means that almost 80 percent of these areas are very homogenous along the education dimension (since they have either less than 25 percent or more than 85 percent of high-skilled workers).

Figure 2: Distribution of neighborhoods by average human capital quality



Second, Figure 3 displays the relationship between the average neighborhood human capital quality and the average frequency that a child, having parents of type $i = h, l$, is of type $j = h, l$. To be more precise, for each ward k, we calculate the following empirical probability:

$$p_k^{ij} = \frac{1}{N_k} \sum_{n=1}^{N_k} s_n^{ij}$$

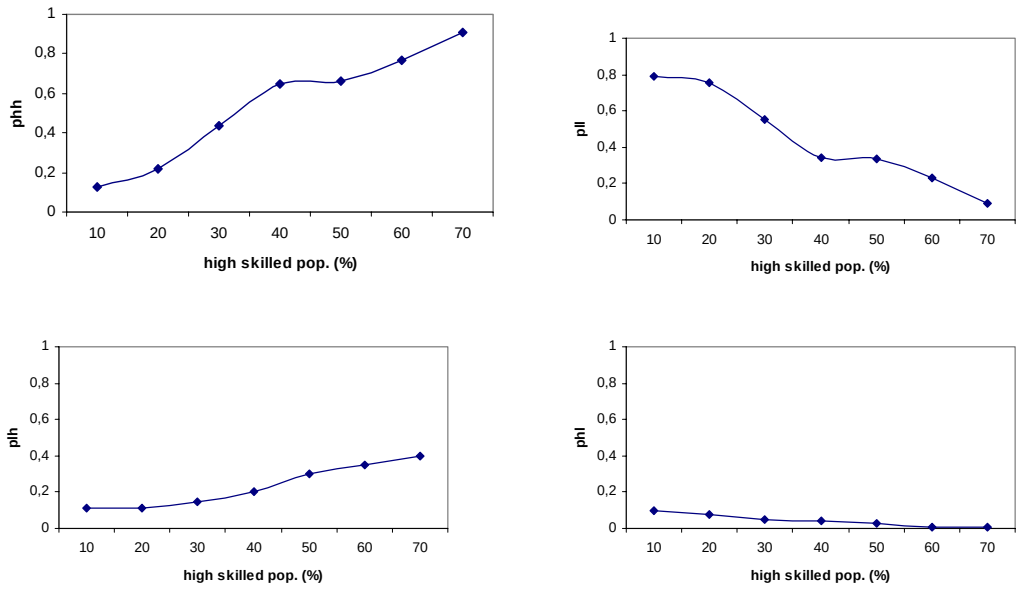
where N_k denotes the number of observations (children) in each area k, s_n^{ij} is a dummy variable that is equal to one if a child n $(n = 1, ..., N_k)$, of type $j = h, l$, who resides in ward

k ($k = 1, ..., K$), has a parent of type $i = h, l$, and is equal to zero otherwise. Thus, for each ward k, the four probabilities p_k^{hh}, p_k^{ll}, p_k^{lh} and p_k^{hl} sum up to one. These empirical frequencies are then averaged over the areas with a similar proportion of high skilled population, that is:

$$p_r^{ij} = \frac{1}{K_r} \sum_{k_r=1}^{K_r} p_{k_r}^{ij}$$

where K_r is the number of wards having an observed percentage interval r ($r = 10, ..., 70$) of high-skilled residents and $p_{k_r}^{ij}$ is the empirical probability for each area k in the different groups.

Figure 3: Average frequencies per neighborhood quality*



*p^{lh} is rescaled by 100

For example, for $r = 20$, p_{20}^{ij} is the observed average frequency of children of type j whose parents are of type i and who reside in neighborhoods with a percentage of high-skilled people between 10 and 20 percent. The values p^{ij} (i.e. the p_r^{ij} for different values of r) are reported in Figure 3, where a line has been drawn between the different points in each panel. It is striking to observe the patterns of the correlations between neighborhood quality and the probability to be educated. Indeed, both p^{hh} and p^{lh} appears as increasing functions of the residential neighborhood quality whereas both p^{ll} and p^{hl} are decreasing functions of the

same variable. This documents a positive assortative matching of these frequencies along the neighborhood human capital quality, which suggests that, irrespective of parental education, better quality neighborhoods might be associated with higher chances to be educated. If we investigate further, it also appears that parents' education may play a role. Take for example $r = 10$, i.e. very low quality neighborhoods (less than 10 percent high-skilled workers). If one compares children whose parents have different backgrounds, then, conditionally on the neighborhood quality, the chance to be uneducated seems to be much higher for those with low-educated parents ($p_{10}^{ll} \approx 80\%$) than with high-educated parents ($p_{10}^{hl} \approx 10\%$). At the opposite, in a good quality neighborhood (60-70 percent are skilled), the chance to be uneducated for a child whose parents are educated (p_{70}^{hl}) appears virtually zero while for a child whose parent is not educated (p_{70}^{ll}) it is roughly 10 percent.

Even though this evidence can be driven by an endogenous sorting of families into residential locations, it suggests the possibility that both the family background and the quality of the neighborhood may affect the educational attainment of children growing up in the area. In order to better understand these results, we now test our theoretical model where the relationship between parents' involvement in education, the quality of the neighborhood and children's educational attainment is explicitly analyzed.

6 Estimation results

Before commenting the estimation results, we discuss Table 2, which gives simple information on parents' effort e^{i*}. It provides some evidence in line with Proposition 1 (i), which stated that more educated parents put more effort in educating their children than less educated parents. Indeed, depending on the way parental interest is measured and at which age of the child it is calculated, between 60 and 80 percent of children have educated parents who are highly interested in their education while it is roughly between 20 and 40 percent for children of less educated parents. For example, if we consider parents taking the initiative to discuss about their children with the headmaster or any member of the teaching staff at the child's age 7, then 74 percent of them are highly-educated while only 26 percent are not. As stated above, these are evidence that low-educated parents prioritize less their children's education than high-educated parents.

Let us now focus on the estimation results of our bivariate probit model (14)-(15), run on the *movers* and *stayers* sub-samples separately. For exposition purposes we report in Table 3 the results concerning equation (14) and in Table 4 those referring to equation (15). In other words, we proceed as follows. Firstly, we provide evidence of whether or not the educational

composition of a neighborhood affects high-skilled and low-skilled parents' involvement in their children's education (Table 3, tests of Propositions 1 and 3). Secondly, we assess the relative importance of high-skilled (low-skilled) parents' involvement in the child's education and average neighborhood human capital on the child's ultimate educational attainment (Table 4, tests of Proposition 2 and 3).

6.1 Neighborhood quality and parental behavior

Table 3 reports the marginal effects (at the sample means) and standard errors (in parentheses) of our measure of neighborhood quality (i.e. percentage of high-skilled population) based on the maximum likelihood estimation of model (14)-(15), when using the most extensive set of controls (listed in Table A2). The second column shows the results for the *movers* sub-sample, whereas the third column those for the *stayers* sub-sample. We also report the results on the regressors used as identifying restrictions. Their estimated effects are always significant and with the expected sign, in both columns. Focussing the attention on our variable of interest, we find the following evidence.[15]

[*Insert Table 3 here*]

All estimated coefficient are always much higher in magnitude for *movers* than for *stayers*, irrespective of the parents' education level. This may suggest that parents who decide to relocate are those that are more concerned about their children's education, even within skill groups. Nevertheless, the estimated effect obtained on the sub-sample of *stayers* is still statistically different from zero. Specifically, for high-educated parents, we find a positive and statistically significant effect of neighborhood quality on parental effort in the child's education. This suggests *cultural complementarity* in parents' behavior since the better the quality of the neighborhood, the more they invest in their child's education. Regarding the magnitude of the effect, a marginal increase (1 percent increase) in the average level of education of the neighborhood increases the probability that the parents devote more effort in their children education by 0.04. For low-educated parents, the estimated effect is always positive and significant but largely reduced in magnitude (slightly more than 0.01). In conformity with Propositions 1 and 3, this may be due to the fact that δ, the returns to parents' effort or parents' priority in terms of education, is quite small. A $t-$test on

[15] We focus on the estimated impact of our target variable for movers and stayers and skill groups. For the sake of brevity, we do not comments the results related to our control variables, which are anyway in line with the expectations. Alternative dependent variables and different selections of control variables have also been used, but the qualitative results remain unchanged.

the contrast between the estimated coefficients for high and low-educated parents rejects always the null hypothesis of equal effects. This provides a formal test for differences in neighborhood effects on parents of different education levels.

6.2 Child's educational attainment

Let us now focus on the test of the implications of the model with respect to the expected school achievement of each individual (test of Proposition 2).

Tables 4 has the same structure as Table 3. It reports the marginal effects (at the sample means) and standard errors (in parentheses) of parental effort in the child's education and residential community human capital divided between *movers* and *stayers*, when using the most extensive set of controls.[16,17] The dependent variables are the transition probability π^{ij} described in Proposition 2. Clearly, conditionally on parental education, the probabilities that a child achieves A-level degree and that she/he does not, sum up to one. Thus, we only report the results for π^{hh} and π^{ll} ($\pi^{hl} = 1 - \pi^{hh}$ and $\pi^{lh} = 1 - \pi^{ll}$).[18]

[*Insert Table 4 here*]

Table 4 clearly confirms the important role of spatial sorting of families by showing, in particular, an estimated impact of the average neighborhood human capital that is roughly five time bigger in magnitude for *movers* than for *stayers*, for both skill groups. However, the finding of statistical significant effects for the *stayers* indicates that selection issues are

[16] Note that among the individual-level variables we include both arithmetic and reading test scores when the child is 7 years old aiming to control for child's ability. Also, among the area-level controls, we include the proportion of employed in agriculture to account for area industry specialization on education choices and total area population to control for agglomeration effects. Indeed, children that grow up in agricultural areas will be more likely to leave school to continue in agriculture work, regardless of the parental involvement in their education; and any intergeneration link may be affected by different degrees of individuals' social networks and physical proximity, that distinguish urban from rural areas.

[17] The inclusion of the interaction term leaves the effects of the other variables almost unchanged (only slightly lower in absolute values), and its estimated coefficient (estimate of ρ in equation (15)) is never significant in any model specification. This indicates that the effect of parental effort on children's education attainment does not vary with neighborhood human quality. Thus, we focus our attention on the results contained in Table 4, which exclude the interaction term from the regressors.

[18] Also in this case, we focus our attention on the estimated impact of our target variables for movers and stayers and skill groups. The complete set of results for all the control variables in the different model specifications and sub-samples is in line with the expectations. Also in this case, different sets of control variables have been used, but the qualitative results remain unchanged.

not the only responsible for this evidence. The mechanisms highlighted by our theoretical model seem at work.

Let us start with high-educated parents and thus focus on π^{hh} (and π^{hl}). All estimated coefficients are always significant both for parental interest and neighborhood quality and with a positive sign (thus negative for π^{hl}). In words, children whose parents are educated are more likely to be *educated* if parents spend time educating them and if the neighborhood where they live is of good quality. On the contrary, the less parents are interested in their child's schooling and the worse the neighborhood quality, the more likely children, whose parents are educated, will be *uneducated*. Comparing the magnitudes of the effects, it appears that parental interest has a higher impact than neighborhood quality. Looking at the sub-sample of *stayers*, i.e. where sorting issues are mitigated, when all the controls are considered, a marginal increase in the quality of the neighborhood raises the average probability of a successful transmission of the parental level of education by about 0.06, whereas for parental interest, the effect is roughly 0.09. This indicates that the latter effect is more potent than the former one.

Concerning low-educated parents, i.e. π^{ll} (and π^{lh}), it is instead the quality of the neighborhood that shows the more influential effect on children's educational attainment. Going back to the model and having in mind Table 2, one can conjecture that this result is due to a low δ, that is low-educated parents prioritize less education than high-educated parents. Looking again at the *stayers* sub-sample, the estimated impact of parental effort slightly reaches 0.01 in absolute value. Concerning the influence of the neighborhood, we obtain the expected sign, that is negative for π^{ll} (and positive for π^{lh}). In words, a better quality neighborhood has a negative impact on the chance to be uneducated (and a positive impact on the chance to become educated). Observe that, when all the controls are considered, when the quality of the neighborhood increases marginally, the decrease in the average probability that a child remains low educated is -0.11. This implies that the effect of neighborhood quality is larger (almost twice as much) for π^{lh} (0.11) than for π^{hl} (-0.06), which means that the percentage of high-skilled population in a neighborhood plays a major role in determining the chance to become educated for a child whose parents are not educated (in respect of its influence in decreasing the chance to remain uneducated if the parents are highly educated).

These results suggest that a failure in transmitting education for high-educated parents is more related to their lack of interest or time rather than to a negative influence of the local environment. On the other hand, children whose parents are low-skilled have some chance to obtain higher degrees if they live in a good neighborhood. In that case, parental dedication in education seems to play a minor *direct* role, possibly because they do not

prioritize education. Clearly, however, parental interest may have an *indirect* affect on children's education, as parents choose the residential neighborhood, and this decision may be driven by considerations about their children's life chances

6.3 Robustness checks

As a robustness check, we consider an alternative approach to mitigate the effects of a possible spatial sorting of families. We estimate equations (14)-(15) only on council tenants.[19] This is because council houses are allocated to tenants by local authorities. Thus, council tenants' neighborhood of residence should not be the result of their ability to pay, education level and concerns about their children's education. Even though one may imagine a discretion in the allocation process (e.g. if better educated tenants push for accommodations in better quality neighborhoods), council tenants' choice of residential location are definitely less correlated with their ability to pay for housing and preferences than is the choice of home-buyers and private renters.[20] Clearly we over-sample here low income families, but we still find enough variation to appreciate differences in our target variables. We also maintain the selection of adopted children only, to cope with unobserved inherited characteristics.

Table 5 collects the results for our target variables of this exercise for equations (14) and (15).[21] The evidence is virtually qualitatively unchanged and the magnitudes of the estimated coefficients resemble those obtained for the *stayers* sub-sample in Tables 3 and 4. These results thus confirm the important role of unobservables in our analysis. At the same time, however, they provide further empirical evidence in line with the model predictions. Indeed, Table 5 reveals the predicted presence of important community/family interactions. In particular, we still find that for high-educated parents the child's educational attainment is more affected by the parents' involvement than by the neighborhood quality while, for low-educated parents, the neighborhood quality seems to play the major role.

[*Insert Table 5 here*]

[19] Council tenants refer to families who live in council provided accommodations. These programs are equivalent to the housing projects in the US. In our whole sample, 36 percent of families are under this scheme.

[20] Indeed, we find a significant correlation between neighborhood quality and tenants' education level (using local area controls) but its magnitude (0.002) is almost twenty times lower than the one estimated for home-owners and private renters (0.038).

[21] The complete lists of results for the control variables (that remain those listed in Table A2) are qualitatively unchanged. They are not reported here for brevity (available upon request).

7 Conclusion

This paper has proposed a microeconomic mechanism of neighborhood effects on educational attainment based on parents' involvement in education. The theoretical model postulates potential interactions between residential community environment and parental investment in a child's education. The residential neighborhood average human capital affects parents' effort in their children's education, which in turn play an important role in determining children's schooling achievement. These potential interactions depend on the parents' education level. The empirical evidence supports these predictions. In particular, the children's educational attainment of high-educated parents appears to be more influenced by parents' involvement while, for low-educated parents, it is instead the neighborhood quality that plays the more influential role.

References

[1] Becker, G.S. (1974), "A theory of social interactions," *Journal of Political Economy* 82, 1063-1093.

[2] Becker, G.S. (1991), *A Treatise on the Family*, Cambridge: Harvard University Press.

[3] Becker, G.S. and N. Tomes (1979), "An equilibrium theory of the distribution of income and intergenerational mobility," *Journal of Political Economy* 87, 1153-1189.

[4] Benabou, R. (1993), "Workings of a city: Location, education, and production", *Quarterly Journal of Economics* 108, 619-652.

[5] Bisin, A., Topa, G. and T. Verdier (2004), "Religious intermarriage and socialization in the United States," *Journal of Political Economy* 112, 615-664.

[6] Bisin, A. and T. Verdier (2000), "Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits", *Quarterly Journal of Economics* 115, 955-988.

[7] Bisin, A. and T. Verdier (2001), "The economics of cultural transmission and the dynamics of preferences", *Journal of Economic Theory* 97, 298-319.

[8] Borjas, G.J. (1992), "Ethnic capital and intergenerational mobility," *Quarterly Journal of Economics* 107, 123-150.

[9] Boyd, R. and P. Richerson (1985), *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.

[10] Cavalli-Sforza, L. and M. Feldman (1981), *Cultural Transmission and Evolution. A Quantitative Approach*, Princeton: Princeton University Press.

[11] Card, D. and A. Krueger (1992), "Does school quality matter? Returns to education and the characteristics of public schools in the United States," *Journal of Political Economy* 100, 1-40.

[12] Card, D. and A. Krueger (1996), "Labor market effects of school quality: Theory and evidence," in G. Burtless (ed.), *Does Money Matter? The Link between Schools, Student Achievement and Adult Success*, Washington D.C: The Brookings Institution.

[13] De Bartolome, C. (1990), "Equilibrium and inefficiency in a community model with peer group effects," *Journal of Political Economy* 98, 110-133.

[14] Durlauf, S.E. (2004), "Neighborhood effects," in J.V. Henderson and J-F. Thisse (Eds.), *Handbook of Regional and Urban Economics Vol. 4*, Amsterdam: Elsevier Science, Ch. 50, pp. 2173-2242.

[15] Ermisch, J. and M. Francesconi (2001), "Family matters: Impacts of family background on educational attainments," *Economica* 68, 137-156.

[16] Hanushek, E. (2002), "Publicly provided education," NBER Working Paper No. 8799.

[17] Haveman, R. and B. Wolfe (1995), "The determinants of children's attainments: A review of methods and findings," *Journal of Economic Literature* 33, 1829-1878.

[18] Ioannides, Y.M. (2002), "Nonlinear neighbourhood interactions and intergenerational transmission of human capital," In: *Essays in Economic Theory, Growth and Labour Markets: A Festschrift in Honour of Emmanuel Drandakis*, G. Bitros and Y. Katsoulakos (Eds.), Gloucester: Edward Elgar, pp. 75-112.

[19] Ioannides, Y.M. (2003), "Empirical nonlinearities and neighbourhood effects in the intergenerational transmission of human capital," *Applied Economics Letters* 10, 535-539.

[20] Katz, L.F., Kling, J.R. and J.B. Liebman (2001), "Moving to opportunity in Boston: Early results of a randomized mobility experiment," *Quarterly Journal of Economics* 116, 607-654.

[21] Kremer, M. (1997), "How much does sorting increase inequality?" *Quarterly Journal of Economics* 112, 115-139.

[22] Laferrère, A. and F.-C. Wolff (2004), "Microeconomic models of family transfers," in L.-A. Gérard-Varet, S.-C. Kolm and J. Mercier Ythier (Eds.), *Handbook on the Economics on Giving, Reciprocity and Altruism*, Amsterdam: Elsevier Science, Ch. 12.

[23] Leibowitz, A. (1974), "Home investments in children," *Journal of Political Economy* 82, S111-S131.

[24] Office of National Statistics (2003), "Trends in education and skills," Department for Education and Skills, London, http://www.dfes.gov.uk/trends.

[25] Oreopoulos, P. (2003), "The long-run consequences of living in a poor neighborhood", *Quarterly Journal of Economics* 4, 1533-75.

[26] Plug, E. and W. Vijverberg (2003), "Schooling, family background, and adoption: Is it nature or is it nurture?", *Journal of Political Economy* 111, 611-641.

[27] Sacerdote, B. (2002), "The nature and nurture of economic outcomes," *American Economic Review Papers and Proceedings* 92, 344-348.

[28] Solon, G., Page, M.E. and G.J. Duncan (2000), "Correlations between neighboring children in their subsequent educational attainment," *Review of Economics and Statistics* 82, 383-392.

Table 1. Mother's attitude during pregnacy, parental effort in the child's education and child's education attainment

	Mother quits smoking	
	high educated parents	low educated parents
parents read to child	0.6592***	0.5993**
	(0.0791)	(0.0881)
single parent families	0.1618	−0.1919
	(0.1942)	(0.2055)
parents' education	0.0475	0.0725
	(0.0535)	(0.0849)
parents' social class	−0.0302	0.0922
	(0.0934)	(0.0937)
child's education attainment	0.0607	0.0966
	(0.0820)	(0.1009)
	Mother's antenatal visits	
	high-educated parents	low-educated parents
parents read to child	0.7881***	0.8495***
	(0.1006)	(0.1109)
single parent families	0.1338	−0.0619
	(0.1411)	(0.0844)
parents' education	0.1515	0.2099
	(0.1705)	(0.2094)
parents' social class	−0.0046	0.0151
	(0.0052)	(0.0207)
child's education attainment	0.1051	0.1412
	(0.1063)	(0.1555)

Notes:

- correlation coefficients and standard errors in parentheses

- coefficients marked with one (two) [three] asterisks

are significant at the 10 (5) [1] percent level

Table 2. Parental involvement by education group

Variable	high educated parents	low educated parents
1) parental interest age 7	60%	40%
2) parental interest age 11	62%	38%
3) parents read to child age 7	57%	43%
4) parental initiative to discuss child age 7	74%	26%
5) parental initiative to discuss child age 11	69%	31%
6) parents/teachers discussions age 16	79%	21%
7) parental anxiety education age 16	64%	36%

Notes: we report the precise sweep, interview and question from the NCDS and how we use this information.

1)-2) NCDS1-2 school interview: "With regard to the child's educational progress, do the mother/father appear: over concerned about the child's progress and/or expecting too high a standard? Very interested ? To show some interest? To show little or no interest ? Can't say or inapplicable." We consider the percentage of children having each parent over concerned or very interested.

3) NCDS1 parental interview. "Does the mother/father read to, or read with, the child?" "Yes, at least every week; Yes occasionally; Never or hardly never; Don't know or inapplicable." We consider the percentage of children having both parents reading at least every week.

4) NCDS1 school interview: "Since September, 1964, have the parents taken the initiative to discuss the child, even briefly, with you (headmaster) or any member of your teaching staff?" We consider the percentage of children having "yes" to this question.

5) NCDS2 school questionnaire: "Since the beginning of the school year, has either parent taken initiative to discuss the child even briefly with you (headmaster) or any member of your teaching staff?" We consider the percentage of children having the father and/or the mother that took such initiative.

6) NCDS3 parental interview: "Ask the parent how many times during the past twelve months he/she has discussed the study child's school progress with his/her teachers. If no such discussion write 0 in box. If 9 or more, please write 9." The variable, ranging from 0 to 9, has a mean of 1.42 and a standard deviation of 1.59. We consider the percentage of children having the parents answering more than 3 to this question.

7) NCDS3 individual interview: "How anxious do you think your parents are that you should do well at school?" "Very anxious, Fairly anxious, Content if I do my best, They don't mind one way or another, Uncertain." We consider the percentage of children having each parent very anxious or fairly anxious.

Table 3. Maximum Likelihood estimation of the bivariate probit model (14)-(15)

Marginal effects for equation **(14)**

high educated parents: HEP; low educated parents: LEP

dep. var.: parents read to child

	Movers	Stayers
HEP-high skilled population	0.1815***	0.0441***
	(0.0347)	(0.0152)
HEP-instrumental variables:		
mother quits smoking	0.0549***	0.0344***
	(0.0183)	(0.0105)
mother's antenatal visits	0.0629**	0.0465**
	(0.0286)	(0.0217)
LEP-high skilled population	0.0645**	0.0124**
	(0.0302)	(0.0061)
LEP-instrumental variables:		
mother quits smoking	0.0539**	0.0312**
	(0.0257)	(0.0148)
mother's antenatal visits	0.1360***	0.0789***
	(0.0397)	(0.0255)
Control set:		
child variables	yes	yes
neighborhood variables	yes	yes
family background variables	yes	yes
school variables	yes	yes
regional dummies	yes	yes

Notes:

- precise list and definitions of control variables by group in Table A2, A4, A6

- marginal effects at the sample means and standard errors in parentheses

- coefficients marked with one (two) [three] asterisks

are significant at 10 (5) [1] percent level

- errors are clustered at the neighborhood level

- sub-samples sizes: movers 2,723; movers with high educated parents 1,579;

stayers 1,981; stayers with high educated parents 911

Table 4. Maximum Likelihood estimation of the bivariate probit model (14)-(15)

Marginal effects for equation **(15)**

	Movers	Stayers
	dep. var.: π^{hh}	
parents read to child	0.3612***	0.0923***
	(0.0674)	(0.0202)
high skilled population	0.3066**	0.0588***
	(0.0582)	(0.0161)
	dep. var.: π^{ll}	
parents read to child	−0.1055**	−0.0102**
	(0.0504)	(0.0049)
high skilled population	−0.5428***	−0.1105***
	(0.0712)	(0.0260)
Control set:		
child variables	yes	yes
neighborhood variables	yes	yes
family background variables	yes	yes
schooling variables	yes	yes
regional dummies	yes	yes

Notes:

- precise list and definitions of control variables by group in Table A2, A4, A6

- marginal effects at the sample means and standard errors in parentheses

- coefficients marked with one (two) [three] asterisks

are significant at 10 (5) [1] percent level

- errors are clustered at the neighborhood level

- sub-samples sizes: movers 2,723; movers with high educated parents 1,579;

stayers 1,981; stayers with high educated parents 911

Table 5. Robustness checks

Sample of Council tenants with adopted children

high educated parents: HEP; low educated parents: LEP

	Marginal effects for Equation (14)
	dep. var.: parents read to child
HEP-high skilled population	0.0512^{***}
	(0.0170)
LEP-high skilled population	0.0215^{**}
	(0.0106)

	Marginal effects for Equation (15)
	dep. var.: π^{hh}
parental interest	0.1005^{***}
	(0.0302)
high skilled population	0.0701^{***}
	(0.0229)
	dep. var.: π^{ll}
parental interest	-0.0148^{***}
	(0.0051)
high skilled population	-0.1216^{***}
	(0.0300)

Control set:	
child variables	yes
neighborhood variables	yes
family background variables	yes
schooling variables	yes
regional dummies	yes

Notes:

- precise list and definitions of control variables by group in Table A2, A4, A6

- marginal effects at the sample means and standard errors in parentheses

- coefficients marked with one (two) [three] asterisks

are significant at 10 (5) [1] percent level

- errors are clustered at the neighborhood level

- sub-samples sizes: whole sample 2,109; high educated parents 788

Appendix

Table A1. Description of key variables

Variable		Stayers	Movers
high skilled population***	Census ward proportion of over-18s persons with A-levels (highest grade at age 16 exams) or above qualifications	13.02 (.3.58)	14.34 (3.34)
parents read to child***	dummy variable taking value one if the parents read at least every week to the child	0.45 (0.15)	0.74 (0.20)
parent education***	average completed years of schooling (derived from age left full-time education) of the mother and the father	9.68 (2.39)	12.72 (3.11)
high educated parents***	dummy variable taking value one if the average completed years of schooling of the parents is greater than 12	0.45 (0.19)	0.58 (0.24)
n. obs.		1,981	2,723

Notes:

- A parent is the mother or the father or the person acting as mother or father respectively

- Mean values and standard deviations (in parentheses) are reported

- t-tests for differences in means are performed

- Variables marked with one (two) [three] asterisks denote differences of mean values that are significant at the 10 (5) [1] percent level

Table A2. Description of control variables used in the regressions

		Stayers	Movers
✓ child variables			
special education	dummy variable taking value one if the child has been ascertained as in need of special education (speech defect, physically handicapped, partially sighted/ hearing)	0.23 (0.42)	0.21 (0.59)
arithmetic test score***	child's age-7 arithmetic test scores, coded 0 to 10	5.49 (3.14)	6.25 (3.81)
reading test score***	child's age-7 reading test scores, coded 0 to 30	22.32 (7.05)	24.68 (5.64)
female	dummy variable taking value one if the child is female	0.53 (0.51)	0.51 (0.54)
✓ family background variables			
parents income***	weekly net wage of father (or mother if no father figure), 12 bands, mid-points of each range considered	24,22 (11,15)	29.86 (13.42)
parents age	average parents' age at the childs's age 16	41.55 (8.65)	41.24 (10.14)
parents social class**	social class of father (or mother if no father figure), coded 1 to 5: unskilled, semi-skilled manual, skilled manual, skilled non-manual, professional	3.86 (1.51)	3.97 (1.97)
parents employed***	dummy variable taking value one if both parents are working	0.64 (0.27)	0.82 (0.48)
parents born in UK*	dummy variable taking value one if both parents are born in Great Britain	0.79 (0.66)	0.76 (0.44)
single parent families	dummy variable taking value one if there is no regular father figure or there is no natural mother	0.003 (0.055)	0.004 (0.047)
household health problems	dummy variable taking value one if the family experienced health-related difficulties (serious ill-health of a member of the household, including death of mother or father)	0.09 (0.15)	0.09 (0.14)
household financial problems***	dummy variable taking value one if the family experienced financial difficulties	0.08 (0.24)	0.04 (0.18)
house size*	number of rooms in household accommodation	4.01 (1.42)	4.54 (1.73)
family size**	number of people in household	4.88 (2.64)	4.70 (2.75)

✓ school variables			
school composition***	proportion of boys or girls studying for GCE and SCE O-levels in the school attended by the child at age 16, 9 bands, coded 1 to 9	4.60 (2.24)	4.92 (2.35)
school private***	dummy variable taking value one if the school attended by the child at age 16 is private	0.21 (0.22)	0.27 (0.13)
school grammar***	dummy variable taking value one if the school attended by the child at age 16 is grammar	0.07 (0.19)	0.14 (0.16)
school secondary modern***	dummy variable taking value one if the school attended by the child at age 16 secondary modern	0.09 (0.09)	0.11 (0.13)
✓ neighborhood variables	(Census 1971)		
young population*	Census ward proportion of persons aged less than 21	31.02 (21.44)	32.24 (25.77)
total population (thousands)*	Census ward total residing population	52.68 (53.27)	55.24 (60.43)
unemployment rate***	Census ward unemployed over active population	0.13 (0.05)	0.07 (0.07)
activity rate	Census ward active population (aged more than 15) over present population	0.60 (0.35)	0.58 (0.55)
professional employment***	Local Authority professional and managerial employees over active population	0.14 (0.08)	0.20 (0.14)
unskilled employment*	Local Authority unskilled manual employees over active population	0.06 (0.16)	0.05 (0.23)
agriculture employment***	Local Authority proportion of workers in agriculture employment	0.04 (0.04)	0.02 (0.01)
amenities***	Census ward proportion of households lacking or sharing hot water and/or inside toilet	0.15 (0.19)	0.08 (0.18)
council housing***	Census ward proportion of households residing in council houses	0.31 (0.36)	0.04 (0.02)
car access***	Census ward proportion of households with no car	0.24 (0.41)	0.17 (0.46)
n. obs.		1,981	2,723

Notes:

- GCE (General Certificate of Education)and SCE (Scottish Certificate of Education) O-levels (Ordinary levels) were taken at age 16 mainly by pupils in grammar schools and independent schools-nationally the top 20% of the population by ability

- A parent is the mother or the father or the person acting as mother or father respectively

- Mean values and standard deviations (in parentheses) are reported

Table A3. Description of key variables

-stayers sub-sample-

high educated parents: HEP; low educated parents: LEP

Variable		HEP	LEP
high skilled population***	Census ward proportion of over-18s persons with A-levels (highest grade at age 16 exams) or above qualifications	13.79 (2.96)	13.01 (2.55)
parents read to child***	dummy variable taking value one if the parents read at least every week to the child	0.52 (0.11)	0.39 (0.10)
parent education***	average completed years of schooling (derived from age left full-time education) of the mother and the father	12.51 (2.22)	9.12 (2.53)
n. obs.		911	1,070

Notes:

- A parent is the mother or the father or the person acting as mother or father respectively

- Mean values and standard deviations (in parentheses) are reported

- t-tests for differences in means are performed

- Variables marked with one (two) [three] asterisks denote differences of mean values that are significant at the 10 (5) [1] percent level

Table A4. Description of control variables used in the regressions

-stayers sub-sample-

high educated parents: HEP; low educated parents: LEP

		HEP	LEP
✓ child variables			
special education	dummy variable taking value one if the child has been ascertained as in need of special education (speech defect, physically handicapped, partially sighted/ hearing)	0.25 (0.50)	0.22 (0.48)
arithmetic test score**	child's age-7 arithmetic test scores, coded 0 to 10	5.54 (2.82)	5.25 (3.42)
reading test score***	child's age-7 reading test scores, coded 0 to 30	23.01 (6.79)	21.50 (6.41)
female***	dummy variable taking value one if the child is female	0.41 (0.45)	0.61 (0.50)
✓ family background variables			
parents income*	weekly net wage of father (or mother if no father figure), 12 bands, mid-points of each range considered	24.80 (13.19)	23.90 (10.04)
parents age***	average parents' age at the childs's age 16	38.75 (9.68)	41.75 (7.14)
parents social class*	social class of father (or mother if no father figure), coded 1 to 5: unskilled, semi-skilled manual, skilled manual, skilled non-manual, professional	3.95 (2.31)	3.80 (1.26)
parents employed***	dummy variable taking value one if both parents are working	0.71 (0.27)	0.58 (0.21)
parents born in UK	dummy variable taking value one if both parents are born in Great Britain	0.81 (0.70)	0.79 (0.65)
single parent families	dummy variable taking value one if there is no regular father figure or there is no natural mother	0.004 (0.071)	0.003 (0.035)
household health problems	dummy variable taking value one if the family experienced health-related difficulties (serious ill-health of a member of the household, including death of mother or father)	0.09 (0.19)	0.08 (0.14)
household financial problems**	dummy variable taking value one if the family experienced financial difficulties	0.07 (0.21)	0.09 (0.14)
house size	number of rooms in household accommodation	4.01 (1.07)	4.00 (1.94)
family size	number of people in household	4.80 (2.46)	4.90 (2.30)

✓ school variables			
school composition***	proportion of boys or girls studying for GCE and SCE O-levels in the school attended by the child at age 16, 9 bands, coded 1 to 9	4.90 (2.42)	4.59 (2.23)
school private***	dummy variable taking value one if the school attended by the child at age 16 is private	0.24 (0.12)	0.18 (0.14)
school grammar***	dummy variable taking value one if the school attended by the child at age 16 is grammar	0.09 (0.21)	0.05 (0.16)
school secondary modern	dummy variable taking value one if the school attended by the child at age 16 secondary modern	0.10 (0.36)	0.08 (0.41)
✓ neighborhood variables	(Census 1971)		
young population***	Census ward proportion of persons aged less than 21	33.10 (11.91)	31.01 (11.08)
total population (thousands)	Census ward total residing population	52.56 (50.99)	52.42 (50.45)
unemployment rate***	Census ward unemployed over active population	0.11 (0.03)	0.13 (0.05)
activity rate	Census ward active population (aged more than 15) over present population	0.60 (0.33)	0.58 (0.55)
professional employment***	Local Authority professional and managerial employees over active population	0.16 (0.08)	0.12 (0.10)
unskilled employment	Local Authority unskilled manual employees over active population	0.06 (0.16)	0.06 (0.15)
agriculture employment***	Local Authority proportion of workers in agriculture employment	0.01 (0.03)	0.05 (0.01)
amenities***	Census ward proportion of households lacking or sharing hot water and/or inside toilet	0.11 (0.12)	0.18 (0.10)
council housing	Census ward proportion of households residing in council houses	0.29 (0.37)	0.31 (0.33)
car access	Census ward proportion of households with no car	0.23 (0.54)	0.25 (0.45)
n. obs.		911	1,070

Notes:

- GCE (General Certificate of Education)and SCE (Scottish Certificate of Education) O-levels (Ordinary levels) were taken at age 16 mainly by pupils in grammar schools and independent schools-nationally the top 20% of the population by ability

- A parent is the mother or the father or the person acting as mother or father respectively

- Mean values and standard deviations (in parentheses) are reported

Table A5. Description of key variables

-movers sub-sample-

high educated parents: HEP; low educated parents: LEP

Variable		HEP	LEP
high skilled population***	Census ward proportion of over-18s persons with A-levels (highest grade at age 16 exams) or above qualifications	14.52 (2.38)	13.65 (2.19)
parents read to child***	dummy variable taking value one if the parents read at least every week to the child	0.76 (0.16)	0.68 (0.21)
parent education***	average completed years of schooling (derived from age left full-time education) of the mother and the father	13.23 (2.89)	10.01 (2.76)
n. obs.		1,579	1,144

Notes:

- A parent is the mother or the father or the person acting as mother or father respectively
- Mean values and standard deviations (in parentheses) are reported
- t-tests for differences in means are performed
- Variables marked with one (two) [three] asterisks denote differences of mean values that are significant at the 10 (5) [1] percent level

Table A6. Description of control variables used in the regressions

-movers sub-sample-

high educated parents: HEP; low educated parents: LEP

		HEP	LEP
✓ child variables			
special education	dummy variable taking value one if the child has been ascertained as in need of special education (speech defect, physically handicapped, partially sighted/ hearing)	0.22 (0.58)	0.20 (0.45)
arithmetic test score***	child's age-7 arithmetic test scores, coded 0 to 10	7.10 (2.61)	6.17 (2.59)
reading test score***	child's age-7 reading test scores, coded 0 to 30	25.22 (3.90)	23.45 (4.56)
female*	dummy variable taking value one if the child is female	0.51 (0.39)	0.48 (0.54)
✓ family background variables			
parents income***	weekly net wage of father (or mother if no father figure), 12 bands, mid-points of each range considered	30.42 (13.01)	26.09 (15.04)
parents age**	average parents' age at the childs's age 16	40.64 (9.99)	41.69 (13.99)
parents social class***	social class of father (or mother if no father figure), coded 1 to 5: unskilled, semi-skilled manual, skilled manual, skilled non-manual, professional	4.58 (1.12)	3.44 (2.06)
parents employed*	dummy variable taking value one if both parents are working	0.83 (0.20)	0.81 (0.38)
parents born in UK	dummy variable taking value one if both parents are born in Great Britain	0.77 (0.34)	0.76 (0.35)
single parent families	dummy variable taking value one if there is no regular father figure or there is no natural mother	0.005 (0.033)	0.004 (0.022)
household health problems	dummy variable taking value one if the family experienced health-related difficulties (serious ill-health of a member of the household, including death of mother or father)	0.09 (0.15)	0.08 (0.19)
household financial problems	dummy variable taking value one if the family experienced financial difficulties	0.03 (0.27)	0.04 (0.17)
house size*	number of rooms in household accommodation	4.61 (1.55)	4.51 (1.47)
family size	number of people in household	4.65 (2.15)	4.71 (1.01)

✓ school variables			
school composition*	proportion of boys or girls studying for GCE and SCE O-levels in the school attended by the child at age 16, 9 bands, coded 1 to 9	4.96 (1.21)	4.84 (2.13)
school private***	dummy variable taking value one if the school attended by the child at age 16 is private	0.32 (0.20)	0.27 (0.10)
school grammar***	dummy variable taking value one if the school attended by the child at age 16 is grammar	0.17 (0.10)	0.12 (0.11)
school secondary modern***	dummy variable taking value one if the school attended by the child at age 16 secondary modern	0.13 (0.09)	0.10 (0.11)
✓ neighborhood variables	(Census 1971)		
young population***	Census ward proportion of persons aged less than 21	34.10 (10.41)	31.44 (11.99)
total population (thousands)*	Census ward total residing population	57.79 (42.29)	54.52 (51.03)
unemployment rate***	Census ward unemployed over active population	0.06 (0.06)	0.07 (0.10)
activity rate***	Census ward active population (aged more than 15) over present population	0.64 (0.45)	0.56 (0.44)
professional employment	Local Authority professional and managerial employees over active population	0.20 (0.09)	0.20 (0.11)
unskilled employment**	Local Authority unskilled manual employees over active population	0.04 (0.24)	0.06 (0.27)
agriculture employment	Local Authority proportion of workers in agriculture employment	0.02 (0.02)	0.02 (0.03)
amenities***	Census ward proportion of households lacking or sharing hot water and/or inside toilet	0.05 (0.09)	0.09 (0.08)
council housing***	Census ward proportion of households residing in council houses	0.02 (0.03)	0.05 (0.02)
car access**	Census ward proportion of households with no car	0.15 (0.30)	0.18 (0.40)
n. obs.		1,579	1,144

Notes:

- GCE (General Certificate of Education)and SCE (Scottish Certificate of Education) O-levels (Ordinary levels) were taken at age 16 mainly by pupils in grammar schools and independent schools-nationally the top 20% of the population by ability

- A parent is the mother or the father or the person acting as mother or father respectively

- Mean values and standard deviations (in parentheses) are reported