

DISCUSSION PAPER SERIES

No. 4744

**INTERGENERATIONAL EDUCATION
TRANSMISSION: NEIGHBOURHOOD
QUALITY AND/OR PARENTS'
INVOLVEMENT**

Elenora Patacchini and Yves Zenou

PUBLIC POLICY



Centre for Economic Policy Research

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP4744.asp

INTERGENERATIONAL EDUCATION TRANSMISSION: NEIGHBOURHOOD QUALITY AND/OR PARENTS' INVOLVEMENT

Elenora Patacchini, University of Southampton
Yves Zenou, IUI, Stockholm, University of Southampton and CEPR

Discussion Paper No. 4744
November 2004

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Elenora Patacchini and Yves Zenou

CEPR Discussion Paper No. 4744

November 2004

ABSTRACT

Intergenerational Education Transmission: Neighbourhood Quality and/or Parents' Involvement

We develop a model that gives some microfoundation to the impact of residential neighborhood on children's educational attainment and then test it using the UK National Child Development Study. We find that, for high-educated parents, the better the quality of the neighborhood in terms of human capital, the higher the parent's involvement in children's education, indicating cultural complementarity. For highly educated parents, we also 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. Low-educated parents do not spend much time educating their offspring and we show that only the quality of the neighborhood has a significant impact on their children's educational attainment.

JEL Classification: I21, J13 and J24

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

Elenora Patacchini
Università di Roma 'La Sapienza'
Facoltà di Scienze Statistiche
Box 83 - RM62
P. le A. Moro 5
00185 Roma
ITALY
Tel: (39 06) 4991 0843
Fax: (39 06) 4453 246
Email: eleonora.patacchini@uniroma1.it

Yves Zenou
The Research Institute of Industrial
Economics
IUI
Box 5501
114 85 Stockholm
SWEDEN
Tel: (46 8) 665 4535
Fax: (46 8) 665 4599
Email: yvesz@iui.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=158584

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=126027

Submitted 27 October 2004

1 Introduction

The investigation of neighborhood effects on individual's educational attainment has recently received considerable attention (see for example the remarkable survey of Durlauf, 2004). Many studies have focused on the measurement of neighborhood effects on individual's educational attainment using various econometric techniques. The general consensus is 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.

The open question is the mechanism underlying these effects. As pointed out by Durlauf (2004) "...the evidence of neighborhood effects in this literature is largely a black box, i.e. it is difficult to translate the findings of the papers into specific microeconomic mechanism."

The focus of this paper is on the role of neighborhood quality in the intergenerational transmission of education and our aim is to provide a simple microeconomic mechanism that explains the link between the educational attainment of children and the average educational level of their residential neighborhood community. This link is established through parents' involvement in their children's education. Indeed, in our framework, neighborhood quality influences parents' involvement in education and this, in turn, affects children's education attainment.

The literature on children's education is large and most studies have found that school quality (e.g., Card and Krueger, 1992, 1996, and Hanushek, 2002), family background (e.g., Ermisch and Francesconi, 2001, Sacerdote, 2002, Plug and Vijverberg, 2003) and neighborhood quality (e.g., Durlauf, 2004) have a significant and positive impact on the level of education of children.¹ As stated above, we provide here a new mechanism that highlights the role of parents' involvement in education.² 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,³ arguing that the transmission of a

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

²There 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 attainment. Our model links these two approaches and, as a result, gives a mechanism through which both effects affect children's outcomes.

³Indeed, Bisin and Verdier were the first to introduce cultural transmission in an economic framework. They use this concept to explain the transmission of social status (Bisin and Verdier, 1998), ethnicity and religion (Bisin and Verdier, 2000a), political power (Bisin and Verdier, 2000b) and cultural traits (Bisin and Verdier, 2001). More recently, Hauk and Saez-Marti (2002) and Saez-Marti and Zenou (2004) have used

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. To be more precise, we propose a simple theoretical model in which altruistic parents, who can 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 have this behavior while 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). We assume that uneducated parents are less efficient in educating their children than educated parents because either they are less able or do not prioritize education. In this respect, there are two aspects of education: the time spent with children and the quality or intensity of this time. Given the quality of the neighborhood, the crucial decision for each parent is how much time to spend educating their children.

First, we show that educated parents spend more time (or more effort) educating their children than uneducated parents simply because the rewards are higher. If the utility function is such that the higher the effort, the higher the marginal utility of leisure, cultural substitution prevails while if the loss in the marginal utility of leisure following an increase in effort is low, there is cultural complementarity. Second, we demonstrate that, with cultural complementarity, better-quality neighborhood increases the probability to be educated and decreases the probability to be uneducated while this is undetermined with cultural substitution. Finally, if uneducated parents are very inefficient in educating their children (because for example they do not prioritize at all education), then we show that only the neighborhood quality affects the educational attainment of their children.

cultural substitution to explain the transmission of corruption and work ethic, respectively.

We find evidence in line with these predictions 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 are 13 years old). First, adopting a number of explanatory variables to address possible unobserved variables and controlling for endogeneity issues, we find a significant and positive effect of neighborhood quality on highly educated parents' interest in their children's education, suggesting cultural complementarity.⁴ For low-educated parents, we find no significant effect, suggesting that their effort is not very high (as it is documented by comparing the effort levels of educated and uneducated parents). Second, 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 find that low-educated parents do not spend much time educating their offsprings and only the quality of the neighborhood has a significant impact on children's educational attainment.

The remainder of the paper is organized as follows. The next section presents some descriptive evidence on the relationship between education and the quality of the neighborhood. In section 3, the theoretical model and its main predictions are exposed. Section 4 deals with the empirical model and estimation issues while section 5 is devoted to the data and the definition of the variables. Section 6 gives the main empirical results and discusses endogeneity issues. Finally, section 7 concludes.

2 Descriptive evidence

In this section, using our entire data set,⁵ we provide some simple evidence on the link between children's education attainment and the quality of neighborhood where they live in the UK.

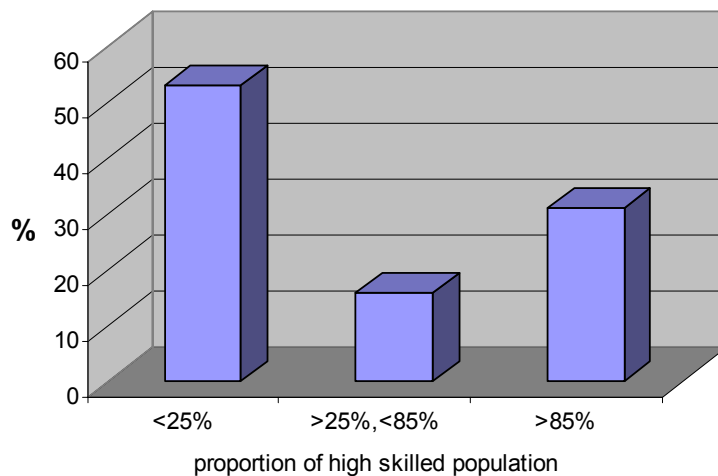
First, looking at the distribution of UK Census wards by average human capital of the residential community (i.e., the proportion of high-skilled workers per ward), we find that most neighborhood residential communities are highly homogeneous with respect to their educational attainment. In Figure 1 we consider the percentage of high-skilled (that is the

⁴To 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 so that the more a religion is in minority in a state the more parents put effort in transmitting it. For education, we find the contrary and this could be interpreted considering that education and the quality of the neighborhood may be quite naturally complement.

⁵See Section 5 for a detailed description of our data.

percentage of persons over 18 years old holding a A-level⁶ 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 1: Distribution of neighborhoods by average human capital quality



Second, Figure 2 displays the relationship between the average neighborhood human capital quality and the average frequency that a child, having parents of type $i = h, l$ (h and l stand respectively for high and low education level),⁷ 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 ($n = 1, \dots, N_k$), of type $j = h, l$, who resides in ward k ($k = 1, \dots, K$), has a parent of type $i = h, l$, and is equal to zero otherwise. These

⁶The A-level in UK is equivalent to the SAT in the US or the baccalaureat in France.

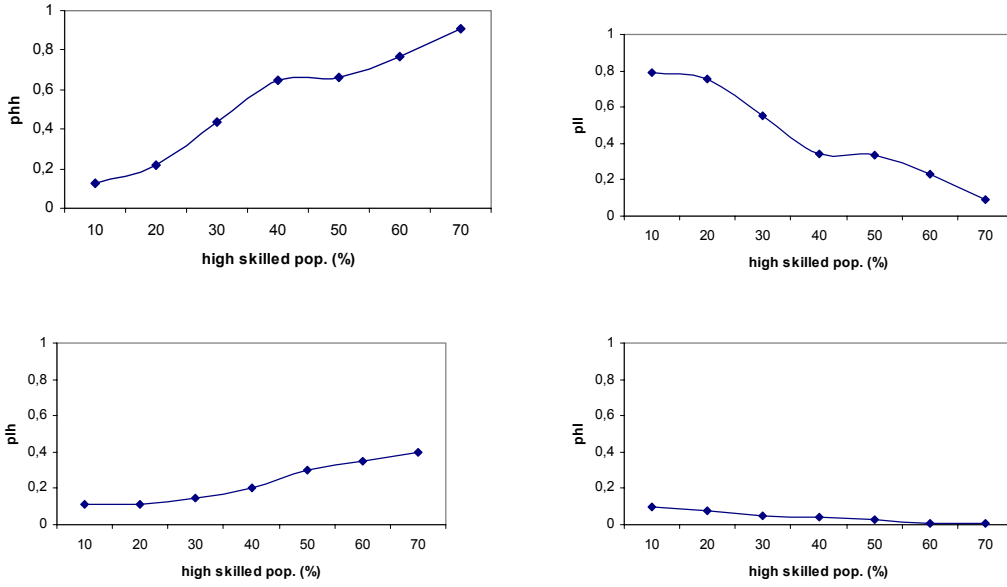
⁷We consider parents as educated if either the father or the mother is educated (having A-level qualification or above), and not educated otherwise.

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 ($r = 10, \dots, 70$) of high-skilled residents and $p_{k_r}^{ij}$ is the empirical probability for each area k in the different groups.

Figure 2: 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 2, where a line has been drawn between the different points in each panel. It is striking to observe the influence of the neighborhood on the chance to be educated. Indeed, both p^{hh} and p^{lh} are 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 means that, irrespective of parental education, better quality neighborhoods yield higher chance to be educated. If we investigate further, it is also easy to see that parents'

education does play a role. Take for example $r = 10$, i.e. very low quality neighborhoods (less than 10 percent high-skill workers). If one compares children whose parents have different backgrounds, then, irrespective of the neighborhood, the chance to be uneducated is much higher for those with low-educated parents ($p_{10}^l \approx 80\%$) than with high-educated parents ($p_{10}^h \approx 10\%$). 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}^h) is zero while for a child whose parent is not educated (p_{70}^l) it is roughly 10 percent.

Even though this evidence can be driven by an endogenous sorting of families into residential locations,⁸ 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 propose a simple theoretical model in the next section where the relationship between parents' involvement in education, the quality of the neighborhood and children's educational attainment is explicitly analyzed.

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 they want to decide how much 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.⁹

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:¹⁰

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

where z^i is the quantity of a consumption good (taken as the numeraire) consumed by the parents, λ^i is the time spent on leisure and e^i is time (effort) they spend with their children trying to educate them. $U(\cdot)$ is assumed to be increasing in λ^i and decreasing in the effort

⁸We discuss in details this issue in Section 6.3.

⁹Observe that the model provided here is quite different to that of Bisin and Verdier. Here, contrary to Bisin and Verdier, parents do not transmit a "trait" but decide how much effort they spend with their children. The fundamental difference is that, since education is not a "trait", and because of altruism, low-educated parents do not want their children to be like them and thus put effort to educate their children.

¹⁰None of our results is affected by the fact that the utility function is separable.

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:

$$w^i T = z^i \quad (1)$$

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.¹¹

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

$$1 - T = \lambda^i + e^i \quad (2)$$

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

By plugging (1) and (2) into the utility function, we obtain the following instantaneous indirect utility for parents of type $i = h, l$:

$$V^i(e^i) = z^i + U(\lambda^i, e^i) = w^i T + U(1 - T - e^i, e^i) \quad (3)$$

Let us now focus on the parent's choice of effort $e^i \in [0, 1]$. Because low-educated parents have less "knowledge" than high-educated parents, we assume that they are less successful in educating their children. In other words, if both high- and low-educated parents put effort e in educating their offsprings, the probability that their kids will be educated is e and δe (with $0 < \delta < 1$) for high- and low-educated parents respectively. Thus, 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 ($0 \leq \delta \leq 1$ captures the relative inefficiency in the education process for the uneducated parents). In a more general interpretation, δ could capture the fact that low-educated parents are less "intense" in educating their kids because they have different priorities than educated parents. In this interpretation, $\delta = 1$ means that education is a top priority for the parents so the time they spend with their kids is of very 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,

¹¹We could have assumed that more educated individuals work more hours than less educated workers. This would not affect any of our results.

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 or intensity of this time.

In both cases, if education is not successful, the child remains without education and gets randomly matched with somebody 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* workers in the economy (or the neighborhood), we have the following transition probabilities:

$$\pi^{hh} = e^h + (1 - e^h)q \quad (4)$$

$$\pi^{hl} = (1 - e^h)(1 - q) \quad (5)$$

$$\pi^{ll} = (1 - \delta e^l)(1 - q) \quad (6)$$

$$\pi^{lh} = \delta e^l + (1 - \delta e^l)q \quad (7)$$

where $0 \leq \delta \leq 1$. Let us interpret equation (4). 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 (5) gives the probability that a child of educated parents is not educated: it is because both the parents and the society were not successful in educating the child. For the non-educated parents (equations (6) and (7)), we have a similar interpretation, with the difference that parents are less efficient in educating 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. 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:¹²

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

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

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 the educated and the uneducated parents are respectively given by (we only focus on interior solutions):¹³

$$-U_{\lambda^i} + U_{e^i} + a \delta^i (1 - q)(w^h - w^l) = 0 \quad (8)$$

where $\delta^h = 1$ and $\delta^l = \delta$. The solution of (8) 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:

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:*

$$\frac{\partial e^{i*}}{\partial q} = \frac{a \delta^i (w^h - w^l)}{U_{\lambda^i \lambda^i} - U_{\lambda^i e^i}} < 0 \quad (9)$$

This is referred to as cultural substitution.

¹²The altruistic model was made famous by Becker (1974, 1991). For a recent survey on these types of models, see Laferrere and Wolff (2004).

¹³For each parent $i = h, l$, the second order condition is given by:

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

and is assumed to be negative.

(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:

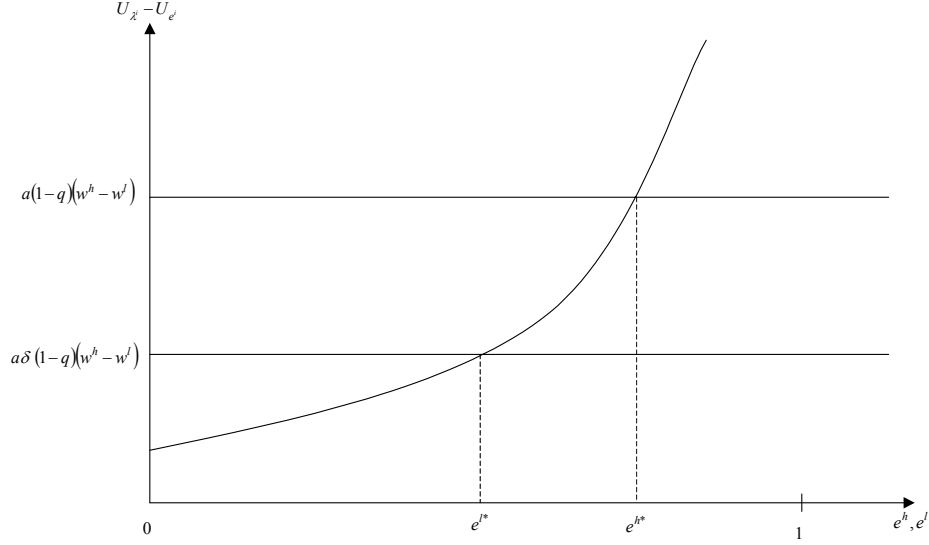
$$\frac{\partial e^{i*}}{\partial q} = \frac{a \delta^i (w^h - w^l)}{U_{\lambda^i \lambda^i} - U_{\lambda^i e^i}} > 0 \quad (10)$$

This is referred to as cultural complementarity.

The first order condition (8) 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 $\delta \approx 0$ 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 3 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 3: 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:

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

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

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

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

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

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

(iii) For both parents (educated or not), if there is cultural complementarity, 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 (4)-(7) 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 lower the efficiency of parents' effort δ , the lower the time spent with their children e^{l*} , i.e. $\frac{\partial e^{l*}}{\partial \delta} > 0$;
- (ii) the lower the efficiency of effort 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. $\delta \rightarrow 0$, 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 ($\delta \rightarrow 0$) or equivalently if the returns of parents' involvement are very low (because

for example the parents themselves do not have the capacity to help their children), 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 level of education is quite high. If, on the contrary, they live in a rundown area with schools of low qualities 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. This is an acute problem since the increase in the number of one-parent families in the past decade has been dramatic. According to the U.S. Bureau of the Census, today, over 20 percent of families with children are maintained by a separated, divorced, widowed, or never-married parent. Among these families, 9 of 10 are maintained by mothers. One-fifth of children under 18 years of age presently live in a one-parent family, an increase of well over 50% in the past 10 years.¹⁴

4 Empirical strategy and econometric issues

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

In our empirical strategy, we adopt a *two-stage procedure*.

In the first stage, we test 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 and we use as indicator variable an ordered response variable based on qualitative information on the parent's interest in his/her child's education. A standard ordered probit

¹⁴This is even more dramatic for African American families living in inner-city ghettos since in most cases the father is absent and does not live with his children (Wilson, 1987). Indeed, the average configuration of black families has changed from the 1960s, when only 25% of African-American children were living in single-parent, primarily poor, homes. In the 1990s, 54% of African American children are born into poverty, and 65% percent are living with a single parent at any given time (Johnson, 1996).

estimator is employed and from this we predict the dedication to the child's education of each parent. From Proposition 1, the following model is considered:

$$e_{z,k,t}^i = \alpha q_{k,t} + \sum_{m=1}^M \beta_m x_{m,t} + \varepsilon_{z,t} \quad z = 1, \dots, Z \quad (11)$$

where $e_{z,k,t}^i$ is the (unobservable) time spent by parent z of type $i = h, l$ who resides in area $k = 1, \dots, K$ at time t for educating his/her child (who is not yet educated); $q_{k,t}$ is the average quality in terms of education of area k at time t ; $x_{m,t}$ (for $m = 1, \dots, M$) 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; $\varepsilon_{i,t}$ is a white noise error term. The time period t has to be the offspring's childhood or early teenage since we want him/her not to be yet educated. A successful test of Proposition 1 would be that, for educated parents ($i = h$), α is significantly different from zero. For non-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). Moreover, in both cases, the sign of α would indicate either cultural substitution (if negative) or cultural complementarity (if positive). Model (11) is estimated separately for high-educated and low-educated parents, considering the pooled sample of mothers and fathers, irrespective to the fact that they are a couple. Information on each individual's partner enters only in the control variables (at the household level).

Next, turning to the second stage, that is the test of Proposition 2 (and Proposition 3), for each child, we need to have a unique indicator of the parental dedication to her/his education, distinguishing between parents of different types $i = h, l$. We define parents of type $i = l$ if both the mother and the father are low educated, parents of type $i = h$ otherwise (i.e. at least one of the two is highly educated). From model (11), the estimated propensity in investing in a child's education for each couple of parents is then averaged over the couple and used in the test of Proposition 2. Specifically, we model the likelihood of a successful or unsuccessful intergenerational transmission of education (transition probabilities in Proposition 2) as follows:

$$\pi_{n,t+1}^{ij} = \phi q_{k,t} + \gamma \widehat{e}_{k,t}^i + \rho(q_{k,t}, \widehat{e}_{k,t}^i) + \sum_{v=1}^V \theta_v x_{v,t} + \eta_{n,t+1} \quad n = 1, \dots, N \quad (12)$$

where $\pi_{n,t+1}^{ij}$ 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 ; $\widehat{e}_{k,t}^i$ is the (average) *predicted* type- i

parents' dedication to a child's schooling at time t in area k ; $\eta_{n,t+1}$ is a white noise error term. Among the other explanatory variables in the set $x_{v,t}$ (for $v = 1, \dots, V$), there are some (but not all) of the variables that entered in the set $x_{m,t}$ from model (11), allowing these variables to have a different impact on $\pi_{n,t+1}^{ij}$ than they had on $e_{z,k,t}^i$. Observe that this is of particular interest for our target variable $q_{k,t}$, which has been separated out in model (11) from the set of control variables for ease of clarity. The four probabilities π_{t+1}^{ij} are analyzed using four different 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 educated parents, both the effect of $q_{k,t}$ and $\widehat{e}_{k,t}^i$ have to be significant, whereas for children of non-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 $\widehat{e}_{k,t}^i$ 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.

Provided that $\widehat{e}_{k,t}^i$ is uncorrelated with $\eta_{n,t+1}$, the two-stage procedure adopted here produces consistent estimates of the parameters in model (12). Only a correction of the variance-covariance matrix is needed in order to account for the sampling variation in the first stage parameters (see Wooldridge, 2002, p. 116). To assume that parental propensity in investing in a child's education (estimated from model (11)) is not correlated with any unobservable factor in model (12) requires that the indicator of parental interest in a child's education should not be driven by the child's schooling performance (otherwise some unobservable factors affecting individuals' ultimate educational attainment, very likely also correlated with their schooling performance, will cause inconsistent estimates of γ). This is ensured by taking the information on parental interest in a child's education in the early childhood, ideally when children have just started school (say 6-7 years old). In other words, the time lag between t and $t+1$ should be large enough to ensure that $\widehat{e}_{k,t}^i$ is exogenous with respect to $\pi_{n,t+1}^{ij}$.

5 Data

5.1 Data set

Our empirical analysis is based on data from the National Child Development Study (NCDS). It is a longitudinal survey that follows 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 Census ward 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 available for 1974 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) 1971 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 neighborhood measure. 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).

As 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. Considering only individuals that did not move out of local area since child's birth (derived from NCDS sweeps one and three) and without missing values in our target variables, our final sample is of 3,477 children and 5,871 parents in 1,689 wards.

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 parent's interest in child's education. The ideal variable to measure parent's effort would have been the number of hours spent investing in child's education (reading to the child, meeting teachers, etc...). Unfortunately, this variable is not directly available in the NCDS. However, the NCDS provides qualitative

information on parental interest in child’s education. In particular, the school questionnaire of different NCDS sweeps (1965, 1969, 1974) contains the following question: “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”, and it is stated to ring the appropriate definition (only one for each parent). We exclude the mothers and fathers answering “Can’t say or inapplicable” category. The interesting feature of this question is that it is not answered by the parents themselves but by the children’s teachers and headmasters. Because in our theoretical model we focus on the impact of the parent’s effort on his/her child’s education, we use the NCDS sweep one (age 7 in 1965), i.e. when the child has just started school and has not yet be “educated”. This choice should ensure that this indicator of parental interest in a child’s education is not driven by the child’s schooling performance, thus solving possible endogeneity issues in the second stage of our empirical strategy (see the end of Section 4). Indeed, the parents’ interest in his/her child education at that early stage as perceived by headmasters (based on parent’s telephone calls asking information about courses and textbooks, frequency of parental visits to the school, meetings with teachers, etc...) should closely proxy parental care (in the sense used in the model, namely the amount of time spent caring for a child’s education) independently from children future schooling achievements.¹⁵

For robustness check, we also report the results obtained by considering an alternative proxy based on parental self-reported information on the frequency they read to their child, also 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”. Here also we exclude the mothers and fathers who answer “Don’t know or inapplicable”.

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 or above qualifications and low-educated individuals otherwise. The information on the parents’ education is derived from

¹⁵The use of other indicators of parental interest from later sweeps (including a continuous variable reporting the number of parents/teachers meetings from NCDS third sweep) does not anyway change qualitatively our results. We also construct a combined variable from different indicators obtained using a factor analysis (reliability assessed by Cronbach’s alpha), but no qualitatively different results are obtained.

the age the parents left school, which is reported in 1974 (NCDS sweep three). Consistently with the aggregation used for cohort's member education, we define high-educated parents the ones that left school at an age greater than 18 years and low-skilled otherwise. So, for example, $\pi_{n,t+1}^{hh}$ is the probability that an adult n of 33 years old in $t + 1 = 1991$, whose at least one of his/her parent has a A-level or above degree, has him(her)self at least a A-level.

Finally, as a proxy for the average neighborhood education quality 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.

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 13-16 (in 1971 – 1974) and the child's education attainment when he/she is 33 years old (in 1991). Precise definitions of these variables and of all the others used in the empirical analysis can be found in Appendix 1.¹⁶ Table 1 contains the summary statistics.

6 Empirical 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 result (i) of Proposition 1. Indeed, as suggested by Proposition 1 (i), Table 2 shows 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 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 75 percent of them are high-educated while only 25 percent are low-educated parents.

Let us now focus on the estimation results of models (11) and (12). Firstly, we provide evidence of whether or not the educational composition of a neighborhood affects high-skilled and low-skilled parents' involvement in their child'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 and 5, tests of Proposition 2 and 3).

¹⁶All 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.

6.1 Neighborhood quality and parental’s behavior

Table 3 reports estimated marginal effects at the sample means and t -statistics (in parentheses) of our proxy for residential community human capital (i.e. percentage of high-skilled population) based on the ordered probit estimation of model (11). It uses alternative dependent variables and different selections of control variables.¹⁷ Column (1) displays the results controlling only for child and neighborhood characteristics. Column (2) adds parental background information and Column (3) includes also controls for the quality and type of the school attended by the child.¹⁸

Observe that the estimation results are qualitatively the same if one uses as the dependent variable either the level of parental interest of a child’s education or the frequency a parent reads to the child. We will thus now focus on the first measure. As stated above, we run different regressions for high- and low-educated parents.

For high-educated parents, all regressions show a positive and statistically significant effects of neighborhood quality of parental interest in 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. However, when moving from Column (1) to Column (3), that is where more control variables on parental background and school characteristics are included, the magnitude of the marginal effect decreases from 0.34 to 0.10. This provides some evidence that neighborhood variables may proxy for unobservable family and school characteristics. Nevertheless, we still find that an incremental change in the proportion of high skilled people living nearby leads to a significant 0.10 increase in the propensity of a parent to increase her/his level of interest in her/his child for an average parent.

For low-educated parents, the coefficients are always positive but no significant (only at the 10 percent level in Column (1)). 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.

The evidence collected so far (Table 2 and 3) seems to indicate that for high-educated parents the child’s educational attainment may be affected by both the neighborhood quality and parents’ involvement while, for low-educated parents, only the neighborhood quality may

¹⁷We focus on the estimated impact of our target variable across different specifications, between skill groups and using alternative dependent variables. For ease of brevity, we do not comments the results related to our control variables, that are anyway in line with the expectations. The complete table with all the control variables is available upon request. Different sets of control variables have also been used, but the qualitative results remain unchanged.

¹⁸Regional dummies as well as a dummy variable indicating if the parent is a female are included in all specifications.

play a role. In the next section, we investigate further these issues by providing a test of Proposition 2 (stage two of our empirical strategy).

6.2 Neighborhood quality, parental behavior and child’s education attainment

Let us now focus on the test of the implications of the model with respect of the expected school achievement of each individual. For high-educated parents, the key feature of the intergenerational education transmission mechanism proposed in Section 3 is that both socialization *inside* the family and socialization *outside* the family (neighborhood quality) plays an important role in the education process of children. This implies that the residential neighborhood *directly* affects the expected school achievement of children and also *indirectly* through its influence on parental effort in a child’s education. For low-educated parents, we might expect a significant effect of the neighborhood quality on children’s educational attainment only.

Table 4 and 5 report marginal effects at the sample means and t -statistics (in parentheses) of our proxy for parental interest in child’s education and residential community human capital based on the probit estimation of model (12).¹⁹ These tables have the same structure as Table 3, i.e. they contain in the different columns the estimates obtained using the different selections of control variables considered in Table 3.²⁰ The tables show four sets of results, one for each transition probability π^{ij} described in Proposition 2.²¹ Table 4 reports the estimated effects of parental interest, neighborhood human quality and their interaction

¹⁹The results obtained using our alternative proxy of parental care in a child’s education (first stage predictions using the frequency parents read to child) are qualitatively the same, thus not discussed and not reported here for brevity.

²⁰Note 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 (that may be passed genetically by parents, see Plug and Vijverberg, 2003). 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.

²¹Also in this case, we focus our attention on the estimated impact of our target variables. The complete set of results for all control variables in the four probit estimations in the different models specifications, and for the two skills groups are not reported for brevity. They are in line with the expectations, and available upon request. Also in this case, different sets of control variables have been used, but the qualitative results remain unchanged.

term whereas Table 5 contains the results after removing the interaction term from the model specification. Comparing Table 4 and Table 5 it can be seen that the impact of the interaction term (estimate of ρ in model (12)) is never significant in any specification and that the effects of the other variables are almost unchanged. This indicates that the effect of parental interest on children’s education attainment does not vary with neighborhood human quality. Thus, we focus our attention on the results contained in Table 5.

Let us start with high-educated parents and thus focus on π^{hh} and π^{hl} . All coefficients are significant both for parental interest and neighborhood quality. However, the sign is exactly the opposite, that is always positive for π^{hh} and 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*. This suggests that, for high-educated parents, both parents’ involvement and neighborhood quality play a crucial role in the intergenerational transmission of education. If one compares the coefficients, then one can see that the magnitude of the effect is higher for parental interest than for neighborhood’s quality. Indeed, when all the controls are considered (Column (3)), for the quality of the neighborhood, the increase in the average probability of a successful transmission of the parental level of education is only of 0.81 percent, and the decrease in the average probability of an unsuccessful transmission is only 1.75 percent. For parental interest, the figures are much higher, that is 13.11 percent and 9.35 percent, respectively. This indicates that the latter effect is more potent than the former one.

Concerning low-educated parents, i.e. π^{lh} and π^{ll} , only the quality of the neighborhood has a significant impact 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 either low-educated parents spend time with their offspring but are not efficient or do not prioritize education and thus are not that interested in the educational attainment of their children. Concerning the influence of the neighborhood, we obtain the expected signs, that is positive for π^{lh} and negative for π^{ll} . In words, a better quality neighborhood has a positive impact on the chance to become educated and a negative impact on the chance to be uneducated. Observe that, when all the controls are considered (Column (3)), when the quality of the neighborhood increases marginally, the decrease in the average probability that a child remains low educated is 2.62 percent and the increase in the average probability a child will end up being highly educated is 5.16 percent. Thus the effect is larger (roughly twice as much) for π^{lh} than for π^{ll} , which means that the percentage of high-skilled workers

in a neighborhood plays a major role in determining the chance to become educated for a child whose parents are not 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 does not seem to play any role either because they are not efficient or because they do not prioritize education.

6.3 Endogeneity issues and further robustness checks

The assessment of the effects of neighborhood level variables on individual outcomes is typically characterized by endogeneity issues arising from individuals' residential location choices, that render the identification and the measurement of these effects problematic. Observe however the slightly different approach of our analysis. Our emphasis is not on the measurement of the size of the neighborhood effects on individual outcomes but rather in the search for evidence consistent with the predictions of the model, that is with the existence of the proposed channel mediating these effects. We look at the mechanism translating neighborhood characteristics into individual outcomes, once residential location choices are made. Specifically, our question is the following. Given the neighborhood parents choose or can afford to live, how is their behavior (in terms of interest in their offspring education) affected by the quality of their residential community? Do they leave to the residential community the main role in educating their children (cultural substitution) or will they put more effort to educate their children the more educated is their residential community (cultural complementarity)?

If parents' concerns about their children education is driving location choices, then high-income and high-educated parents may choose higher educational status neighborhood on the basis of the anticipated effects on their child's education attainment. However, at this stage, two scenarios may take place. On the one hand, anticipating that their children will attain a high level of education, parents may not invest much time in their children education; alternatively, parental effort in a child's education may be higher the higher is the education attainment of the local community. The underlying not obvious education transmission mechanism is the research question addressed in this paper. Thus, the model treats the distribution of agents by neighborhood as fixed and consequently the empirical analysis considers the educational attainment of the local community as exogenous variable with respect to parents' interest in their child education.

Nevertheless, in order to exclude the possibility that the endogeneity of moving decisions²² may cause some unobservable heterogeneity that could drive our results, we address this issue in our empirical analysis. We test for exogeneity of the neighborhood’s human capital level with respect to parental interest in a child’s education by using a two-step instrumental variables methodology with a first stage linear regression and a second stage ordered probit estimation (procedure described in Wooldridge, 2002, p. 474). This procedure simply implies to estimate an augmented version of model (11), that is

$$e_{z,k,t}^i = \alpha q_{k,t} + \sum_{m=1}^M \beta_m x_{m,t} + \rho \hat{v}_k + \varepsilon_{z,t} \quad z = 1, \dots, Z$$

where \hat{v}_k are the residual of a regression of $q_{k,t}$ on a set of exogenous variables that contains the ones in $x_{m,t}$ (for $m = 1, \dots, M$) as a subset. Using the standard probit t -test, a failure to reject the hypothesis $H_0 : \rho = 0$ indicates that $q_{k,t}$ can be taken as exogenous variable. In our case, we obtain a t -test statistic for ρ equal to 1.21, suggesting that the possible endogenous sorting of parents into neighborhoods should not be a major concern in our analysis. In other words, we can assume that parents’ residential location decisions are not essentially driven by their interest in their children future education attainment.

This evidence crucially depends on the chosen sources of exogenous variation when estimating \hat{v}_k , that is on the chosen instruments for $q_{k,t}$. We use as instrument for the average neighborhood education level, the neighborhood population age structure. The motivation is as follows. The UK labor force is characterized by younger cohorts having a higher level of education than older ones.²³ Thus, the proportion of young persons living in the neighborhood is correlated with the proportion of highly qualified persons in the neighborhood but it is not expected to be correlated with any unobserved factor that affects a parent’s interest in his/her child’s education.

Let us turn our attention to the robustness checks of the test of Proposition 2. In order to increase the confidence in our empirical results, we estimate model (12) using a different dependent variable. We take individuals’s years of completed schooling as a proxy for individual’s education achievement (in logs). The OLS results of model (12) estimated separately for high-educated and low-educated parents remain in line with the results displayed in Tables 4 and 5 and thus with the predictions of the theoretical model . Indeed, the estimated effect of parental interest in a child education is found to be positive and significant for children whose parents are high-skilled whereas it is positive but not significant for children

²²Note that we select in our sample only families that do not move after the child’s birth, but it is not possible to control for any previous residential change related to the child’s birth.

²³See in particular Office of National Statistics (2003).

whose parents are low skilled. The average educational attainment of the local community is found to have a positive and significant effect in both cases. The interaction term has no significant effect. These results remain qualitatively unchanged if the average educational attainment of the local community is instrumented using the neighborhood population age structure.

It is worthwhile noting that we do not put much emphasis on the size of the estimated effects. Clearly our empirical analysis suffers from parametric assumptions and the results may not be conclusive. Once more, the attention in this paper is focussed on the attempt to propose a microfoundation of neighborhood effects. The scope of our empirical investigation is to provide some evidence in support of the predictions of the theoretical model.

7 Conclusion

This paper has proposed a microeconomic mechanism of neighborhood effects on educational attainment based on parents' involvement in education. In line with the general consensus that neighborhoods determine only a small proportion of the variation in individual outcomes, we find that average neighborhood human capital quality does influence individual education attainments, but the direct effect is quite limited. The contribution of this paper to the existing literature consists in the finding of a significant neighborhood effect on high-educated parents' interest in their children's education, which in turn play an important role in determining children's schooling achievement. An interesting result in the mechanism proposed here is precisely in the fact that the effects are different according to the level of education of parents. The empirical evidence supports these predictions. Indeed, children's educational attainment of high-educated parents appear to be influenced by both parents' involvement and neighborhood quality while, for low-educated parents, only the neighborhood quality matters.

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 (1998), “On the cultural transmission of preferences for social status”, *Journal of Public Economics* 70, 75-97.
- [7] Bisin, A. and T. Verdier (2000a), “Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits”, *Quarterly Journal of Economics* 115, 955-988.
- [8] Bisin, A. and T. Verdier (2000b), “On the cultural transmission of preferences for social status”, *European Journal of Political Economy* 16, 5-29.
- [9] Bisin, A. and T. Verdier (2001), “The economics of cultural transmission and the dynamics of preferences”, *Journal of Economic Theory* 97, 298-319.
- [10] Boyd, R. and P. Richerson (1985), *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.
- [11] Cavalli-Sforza, L. and M. Feldman (1981), *Cultural Transmission and Evolution. A Quantitative Approach*, Princeton: Princeton University Press.
- [12] 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.
- [13] 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.
- [14] De Bartolome, C. (1990), “Equilibrium and inefficiency in a community model with peer group effects,” *Journal of Political Economy* 98, 110-133.
- [15] Durlauf, S.E. (2004), “Neighborhood effects,” in V. Henderson and J-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol.4, Amsterdam: Elsevier Science.
- [16] Ermisch, J. and M. Francesconi (2001), “Family matters: Impacts of family background on educational attainments,” *Economica* 68, 137-156.

- [17] Hanushek, E. (2002), "Publicly provided education," NBER Working Paper No. 8799.
- [18] 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.
- [19] Hauk, E. and M. Saez-Marti (2002), "On the cultural transmission of corruption", *Journal of Economic Theory* 107, 311-335.
- [20] Johnson, D.J. (1996), "Father presence matters: A review of the literature," National Center of Fathers and Families, Graduate School of Education, University of Pennsylvania.
- [21] 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.
- [22] Leibowitz, A. (1974), "Home investments in children," *Journal of Political Economy* 82, S111-S131.
- [23] Office of National Statistics (2003), "Trends in education and skills," Department for Education and Skills, London, <http://www.dfes.gov.uk/trends>.
- [24] 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.
- [25] Sacerdote, B. (2002), "The nature and nurture of economic outcomes," *American Economic Review Papers and Proceedings* 92, 344-348.
- [26] Saez-Marti, M. and Y. Zenou (2004), "Cultural transmission and discrimination," Unpublished manuscript.
- [27] Wilson, W. J. (1987), *The Truly Disadvantaged*, Chicago: University of Chicago Press.
- [28] Wooldridge, J. (2002), *The Econometric Analysis of Cross Section and Panel Data*, Cambridge: MIT Press.

Appendix 1: Description of variables²⁴

Cohort's members 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 or hearing, delicate, etc...). Source: NCDS1

birth order: child's position in birth order (with respect to brothers and sisters). Source: NCDS2

arithmetic test score: child's age-7 arithmetic test scores. Source: NCDS1

reading test score: child's age-7 reading test scores. Source: NCDS1

schooling: years of completed schooling (derived variable from months of completed schooling up to 1981 and monthly activity information up to 1991. Source: NCDS4 and NCDS5

School variables

school quality: proportion of boys or girls studying for GCE and SCE O-levels²⁵ in the school attended by the child at age 16. Source: NCDS3

school composition: percentage of children with non-manual fathers in the school attended by the child at age 16 (9 bands, 8 dummies). Source: NCDS3

school type: type of school attended by the child at age 16: private, grammar, secondary modern, comprehensive, (4 types, 3 dummies). Source: NCDS3

Family background variables²⁶

female: dummy variable taking value one if the parent is female. Source: NCDS1

parental interest: variable taking value of 4 if the parent is over concerned about the child's education progress and/or expecting too high a standard, 3 if very interested, 2 if he/she shows some interest and 1 if she/he shows little or no interest. Source: NCDS1.

parent reads to child: variable taking value of 3 if the parent reads to the child at least every week, 2 if she/he reads to the child only occasionally and 1 if she/he never or hardly never reads to the child. Source: NCDS1

²⁴The main NCDS sweep data source is reported. In order to maintain sample sizes, this information is often complemented with data coming from earlier or later sweeps to replace missing data.

²⁵GCE (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.

²⁶A parent is the mother or the father or a person acting as mother or father respectively. Because in some cases it was not possible to infer precisely (or obtain the relevant information on) who was the mother or father substitute, we exclude these "parents" from our sample selection.

parent born in UK: dummy variable taking value one if the parent is born in Great Britain. Source: NCDS2

parent income: parent's weekly net wage, 12 bands, mid-points of each range considered. Source: NCDS3

parent age: parent's age in 1974. Source: NCDS1

*parent education**: parent's completed years of schooling (derived from age left full-time education). Source: NCDS3

*high educated parent**: high educated parent's completed years of schooling (derived from age left full-time education greater than 18). Source: NCDS3

*low educated parent**: high educated parent's completed years of schooling (derived from age left full-time smaller or equal to 18). Source: NCDS3

parent employed: dummy variable taking value one if the parent is working. Source: NCDS3

parent social class: social class of parent (or person acting as that parent): professional, intermediate status, skilled non-manual, skilled manual, semi-skilled manual, semi-skilled non-manual, unskilled (7 classes, 6 dummies). Source: NCDS2

parent health problems: dummy variable taking value one if the parent suffered from chronic or serious disability or ill-health, including any hospital in patient admission of 2 weeks or longer. Source: NCDS2

single parent: dummy variable taking value one if there is no regular father figure or there is no natural mother. Source: NCDS2

household financial problems: dummy variable taking value one if the family experienced financial difficulties. Source: NCDS2

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). Source: NCDS2

council house: dummy variable taking value one if the family lived in council provided accommodation.²⁷ Source: NCDS3

house size: number of rooms in household accommodation. Source: NCDS3

family size: number of people in household. Source: NCDS3

*Clearly these variables are not included in the regressions. Parent education is used to define high-skilled and low-skilled parents. We report summary statistics for high educated parent and low educated parent only to provide further information about our sample selection.

²⁷These programs are equivalent to the housing projects in the US.

Residential neighborhood variables (Census Ward level variables or otherwise specified)

high skilled population: proportion of over-18s persons with A-levels (highest grade at age 16 exams) or above qualifications. Source: Census 1971

young population: proportion of persons aged less than 21. Source: Census 1971

total population: total population present. Source: Census 1971

unemployment rate: unemployed over economically active population. Source: Census 1971

activity rate: economically active population (aged more than 15) over present population. Source: Census 1971

activity rate (males): economically active males (aged more than 15) over present population aged more than 15. Source: Census 1971

activity rate (females): economically active females (aged more than 15) over present population aged more than 15. Source: Census 1971

women working hours rate: proportion of women aged less than 45 working weekly less than 8 hours. Source: Census 1971

agriculture employment: proportion of workers in agriculture employment. Source: Census 1971 (Local Authority)

mining and manufacturing employment: proportion of workers in mining and manufacturing employment. Source: Census 1971 (Local Authority)

professional employment: professional and managerial employees over active population. Source: Census 1971 (Local Authority)

unskilled manual employment: unskilled manual employees over active population. Source: Census 1971 (Local Authority)

amenities: proportion of households lacking or sharing hot water and/or inside toilet and/or bath. Source: Census 1971

rooms per household: total rooms over total households. Source: Census 1971

persons per room: proportion of households having 1.5 or more persons per room. Source: Census 1971

lone parent families: proportion of families with children head by a single parent. Source: Census 1971

council housing: proportion of households residing in council houses. Source: Census 1971

car access: proportion of households with no car. Source: Census 1971

distance to jobs: proportion of workers travelling to jobs by train. Source: Census 1971

Table 1: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
special education	3477	0.24	0.43	0	1
birth order	3477	1.42	1.85	1	10
arithmetic test score	3477	5.12	2.49	0	10
reading test score	3477	23.39	7.11	0	30
schooling	3477	12.54	2.002	9.55	20.56
school quality	3477	0.24	0.32	0	1
school composition	3477	4.68	2.41	1	9
school type	3477	3.01	1.29	1	4
parent interest	5981	1.77	0.12	1	4
parent read to child	5981	1.34	0.16	1	3
parent education	5981	14.24	2.37	6	25
high educated parent	3244	19.04	2.68	18	25
low educated parent	2747	9.01	2.41	6	18
parent wage	5981	24,001	11,112	2	69
parent age	5981	43.47	8.98	32	78
parent social class	5981	3.92	1.55	1	7
parent employed	5981	0.79	0.38	0	1
parent born in UK	5981	0.69	0.46	0	1
female	5981	0.51	0.50	6	25
household income	5981	43,624	21,321	4	167
single mother	5981	0.003	0.055	0	1
parent health problems	5981	0.05	0.22	0	1
household health problems	5981	0.09	0.16	0	1
household financial problems	5981	0.08	0.24	0	1
council house	5981	0.18	0.49	0	1
house size	5981	4.93	1.45	1	35
family size	5981	4.88	1.64	0	17

high skilled population	1689	0.13	0.04	0.03	0.69
young population	1689	0.31	0.10	0.09	0.41
total population (thousands)	1689	59.79	73.29	7.08	267.44
unemployment rate	1689	0.05	0.03	0.02	0.18
activity rate	1689	0.61	0.04	0.30	0.72
activity rate (males)	1689	0.82	0.04	0.55	0.91
activity rate (females)	1689	0.43	0.06	0.23	0.59
women working hours rate	1689	0.08	0.15	0.05	0.25
professional employment	1689	0.14	0.08	0.03	0.56
unskilled manual employment	1689	0.06	0.04	0.007	0.17
agriculture employment	1689	0.03	0.05	0.00	0.64
mining and manufacturing employment	1689	0.33	0.16	0.00	0.75
lone parent families	1689	0.04	0.01	0.004	0.09
amenities	1689	0.17	0.098	0.00	0.46
persons per room	1689	0.02	0.02	0.001	0.15
rooms per household	1689	4.10	0.58	2.05	7.86
council housing	1689	0.29	0.35	0	0.79
distance to jobs	1689	0.39	0.25	0	0.70
car access	1689	0.25	0.46	0.09	0.76

Table 2: Parental interest by education groups

Variable*	high educated parents	low educated parents
1) parental interest age 7	63%	37%
2) parental interest age 11	66%	34%
3) parents read to child age 7	59%	41%
4) parental initiative to discuss child age 7	75%	25%
5) parental initiative to discuss child age 11	71%	29%
6) parents/teachers discussions age 16	81%	19%
7) parental anxiety education age 16	69%	31%

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

1)-2) NCDS1 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) NCDS3 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 1.39 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: Model (11) ordered probit estimates

	Column (1)	Column (2)	Column (3)
dependent variable: parental interest			
<i>high educated parents</i>			
high skilled population	0.3451*** (3.03)	0.1429*** (3.12)	0.1021*** (2.92)
<i>low educated parents</i>			
high skilled population	0.2301* (1.75)	0.1202 (1.12)	0.0829 (0.92)
dependent variable: parent reads to child			
<i>high educated parents</i>			
high skilled population	0.2813*** (2.65)	0.1333** (2.53)	0.1075** (2.36)
<i>low educated parents</i>			
high skilled population	0.1282 (0.99)	0.0986 (0.97)	0.0151 (0.46)
Control set:			
child variables	yes	yes	yes
neighborhood variables	yes	yes	yes
family background variables	no	no	yes
schooling variables	no	no	no
regional dummies	yes	yes	yes
female dummy	yes	yes	yes
Notes:			
- precise definitions of control variables by groups in Appendix 1			
- marginal effects at the sample means and t-statistics (in parentheses) based on Huber-White robust standard errors are reported			
- coefficients marked with one (two) [three] asterisks are significant at 10 (5) [1] percent level			
- n. of obs. 5,981, pseudo R^2 in the range 0.2122 - 0.2511			

Table 4: Model (12) probit estimates

	Column (1)		Column (2)		Column (3)	
	π^{hh}					
(predicted) parental interest	0.2012***	(3.86)	0.1487***	(4.01)	0.1301***	(3.49)
high skilled population	0.0666**	(2.14)	0.0944**	(2.10)	0.0080**	(2.05)
high skilled population times parental interest	0.1022	(0.80)	0.0884	(0.76)	0.0027	(0.77)
	π^{ll}					
(predicted) parental interest	-0.0800	(0.96)	-0.0699	(0.93)	-0.0199	(0.91)
high skilled population	-0.0776***	(2.89)	-0.0443**	(2.23)	-0.0231**	(2.44)
high skilled population times parental interest	-0.0452	(1.02)	-0.0234	(0.94)	-0.0211	(0.52)
	π^{hl}					
(predicted) parental interest	-0.1801***	(2.99)	-0.1403**	(2.19)	-0.0889**	(2.05)
high skilled population	-0.1002**	(2.12)	-0.0778**	(2.06)	-0.0101**	(2.03)
high skilled population times parental interest	-0.0542	(0.92)	-0.0204	(0.70)	-0.0098	(0.61)
	π^{lh}					
(predicted) parental interest	0.0009	(1.01)	0.0019	(0.80)	0.0007	(0.75)
high skilled population	0.0989**	(2.10)	0.0801**	(2.06)	0.0500**	(2.05)
high skilled population times parental interest	0.0612	(0.87)	0.0422	(0.78)	0.0111	(0.59)
Control set:						
child variables	yes	yes	yes	yes	yes	yes
neighborhood variables	yes	yes	yes	yes	yes	yes
family background variables	no	no	yes	yes	yes	yes
schooling variables	no	no	no	no	yes	yes
regional dummies	yes	yes	yes	yes	yes	yes
female dummy	yes	yes	yes	yes	yes	yes

Notes: - precise definitions of control variables by groups in Appendix 1

- marginal effects at the sample means and t -statistics (in parentheses) based on Huber-White robust standard errors are reported

- coefficients marked with one (two) [three] asterisks are significant at 10 (5) [1] percent level

- n. of obs. 3,477, pseudo R^2 in the range 0.3799 - 0.4314

Table 5: Model (12) probit estimates

	Column (1)		Column (2)		Column (3)	
			π^{hh}			
(predicted) parental interest	0.2191***	(4.04)	0.1601***	(4.12)	0.1311***	(3.88)
high skilled population	0.0891**	(2.39)	0.0108**	(2.13)	0.0081**	(2.10)
			π^{ll}			
(predicted) parental interest	-0.0811	(1.02)	-0.0721	(1.12)	-0.0209	(0.96)
high skilled population	-0.0791***	(2.92)	-0.0481**	(2.44)	-0.0262**	(2.52)
			π^{hl}			
(predicted) parental interest	-0.1863***	(3.02)	-0.1496**	(2.22)	-0.0935**	(2.12)
high skilled population	-0.1019	(2.15)	-0.0802	(2.03)	-0.0175	(2.26)
			π^{lh}			
(predicted) parental interest	0.0012	(1.15)	0.0021	(0.88)	0.0011	(0.76)
high skilled population	0.1081**	(2.25)	0.0826**	(2.17)	0.0516**	(2.15)
<hr/>						
Control set:						
child variables	yes	yes	yes	yes	yes	yes
neighborhood variables	yes	yes	yes	yes	yes	yes
family background variables	no	no	yes	yes	yes	yes
schooling variables	no	no	no	no	yes	yes
regional dummies	yes	yes	yes	yes	yes	yes
female dummy	yes	yes	yes	yes	yes	yes

Notes:

- precise definitions of control variables by groups in Appendix 1
- marginal effects at the sample means and t -statistics (in parentheses) based on Huber-White robust standard errors are reported
- coefficients marked with one (two) [three] asterisks are significant at 10 (5) [1] percent level
- n. of obs. 3,477, pseudo R^2 in the range 0.3754 - 0.4123