DISCUSSION PAPER SERIES

DP11419

CULTURAL TRANSMISSION AND SOCIALIZATION SPILLOVERS IN EDUCATION

Carlo Del Bello, Fabrizio Panebianco, Thierry Verdier and Yves Zenou

LABOUR ECONOMICS

ISSN 0265-8003

CULTURAL TRANSMISSION AND SOCIALIZATION SPILLOVERS IN EDUCATION

Carlo Del Bello, Fabrizio Panebianco, Thierry Verdier and Yves Zenou

Discussion Paper 11419
Published 29 July 2016
Submitted 29 July 2016

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

CULTURAL TRANSMISSION AND SOCIALIZATION SPILLOVERS IN EDUCATION

Abstract

We propose a model of the intergenerational transmission of education where children belong to either high-educated or low-educated families. Children choose the intensity of their social activities while parents decide how much educational effort to exert. We characterize the equilibrium and show under which condition cultural substitution or complementarity emerges. There is cultural substitution (complementarity) if parents decrease (increase) their education effort when their child socializes more with other children of the same type. By structurally estimating our model to the AddHealth data in the United States, we find that there is cultural complementarity for high-educated parents and cultural substitution for low-educated parents. This means that, for both parents, the more their children interact with kids from high-educated families, the more parents exert educational effort. We also perform some policy simulations. We find that policies aiming at mixing high and low educated children perform well in terms of average educational outcomes. We also show that a policy that gives vouchers to children from high-educated families have a positive and significant impact on the educational outcomes of all children while a policy that gives vouchers to children from low-educated families has a negative effect on the outcomes of both groups.

JEL Classification: D85, I21

Keywords: Social Networks, education, homophily, cultural transmission.

Carlo Del Bello - cl.delbello@psemail.eu
PSE

Fabrizio Panebianco - fabrizio.panebianco@unibocconi.it
Bocconi University and IGIER

Thierry Verdier - thierry.verdier@ens.fr
PSE and CEPR

Yves Zenou - yves.zenou@ne.su.se
Monash University and CEPR

Acknowledgements
The authors aknowledge the financial support of ERC 324004 Grant TECTACOM, ERC STRATEMOTIONS - GA 324219 and the Spanish Ministry of Economia y Competitividad under Grant ECO-2014-58297-R.

Cultural Transmission and Socialization Spillovers in Education*

Carlo L. Del Bello† Fabrizio Panebianco‡ Thierry Verdier§ Yves Zenou¶

July 18, 2016

Abstract

We propose a model of the intergenerational transmission of education where children belong to either high-educated or low-educated families. Children choose the intensity of their social activities while parents decide how much educational effort to exert. We characterize the equilibrium and show under which condition cultural substitution or complementarity emerges. There is cultural substitution (complementarity) if parents decrease (increase) their education effort when their child socializes more with other children of the same type. By structurally estimating our model to the AddHealth data in the United States, we find that there is cultural complementarity for high-educated parents and cultural substitution for low-educated parents. This means that, for both parents, the more their children interact with kids from high-educated families, the more parents exert educational effort. We also perform some policy simulations. We find that policies aiming at mixing high and low educated children perform well in terms of average educational outcomes. We also show that a policy that gives vouchers to children from high-educated families have a positive and significant impact on the educational outcomes of all children while a policy that gives vouchers to children from low-educated families has a negative effect on the outcomes of both groups.

Keywords: social networks, education, homophily, cultural transmission.
JEL Classification: D85, I21, Z13.

*The authors aknowledge the financial support of ERC 324004 Grant TECTACOM, ERC STRATEMOTIONS - GA 324219 and the Spanish Ministry of Economia y Competitividad under Grant ECO-2014-58297-R.
†Paris School of Economics, France. Email: cl.delbello@psemail.eu.
‡Bocconi University and IGIER, Italy. Email: fabrizio.panebianco@unibocconi.it.
§Paris School of Economics, France, and CEPR. Email: thierry.verdier@ens.fr.
¶Monash University, Australia, IFN and CEPR. E-mail: yves.zenou@monash.edu.

1 Introduction

Education is a crucial aspect of any policy that could be implemented in a country. Still, it is unclear which policy should be implemented to increase education in the population. In this paper, we investigate the intergenerational transmission of education by studying how parental educational effort and background (vertical transmission) as well as children's social networks (horizontal transmission) have an impact on educational outcomes.

Intergenerational transmission of education is indeed important for studying social mobility and persistence of economic inequalities. For example Black and Devereux (2011) argue that focusing on education makes sense because it is a good measure of lifetime welfare. What is less clear in the literature is why parental education should be important for children education other than an obvious wealth/income channel. Black and Devereux (2011) propose two channels: time allocation and higher productivity in child-enhancing activities. Indeed it is well documented that involvement in child care increases with parental educational attainment. Guryan et al. (2008) use American Time Surveys data to investigate the issue of time allocation in child care. They find a positive relation between parental years of education and child care. Such a relation is robust to all types of childcare activities, whether educational or not. This suggests that it is very difficult to assess how much productivity matters in education. In this paper, not only we explicitly model parental effort in child care, but we are also able to identify empirically an index of parental effort that is specifically related to education-enhancing child care.

To be more precise, we propose a model with two types (or groups), high-educated (type H) and low-educated (type L) and examine a cultural transmission mechanism where children belonging to each group choose a certain level of socialization effort that determines the probability of forming links with other students and their degree of homophily[1] with respect to their own type. In our model, the way two children of different types socialize and their friendship link is formed essentially depend on four key elements: (i) the *preference bias* captured by the preference of interacting with someone from own versus the other type, (ii) the *meeting bias*, captured by the fraction of own type children in the population but also by their geographical distance, (iii) the *synergy* of social interactions between all children (iv) the ex ante *heterogeneity* in their observable characteristics.

Parents decide their education effort by trading off the cost of such an effort and the benefit of having an educated child. What is key and new in this transmission process is that parents decide upon their effort by looking at the (expected) degree of homophily of their children (i.e. how much they form friendship links with children of their own types). We totally characterize the equilibrium of this model and determine under which condition there is cultural complementarity and cultural substitution. There is cultural substitution (complementarity) if parents decrease (increase) their

[1] Homophily is the tendency of agents to associate with other agents who have similar characteristics. It refers to the fairly pervasive observation in working with social networks that having similar characteristics (age, race, religion, profession, education, etc.) is often a strong and significant predictor of two individuals being connected (McPherson et al., 2001).

education effort when their child socializes more with other children of the same type.

We then structurally estimate our model using the AddHealth data, which study the impact of the social environment (i.e. friends, family, neighborhood and school) on adolescents' behavior in the United States. We find that there is *cultural complementarity* for type$-H$ parents (i.e. high-educated parents exert more educational effort the more their children are homophilous) and *cultural substituability* for type$-L$ parents (i.e. low-educated parents exert more educational effort the more their children are heterophilous).

We then perform some comparative statics exercises with respect to the preference bias of the children and the intensity of the peer effects. We find that, when people of one group become less prejudiced against the other group, they tend to socialize more and thus are more likely to form links with individuals from the other group. As a result, they become less homophilous, which induces high-educated parents to exert less educational effort (cultural complementarity) and low-educated parents to provide more educational effort (cultural substituability). This, in turn, implies that the probability of becoming educated decreases for children from high-educated families while it increases for children from low-educated families.

When the intensity of peer effects increases (i.e. there are more synergies from interacting with other children), children socialize more with each other and thus form more friendship links. Because in our data there is a meeting bias in favor of the children from low-educated families (only 28% of children are from high-educated families), children of type H become less homophilous while those of type L become more homophilous. This leads to the fact that both parents exert less educational effort. As a result, the impact of education outcomes is very small since vertical transmission decreases while the horizontal one increases.

We then look at different policy experiments. First, we consider two-social mixing policies. In the first one, we increase the fraction of children from high-educated families while keeping the total number of children constant. The aim of this policy is to reduce the meeting bias so that children have more chance to meet children from high-educated families (since they only constitute 28% of the population in our dataset). When implementing such a policy, we show that it not only benefits the children from high-educated families but also the ones from low-educated families since they all have a higher chance of becoming more educated. This is because children from high-educated families become more homophilous while those from low-educated families become more heterophilous. This, in turn, raises the educational effort of both types of parents.

In the second social-mixing policy, we reduce the geographical segregation (or distance) between the two types of children. We show that such a policy favors children from low-educated families but harms those from high-educated families. This is because both types of children become less homophilous, which, in turn, reduce the educational effort of type$-H$ parents (cultural complementarity) but increase the educational effort of type$-L$ parents (cultural substituability).

Second, we also examine voucher policies, which consist in giving (monetary or non-monetary) incentives to children for socializing more in after-school activities. In the first voucher policy, we only give vouchers to children from high-educated families while, in the second one, we only give vouchers to children from low-educated families. We show that the former has a positive impact on the education outcomes of both types of children while the latter has a negative impact on the education outcomes of both types of children. In fact, if we increase the voucher by 83% in each of these two policies, we show that the difference in the probability of becoming high-educated is as large as 23% for type$-H$ children and 17.5% for type$-L$ children.

The rest of the paper unfolds as follows. In Section 2, we relate our paper to the relevant literature and in Section 3 we propose some initial evidence from the data motivating our analysis. In Section 4, we describe our model while, in Section 5, we characterize the equilibrium and show under which condition it is unique. In Section 6, we structurally estimate our model using the AddHealth data. Section 7 is devoted to the comparative statics exercises while Section 8 examine the different policy experiments and their impacts on educational outcomes. Finally, Section 9 concludes.

2 Related literature

There is a very large literature on education and, in particular, on why children with different backgrounds end up with different levels of education. Most studies have found that school quality (e.g., Card and Krueger, 1992; Hanushek, 2002), family background (e.g., Ermisch and Francesconi, 2001; Sacerdote, 2002; Plug and Vijverberg, 2003), parental involvement in children's education (Jencks and Phillips, 1998; Patacchini and Zenou, 2009, 2011), and peer and network effects (Sacerdote, 2001, 2011; Calvó-Armengol et al., 2009) have a significant and positive impact on the level of education of children.

In this paper, we propose a different approach by studying the role of parents, geographical distance and children's network of friends in the education outcomes of children. We also study the efficiency of different policies (spatial and social) in this framework.

There is also an important theoretical and empirical literature on cultural transmission initiated by the seminal papers of Bisin and Verdier (2000, 2001).[2] In this literature, cultural transmission is conceptualized as the result of interactions between purposeful socialization decisions inside the family (*direct vertical socialization*) and other socialization processes like social imitation and learning which govern identity formation (*oblique and horizontal socialization*). 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. Cultural traits are endogenous in this context. Allowing for interesting socio-economic effects interacting with the socialization choices of parents, the basic cultural transmission model of Bisin and Verdier has

[2] For an overview, see Bisin and Verdier (2011).

been extended and applied to several different environments and cultural traits and social norms of behavior, from preferences for social status (Bisin and Verdier, 1998), to corruption (Hauk and Sáez-Martí, 2002), hold up problems (Olcina and Penarubbia, 2004), development and social capital (Francois, 2002), intergenerational altruism (Jellal and Wolf, 2002), labor market discrimination and work ethics (Sáez-Martí and Zenou, 2012), globalization and cultural identities (Olivier et al., 2008), education (Botticini and Eckstein, 2005, 2007; Patacchini and Zenou, 2011), identity (Bisin et al., 2011, 2016), leadership (Verdier and Zenou, 2015, 2016) and religion (Bisin et al., 2004; Patacchini and Zenou, 2016).

Our approach brings several innovations with respect to this literature on cultural transmission.

First, contrary to this literature where peer effects are conceived as an average intra-group externality with arbitrary group boundaries and at a quite aggregate level (residential neighborhood, state or even country level), here peers will be defined by the smallest unit of analysis for peer effects, that is the *social network* of social relationships.[3]

Second, in our paper, the *social network is endogenous* and is formed by socialization efforts from children. In other words, we consider a model in which children play an active role in the socialization process by choosing their socializing activity that shapes the distribution of type in their neighborhood. In this respect, we provide one of the very few models with endogenous oblique socialization.[4] The standard literature and, in particular, the literature on discrete cultural traits, considers a population based framework in which, during the horizontal socialization, all children face the same distribution of types in their neighborhood. This can be considered as a *mean-field* approximation of a network approach. The only model that considers the role of a (fixed and exogenous) network is Panebianco and Verdier (2015), which also uses a mean-field approach. In our model, the socialization choice of the children gives rise to a network, which, in turn, defines the degree of homophily of each child.[5]

Third, we have a setting in which children are first and temporarily socialized to the parent trait (early socialization), then children choose socialization effort and, at last, parents exert an educational effort taking into account their offsprings' choice. This allows us to have parental socialization costs also to depend on the homophily choices of their children.

There are other papers that explicitly model network formation and homophily but do not study the intergenerational transmission of a specific trait (such as education here). Currarini et al. (2009, 2010) is a seminal contribution in the economic theory of homophily in networks. As in our paper, they highlight two possible causes of homophily, *meeting bias* and *preference bias*, and structurally estimate their model with the AddHealth data to empirically measure these two biases

[3]The economics of networks is a growing field. For overviews, see Jackson (2008), Blume et al. (2011), Ioannides (2012), Boucher and Fortin (2016), Graham (2015), Jackson and Zenou (2015) and Jackson et al. (2016).

[4]To the best of our knowledge, Vaughan (2013) provides the only model of cultural transmission that attempts to endogenize the horizontal socialization process. However, he does not propose a network approach.

[5]There is also some recent literature (Panebianco, 2014, and Buechel et al., 2014) that studies the impact of network structure in horizontal socialization but the network is exogenous and cultural traits are continuous.

for different racial groups. Golub and Jackson (2010, 2012a,b) propose a random model of network formation and study time to convergence of a learning/behavior-updating process. They show that homophily induces a lower speed of social learning. They also examine how well these theoretical results match the process when it is simulated on empirically observed high school friendship networks (AddHealth data). Finally, using the framework of Cabrales et al. (2011), Albornoz et al. (2014) develop a model that integrates productive and socialization efforts with network choice and parental investments. Individuals not only choose their efforts within a network but also the network to which they belong. Because of complementarities, they show that individuals underinvest in productive and social effort. They also show that there are more people than the socially desirable number in the network whose distribution of types has a larger mean if individual productivities are uniformly distributed.

Compared to this literature, we also propose a simple model of network formation based on socialization efforts of the children. What is new in our framework is the fact that parents make educational effort decisions based on the degree of homophily of their children. We also study social-mixing as well as voucher policies by structurally estimating our model to the AddHealth data.

3 A motivating example from the data

Consider two types of parents: type $t = H$ are parents who are highly educated while type $t = L$ are parents who have a low level of education. Using the AddHealth data (see Section 6.1 below for a description of this dataset), in Table 1, we report the results of a simple linear probability model where the endogeneous variable is the probability of going to college (dummy variable) for a student of type t while the key explanatory variables are the parental effort (how much time parents spend educating their children) when the student is a teenager, the student's homophily (measured by the percentage of friends of the same type) when the student is a teenager and the interaction of these two variables.[6] We also add an array of control variables and school fixed effects to control for school characteristics. We do not claim any causal relationship just simple correlations.

[*Insert Table* 1 *here*]

We can see, in columns (1) and (2) in Table 1, where we only consider parental effort as explanatory variable, that the more parents put effort in educating their children, the higher is the chance for the children to eventually become educated (i.e. go to college). The effect is significant for both high- and low-educated parents, but the magnitude is higher for parents of type $t = L$. Then, when we put together both parental effort and children's homophily as explanatory variables in columns (3) and (4), the effect of parental effort disappears for high-educated parents

[6] All these variables are precisely defined in Section 6.1

but remains, albeit lower, for low-educated parents. Finally, when we add the interaction term between these two variables (columns (5) and (6)), we observe that the education effort of parents has no direct impact on the education level of their children while the children's homophily has a strong impact on their education outcomes for children whose parents are highly educated but no impact if parents have a low level of education.

Table 1 suggests that both parental effort and children's homophily are a key determinant of the education outcomes of children but that the homophily variable erases the impact of parental effort when these two variables are put together in the same regression. This is an interesting suggestive evidence from a simple reduced-form regression.

We would now like to investigate further this issue by understanding more precisely the exact mechanism behind the role of parental effort and children's homophily on the education outcomes of children. For that, we propose a theoretical model where both parental effort and children's homophily are endogenously determined in equilibrium and examine how they interact with each other in explaining children's education outcomes. We will then structurally estimate our model to evaluate the role of each of these variables on the education outcomes of children and study the relative importance of parental effort versus children' homophily choices.

4 The Model

Consider a cultural transmission model with a *two-cultural trait population* of individuals. We build on the model of cultural transmission of Bisin and Verdier (2000, 2001) in which vertical socialization inside the family interacts with horizontal socialization outside the family. Contrary to Bisin and Verdier (2001), we assume that children are active in the socialization process by choosing how much social interactions they have with other children.

To be more specific, define T as the set of possible types of traits in the population. Assume $T = \{H, L\}$, where H refers to "high educated" and L to "low educated". We will mainly study children who have not yet been educated. As a result, when we say that a child has trait $t \in T$, it means that this child has a parent who is of type t. The fraction of children with trait $t = H$ is denoted by q so that $1 - q$ refers to the fraction of children in the population who have trait $t = L$. Families are composed of one parent and a child, and hence reproduction is a-sexual. All children are first and temporarily socialized to the trait of their parent (*early socialization*). Then, children choose their social interactions with other children of different types and a network is formed. This is the homophily stage. Then, parents react to children' choices by deciding the level of education effort τ^i. Interestingly, this allows us to have parental education costs that depend on the homophily choices of children. Finally, children become educated or not, i.e. they become of type H or L.

Timing The timing is as follows. First, children choose how much socialization effort (or activity) they are exerting with other children. This socialization choice is made by each child knowing that it will have an impact on the probability that she will form links with other children. Thus, the choice of socialization activities of all children endogenously determines the expected friendship network structure. This is the network the child will be exposed to during horizontal socialization, if the parent's education effort in transmitting their trait fails.

Then, each parent, knowing this "expected" network,[7] will decide how much education effort to exert in transmitting their trait. This education effort is costly, which has a direct cost but also a cost that depends on the socialization choices made by children. Indeed, parents may have lower costs if their offsprings make socialization choices that are in line with the educational status of the parent. For example, if a child, whose parent is highly educated, decides to mainly interact with children with high-educated parents (*homophily*), then the parent of this child will experience a lower socialization cost. Observe that all parents want their kids to be educated, even the low-educated parents. As a result, they all exert education effort, which helps their kids to become highly educated.

Finally, given the vertical and horizontal transmission mechanisms described above, children will adopt either trait H or L.

Network formation Let us solve the children' socialization choices. We consider a simultaneous move game of network formation (or social interactions) where each child i selects a socialization effort, $s_i \geq 0$. Let $\mathbf{s} = (s_1, ..., s_n)$ be a profile of socialization efforts. Then, children i and j, with types $t = H, L$ interact with a link intensity given by:

$$p_{ij} = \frac{1}{4}\delta_{ij}\rho_{ij}s_i s_j \tag{1}$$

In (1), p_{ij} is the probability of forming the link ij and $\delta_{ij} = 1$ if i and j are of the same type and $\delta_{ij} < 1$ if they are of different types. Furthermore, ρ_{ij} is an index of geographical proximity between any two children i and j and, thus, is symmetric ($\rho_{ij} = \rho_{ji}$); this index decreases with the geographical distance between i and j. Finally, $1/4$ is a normalization scalar that ensures that p_{ij} is always between 0 and 1.[8] In other words, i and j are more likely to form a friendship link, if they are of the same type (captured by δ_{ij}) and if they reside close to each other. Moreover, the higher is both socialization efforts s_i and s_j, the more likely a link will be formed. Observe that δ_{ij} has been called *preference bias* by Currarini et al. (2009) since it says that people of the same type are more likely to form a link than people of different types. For simplicity, we assume that preference bias is homogenous among agents of the same type and take values 1 for children of the

[7]In fact, as we will see below, the parent does not need to know the whole (expected) network. He or she only needs to know the homophily behavior of his/her offspring, which amounts to know the number of friends the child has and how often they interact with each other.

[8]See Appendix A for further details.

same type, i.e. $\delta_{ij} \equiv \delta^{HH} = \delta^{LL} = 1$, and takes values less than one for children of different types, i.e., $\delta_{ij} \equiv \delta^{HL} < 1$ (for a child of type H forming a link with child of type L) and $\delta_{ij} \equiv \delta^{LH} < 1$ (for a child of type L forming a link with child of type H). We leave the possibility that δ^{HL} may be different from δ^{LH}, which means that the network may be asymmetric since p_{ij} may have a distinct value than p_{ji}. In the empirical analysis below, we will show that it is indeed the case in our data.

As in Cabrales et al. (2011), in (1), the exact identity of the interacting partner is not an object of choice. Rather, each child i chooses an aggregate level of socialization effort s_i. This total effort is then distributed across each and every possible bilateral interaction in proportion to the partner's socialization effort. This interaction pattern arises naturally when meetings result from casual encounters rather than from an earmarked socialization process. In our context, children at schools participate to after-school activities (such as dance, music, honors club, foreign language clubs, etc.) and s_i reflects how many activities and how often they practice these activities. Two children who spend a lot of time practicing these after-school activities are then more likely to be friends than those who don't.

Payoffs The utility of child i choosing socialization effort s_i is given by:

$$u_i = b_i s_i + \phi \sum_j p_{ij} - \frac{1}{2} s_i^2 \tag{2}$$

where $\phi > 0$. By plugging p_{ij} from (1) into (3), we obtain:

$$u_i = b_i s_i + \frac{1}{4} \phi \sum_j \delta_{ij} \rho_{ij} s_i s_j - \frac{1}{2} s_i^2 \tag{3}$$

where b_i is the ex ante heterogeneity (observable characteristics such as gender, race, etc.) of individual i. In the utility function (3), the returns from socialization effort s_i for individual i are the sum of a private component $(b_i s_i - \frac{1}{2} s_i^2)$ and a synergistic component $(\frac{1}{4} \phi \sum\limits_j \delta_{ij} \rho_{ij} s_i s_j)$. Quite naturally, we assume strategic complementarities between the different socialization efforts of linked individuals so that:

$$\frac{\partial^2 u_i}{\partial s_i \partial s_j} = \frac{\phi \delta_{ij} \rho_{ij}}{4} \geq 0$$

Parents' education effort Let us now focus on the parent's effort choice. The parents are assumed to know the socialization choice of their own child and that of her friends and thus the resulting values of p_{ij}^*, between their child i and her friends j.[9] For each parent of a child i, denote by $d_i^* = \sum_j p_{ij}^*$ the expected degree of her child i, which is basically the number of friends of i multiplied by probability that each pair forms a link. Denote also by p_{ij}^{t*} the probability of forming

[9] All variables with a star refer to equilibrium or choice variables.

a link between children i and j when $t_i = t_j$ (i.e. i and j are of the same type). For example, p_{ij}^{H*} is the (equilibrium) probability that a child i of type H (which means here that her parent is of type H) forms a link with a child j of type H. Define $\widehat{p}_{ij}^{t*} \equiv p_{ij}^{t*}/d_i^*$, for trait $t = H, L$. Then, for an individual i of type t,

$$h_i^{t*} \equiv \sum_j \widehat{p}_{ij}^{t*} = \sum_j \frac{p_{ij}^{t*}}{d_i^*} \tag{4}$$

is a measure of *homophily* of individual i of type t in the expected equilibrium network. Indeed, h_i^{t*} measures the homophily of child i since it divides the (expected) number of links with own type t by the (expected) total number of links of individual i. Let us now consider, in turn, the choice of the parents of type H and, then, of type L.

Consider a parent and a child of type H. Denote by π_i^{HH} the probability that a child i from a parent of type H becomes of type H when adult (i.e. the child becomes high educated). Similarly, denote by π_i^{HL} the probability that a child i from a parent of type H becomes of type L. The education mechanism described above can then be characterized by the following transition probabilities:

$$\pi_i^{HH} = \tau_i^H + (1 - \tau_i^H)h_i^{H*} \tag{5}$$

$$\pi_i^{HL} = (1 - \tau_i^H)(1 - h_i^{H*}) \tag{6}$$

where $0 \leq \tau_i^H \leq 1$ is the socialization effort of a type$-H$ parent who has a child i. It is also the probability for which *direct vertical socialization* to the parent's trait ($t = H$) occurs. Consider equation (5). The child i will be socialized to trait H if either the direct socialization from her parent H succeeds (this occurs with probability τ_i^H) or, if it does not succeed (this occurs with probability $1 - \tau_i^H$), she is then subject to *horizontal socialization*, which is here endogenous and determined by the network of social interactions described above. To be more precise, it is determined by h_i^{H*}, the degree of homophily of child i or more exactly the (expected) fraction of child i's friends who are of type H. The more the child has (expected) friends of type H, the more likely she will be of type H. Observe that, instead of considering the average population with trait H as in Bisin and Verdier (2000), here we look at the average population share with a specific level of homophily in *each* child's environment. This has the striking implication of preventing us to have a unique equation representing the whole set of agents in a type. This is why we need to index each transition probability by i since it depends on the social behavior of the child (homophily) and not on the average population with trait t. In this respect, the Bisin-Verdier model can be seen as a mean-field approximation of this process, with the additional simplification of network exogeneity.

We can now define the utility function for a parent i of type H as follows:

$$U_i^H = \underbrace{\left(\pi_i^{HH} V^{HH} + \pi_i^{HL} V^{HL}\right) - \frac{1}{2}\left(\tau_i^H\right)^2}_{\text{Bisin-Verdier}} + \underbrace{\alpha^H h_i^{H*} \tau_i^H}_{\text{Reciprocate Homophily}} \tag{7}$$

where h_i^{H*} is the homophily of child i (defined in (4)) that arises from the endogenous network formation and where $V^{HH} > 0$ and $V^{HL} > 0$ denote the utility a type$-H$ parent derives from having a child of type H and L, respectively. Quite naturally, we assume that $V^{HH} > V^{HL}$ so that $\Delta V^H \equiv V^{HH} - V^{HL} > 0$, i.e. there is a positive utility associated with having a high educated child for a high educated parent.

The utility (7) is composed of two parts: (i) the standard utility function used in the Bisin and Verdier framework, which depends on the benefits and costs of socialization, (ii) a new part, which we refer to as "reciprocate homophily ". To be more precise, when the parent considers h_i^{H*}, she infers whether her child tends to interact more with other children of the same type as herself (homophily) or with children from the other type (heterophily). Then, we assume that the higher is h_i^{H*}, i.e. the more homophilous is the child, the higher is the (marginal) utility of her parent of exerting education effort, i.e. $\frac{\partial^2 U_i^H}{\partial \tau_i^H \partial h_i^{H*}} = \alpha^H > 0$. If h_i^{H*} is low, then there is a conflict between the child' choice and the parent's preferences and thus the more heterophilous is the child, the less utility her high-educated parent derives from exerting education effort. This is a reasonable assumption, which is in fact verified in our data since we will show below that $\alpha^H > 0$ (see Table 3).

The problem for low-educated parents is slightly different because of opposite incentives with respect to parents of type H in the education effort choices. Indeed, even low-educated parents want their child to be educated and thus exert effort in this direction. We thus assume that parents of type L would like to have children of type H. Therefore, their socialization effort is exerted in order to increase the probability that their child becomes of type H. The transition probabilities can then be written as follows:

$$\pi_i^{LH} = \tau_i^L + (1 - \tau_i^L)(1 - h_i^{L*}) \tag{8}$$

$$\pi_i^{LL} = (1 - \tau_i^L) h_i^{L*} \tag{9}$$

The first equation states that parents of type L exert an effort τ_i^L to induce their kids to become of type H. If they are unsuccessful, the probability that their children will become of type H is just equal to their rate of *heterophily*, given by $1 - h_i^{L*}$. The second equation has a similar interpretation. We can then write the utility of type$-L$ parents as follows:

$$U_i^L = \pi_i^{LL} V^{LL} + \pi_i^{LH} V^{LH} - \frac{1}{2}\left(\tau_i^L\right)^2 + \alpha^L (1 - h_i^{L*}) \tau_i^L \tag{10}$$

where $V^{LL} > 0$ and $V^{LH} > 0$ denote the utility a type$-L$ parent derives from having a child of type L and H, respectively. Since low-educated parents prefer their kids to be educated, we assume that $V^{LL} < V^{LH}$ so that $\Delta V^{L} \equiv V^{LL} - V^{LH} < 0$. This utility has a similar interpretation as (7) since it has the standard benefits and costs of education effort and a new part that we can call "reciprocate heterophily ". Indeed, we assume that the higher is $1-h_i^{L*}$, i.e. the more heterophilous is the child, the higher is the utility of her low-educated parent of exerting education effort, i.e. $\frac{\partial^2 U_i^L}{\partial \tau_i^L \partial (1-h_i^{L*})} = \alpha^L > 0$. This is again a reasonable assumption, which is in fact verified in our data since we will show below that $\alpha^L > 0$ (see Table 3).[10]

5 Equilibrium analysis

Let us now determine the equilibrium of this model. Observe that each child i's socialization effort choice s_i is independent of the education effort τ_i^t of her parent of type t. However, τ_i^t is a function of s_i since each parent of a child i needs to know the equilibrium homophily index h_i^{H*} of her child in order to decide how much education effort to exert. Let us solve first the child's problem and then the parent's one.

5.1 Children' Socializing Activity

General result By maximizing utility (3) with respect to s_i, we easily obtain for each child i:

$$s_i^* = b_i + \frac{1}{4}\phi \sum_j \delta_{ij}\rho_{ij}s_j^* \tag{11}$$

Consider a total population of n parents/children (since the parent and the child are defined in the same unit) with n^H and n^L parents/children of type H and L, respectively, and with $n^H + n^L = n$. This means that we can define by $q = n^H/n$ the fraction of high-educated people in the population and by $1 - q = n^L/n$ the fraction of low-educated people in the population. Define $\mathbf{R} = (\rho_{ij})$ as the $(n \times n)$ matrix of geographical proximities with $\rho_{ii} = 0$. Define also $\widetilde{\mathbf{D}} = (\delta_{ij})$ as the $(n \times n)$ matrix of preference biases where clearly $\delta_{ii} = 0$. Remember that we assumed that δ_{ij} could only take 3 values: (i) $\delta_{ij} \equiv \delta^{HH} = \delta^{LL} = 1$ if i and j are of the same type, (ii) $\delta_{ij} \equiv \delta^{HL} < 1$ for a child i of type H interacting with child j of type L, (iii) $\delta_{ij} \equiv \delta^{LH} < 1$ for a child i of type L interacting with child j of type H. If by convention, and without loss of generality, in vectors and matrices we write first the kids of type H and then those of type L, we can define $\widetilde{\mathbf{D}}$ as follows:

$$\widetilde{\mathbf{D}} = \begin{pmatrix} \mathbf{J}^H & \widetilde{\mathbf{D}}^H \\ \widetilde{\mathbf{D}}^L & \mathbf{J}^L \end{pmatrix} \tag{12}$$

[10]Observe that we could have assumed that the cost of exerting education effort is higher for low-educated than high-educated parents by adding some parameter in the cost function, but this will not change any of our results.

where $\mathbf{J}^t$ (for $t = H, L$) is a $(n^t \times n^t)$ matrix with zeros on the diagonal and ones off the diagonal, $\widetilde{\mathbf{D}}^H$ is a $(n^H \times n^L)$ matrix with δ^{HL} everywhere (both on and off the diagonal) and $\widetilde{\mathbf{D}}^L$ is a $(n^L \times n^H)$ matrix with δ^{LH} everywhere (both on and off the diagonal).

In matrix form, we can then write (11) as follows:

$$\mathbf{s}^* = \mathbf{b} + \frac{\phi}{4}\mathbf{D}\,\mathbf{s}^* \tag{13}$$

where $\mathbf{b}$ is the $(n \times 1)$ vector of b_is and $\mathbf{D} \equiv \widetilde{\mathbf{D}} \circ \mathbf{R}$, where $\circ$ is the Hadamard product (i.e. entrywise product). Define $\mu_1(\mathbf{D})$ as the spectral radius of $\mathbf{D}$. We obtain the following result:

Proposition 1 *If $\mu_1(\mathbf{D}) < 4/\phi$, then there is a unique interior equilibrium of children' socialization choices given by*

$$\mathbf{s}^* = \left(\mathbf{I} - \frac{\phi}{4}\mathbf{D}\right)^{-1} \mathbf{b} \tag{14}$$

The proof of this proposition is straightforward by observing that:

$$\left(\mathbf{I} - \frac{\phi}{4}\mathbf{D}\right)^{-1} = \sum_{k}^{\infty} \left(\frac{\phi}{4}\right)^k \mathbf{D}^k$$

which is the Neuman series and so there exists a unique equilibrium if the condition given in the proposition is satisfied. Furthermore, by looking at the first-order condition (11), it is clear that the solution is interior. Interestingly, $\mathbf{s}^*$ is basically a weighted Bonacich centrality (see Ballester et al., 2006) of the network of proximity $\mathbf{D}$, which takes into account both geographical proximity and preference biases. Notice that $\mathbf{D}$ has the same entries for agents of the same type and at the same distance. Thus, heterogeneity in the link formation comes from the heterogeneity in the vector of idiosyncratic characteristics $\mathbf{b}$ and in the geographical distance. Notice also that, in the extreme case where there is no preference biases and all agents are located in the same location, $\mathbf{D}$ becomes a matrix of ones, with zero diagonal. In that case, $\mu_1(\mathbf{D}) = n - 1$ and the condition for the existence and uniqueness of equilibrium is given by $\phi < \frac{4}{n-1}$.

Given the equilibrium $\mathbf{s}^*$, we can calculate the expected network defined by the $(n \times n)$ matrix $\mathbf{P}^* = \left(p^*_{ij}\right)$, which represents all possible probabilities of forming a directed link between all agents two by two. Indeed, using (1), we can write its matrix form equivalent as:

$$\mathbf{P}^* = \frac{1}{4}\mathbf{D} \circ \mathbf{S} \tag{15}$$

where $\mathbf{S} \equiv \mathbf{s}^*\,\mathbf{s}^{*\mathrm{T}}$ and where the diagonal is made of zeros. This means that, given $\mathbf{P}^*$, we do not know the specific realization of the network but the expected network. Notice that $p^*_{ij} = p^*_{ji}$ only if $\delta^{HL} = \delta^{LH}$, which will generally not be true. As a result, the expected network is *asymmetric.*

Let us define more precisely the $(n \times n)$ matrix $\mathbf{P}^*$. We have:

$$\mathbf{P}^* = \begin{pmatrix} \mathbf{P}^{HH*} & \mathbf{P}^{HL*} \\ \mathbf{P}^{LH*} & \mathbf{P}^{LL*} \end{pmatrix}$$

where $\mathbf{P}^{tt'*}$ is a $\left(n^t \times n^{t'}\right)$ matrix that keeps track of the probability of links between children of type t and of type t'.

To summarize, the way two children i and j of different types socialize and their friendship link is formed essentially depend on four key elements: (i) the *preference bias* captured by δ^{HL} and δ^{LH}, (ii) the *meeting bias*, captured by $q = n^H/n$ but also by ρ_{ij}, the geographical distance between i and j,[11] (iii) the *synergy* of social interactions, captured by ϕ, (iv) the ex ante *heterogeneity* in their observable characteristics, b_i and b_j. If i and j are of the same type, then only (ii), (iii) and (iv) would matter.

We will now illustrate these different elements with some examples. In details, the first example shows the role of the standard *preference bias* and *meeting bias*; The second and third examples show, respectively, the role of *synergy* and *heterogeneity*.

Example 1 Consider a very small society composed of $n = 7$ parents/children, two of them are of type H and five of them of type L, i.e. $n^H = 2$ and $n^L = 5$ so that $q = 2/7 = 0.28$ (which is the fraction of type$-H$ students we have in our dataset; see Section 6.1 below). Assume that $\delta^{HL} = 0.6284$ and $\delta^{LH} = 0.6741$, which are the estimation values that we obtain in our regressions (see Table 3 below). This means that children of type L enjoy slightly more interacting with children of type H than the reverse. Then, $\widetilde{\mathbf{D}}$ is a (7×7) matrix defined as follows:

$$\widetilde{\mathbf{D}} = \begin{pmatrix} 0 & 1 & 0.6284 & 0.6284 & 0.6284 & 0.6284 & 0.6284 \\ 1 & 0 & 0.6284 & 0.6284 & 0.6284 & 0.6284 & 0.6284 \\ 0.6741 & 0.6741 & 0 & 1 & 1 & 1 & 1 \\ 0.6741 & 0.6741 & 1 & 0 & 1 & 1 & 1 \\ 0.6741 & 0.6741 & 1 & 1 & 0 & 1 & 1 \\ 0.6741 & 0.6741 & 1 & 1 & 1 & 0 & 1 \\ 0.6741 & 0.6741 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \tag{16}$$

Assume that $\phi = 0.0011$ (as in our Table 3 below). Assume also, for simplicity, that all agents reside at the same location so that the distance between any pair of agents is normalized to 1 so that the $\mathbf{R}$ matrix is a matrix of ones. Then, $\mathbf{D} = \widetilde{\mathbf{D}}$. Assume, finally, that there is no ex ante heterogeneity so that $\mathbf{b} = \mathbf{1}$, where $\mathbf{1}$ is a vector of ones. Denote by $\mathbf{M} \equiv \left(\mathbf{I} - \frac{\phi}{4}\mathbf{D}\right)^{-1}$, which is

[11] In a sense, our meeting bias is a mix of population shares and choices of some socialization activities which, impact the possibilities to meet other people. It is worth noticing that the parameter ρ_{ij} also captures the meeting bias but at the neighborhood level.

the matrix of (expected) discounted paths. Then,[12]

$$\mathbf{s}^* = \mathbf{M}\,\mathbf{b} = \begin{pmatrix} 1.0011 \\ 1.0011 \\ 1.0015 \\ 1.0015 \\ 1.0015 \\ 1.0015 \\ 1.0015 \end{pmatrix}$$

We see that, in terms of socialization efforts, type$-H$ children socialize slightly less than type$-L$ children. This is only due to the preference bias difference, since we assumed that $\delta^{HL} < \delta^{LH}$. Observe that the difference is tiny because ϕ is small. We can now obtain the matrix of probabilities of each link. Using (1) or (15), we have:

$$\mathbf{P}^* = \begin{pmatrix} 0 & 0.25 & 0.158 & 0.158 & 0.158 & 0.158 & 0.158 \\ 0.25 & 0 & 0.158 & 0.158 & 0.158 & 0.158 & 0.158 \\ 0.169 & 0.169 & 0 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.169 & 0.169 & 0.25 & 0 & 0.25 & 0.25 & 0.25 \\ 0.169 & 0.169 & 0.25 & 0.25 & 0 & 0.25 & 0.25 \\ 0.169 & 0.169 & 0.25 & 0.25 & 0.25 & 0 & 0.25 \\ 0.169 & 0.169 & 0.25 & 0.25 & 0.25 & 0.25 & 0 \end{pmatrix}$$

which implies that

$$\mathbf{P}^{HH*} = \begin{pmatrix} 0 & 0.25 \\ 0.25 & 0 \end{pmatrix}, \mathbf{P}^{HL*} = \begin{pmatrix} 0.158 & 0.158 & 0.158 & 0.158 & 0.158 \\ 0.158 & 0.158 & 0.158 & 0.158 & 0.158 \end{pmatrix}$$

$$\mathbf{P}^{LH*} = \begin{pmatrix} 0.169 & 0.169 \\ 0.169 & 0.169 \\ 0.169 & 0.169 \\ 0.169 & 0.169 \\ 0.169 & 0.169 \end{pmatrix}, \mathbf{P}^{LL*} = \begin{pmatrix} 0 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0 \end{pmatrix}$$

Since we impose that, within a group, all children are homogenous, then we obtain only four link probabilities, i.e. $p^{HH} = 0.25$, $p^{HL} = 0.158$, $p^{LH} = 0.169$ and $p^{LL} = 0.25$. For example, $p^{HH} = 0.25$ means that the probability that a child of type H (i.e. whose parent is high educated) will form a link with a given child of type H is 25%. We can see that $p^{LL} = p^{HH} > p^{LH} > p^{HL}$. This implies that type$-L$ students are more eager to form links with type$-H$ students than the

[12] It is easily verified that the condition $\mu_1(\mathbf{D}) < 4/\phi$ states in Proposition 1 holds here since $\mu_1(\mathbf{D}) = 5.0468$ and since $\phi = 0.0011$, we have: $5.0468 < 4/0.0011 = 3636.4$.

reverse. This is also due to the fact that we assumed that $\delta^{LH} > \delta^{HL}$ so that the preference bias is higher among type$-H$ children than among type$-L$ children. Notice that, if we had introduced some heterogeneity in the geographical proximity of agents (i.e. the ρ_{ij}) or in the idiosyncratic parameters (i.e. the b_i), then this would induce different weights in the Bonacich centrality, with agents with a higher level of b_i having a larger probability of forming links with others. Notice at last that the matrix of probabilities of links is asymmetric, treating each direction of the link as independent.

We can now calculate the homophily behavior of all children in this network. Denote by $\mathbf{d}^*$ the $(n \times 1)$ vector of (equilibrium) degrees, where the degree of child i is $d_i^* = \sum_j p_{ij}^*$ so that $\mathbf{d}^* = \mathbf{P}^*\mathbf{1}$. Denote also by $\widetilde{\mathbf{P}}^* = \left(\widetilde{p}_{ij}^*\right)$ the $(n \times n)$ matrix of (equilibrium) relative link probabilities for all children, where $\widetilde{p}_{ij}^* \equiv p_{ij}^*/d_i^*$and $\widetilde{\mathbf{P}}^{tt'*}$ the corresponding $(n^t \times n^{t'})$ matrix for children of type t so that

$$\widetilde{\mathbf{P}}^* = \begin{pmatrix} \widetilde{\mathbf{P}}^{HH*} & \widetilde{\mathbf{P}}^{HL*} \\ \widetilde{\mathbf{P}}^{LH*} & \widetilde{\mathbf{P}}^{LL*} \end{pmatrix}$$

Finally, denote by $\mathbf{H}^{t*} = \widetilde{\mathbf{P}}^{tt*}\mathbf{1}$, for $t = H, L$, the $(n^t \times 1)$ vector of (equilibrium) *homophily* behaviors for children of type t and by $\mathbf{H}^* = \begin{pmatrix} \mathbf{H}^{H*} \\ \mathbf{H}^{L*} \end{pmatrix}$ the $(n \times 1)$ vector of (equilibrium) homophily behaviors for all children. In our example,

$$\mathbf{d}^* = \begin{pmatrix} 1.04 \\ 1.04 \\ 1.338 \\ 1.338 \\ 1.338 \\ 1.338 \\ 1.338 \end{pmatrix}, \widetilde{\mathbf{P}}^* = \begin{pmatrix} 0 & 0.240 & 0.152 & 0.152 & 0.152 & 0.152 & 0.152 \\ 0.240 & 0 & 0.152 & 0.152 & 0.152 & 0.152 & 0.152 \\ 0.126 & 0.126 & 0 & 0.187 & 0.187 & 0.187 & 0.187 \\ 0.126 & 0.126 & 0.187 & 0 & 0.187 & 0.187 & 0.187 \\ 0.126 & 0.126 & 0.187 & 0.187 & 0 & 0.187 & 0.187 \\ 0.126 & 0.126 & 0.187 & 0.187 & 0.187 & 0 & 0.187 \\ 0.126 & 0.126 & 0.187 & 0.187 & 0.187 & 0.187 & 0 \end{pmatrix} \text{ and } \mathbf{H}^* = \begin{pmatrix} 0.240 \\ 0.240 \\ 0.748 \\ 0.748 \\ 0.748 \\ 0.748 \\ 0.748 \end{pmatrix}$$

which means, in particular, that

$$\mathbf{H}^{H*} = \begin{pmatrix} 0.240 \\ 0.240 \end{pmatrix} \text{ and } \mathbf{H}^{L*} = \begin{pmatrix} 0.748 \\ 0.748 \\ 0.748 \\ 0.748 \\ 0.748 \end{pmatrix} \tag{17}$$

We see that there is inbreeding homophily for the type$-L$ kids ($h^{L*} = 0.748$) since, in expectation, 74.8% of their friends should be of type L, while, there is strong heterophily for type$-H$ kids ($h^{H*} = 0.240$) since, in expectation, only 24% of their friends should be of type H. This result is due to both the preference bias (captured by δ^{HL} and δ^{LH}) and the meeting bias (captured by $q = 0.28$). Indeed, even though type$-H$ children have a relatively strong preference bias ($\delta^{HL} = 0.6284$), they have a very high heterophily behavior because there is a very high meeting bias against them since they have only 28% chance to meet type$-H$ children. On the contrary, type$-L$ children, who have both a strong preference bias ($\delta^{LH} = 0.6741$) and a strong meeting bias (the chance to meet a type$-L$ child is 72%), will mainly interact (74.8% of their time) with children of type L.

Example 2 In the previous example, we used for ϕ the parameter value that we obtain (see below) in our empirical analysis. Because ϕ turned out to be quite small ($\phi = 0.0011$) and because we assumed that children were ex ante identical in both geographical distances and observable characteristics, we mainly highlighted two channels: preference and meeting bias. Imagine, now, that we keep exactly the same parameter values but just increase ϕ so that children will socialize more. For example, take a $\phi = 0.2$. In that case, $\widetilde{\mathbf{D}}$ and thus $\mathbf{D}$ are still given by (16) but, now, children's socialization efforts are given by:

$$\mathbf{s}^* = \begin{pmatrix} 1.277 \\ 1.277 \\ 1.358 \\ 1.358 \\ 1.358 \\ 1.358 \\ 1.358 \end{pmatrix}$$

We now see that the socialization efforts are much higher and that type$-L$ children socialize more than type$-H$ children due to different preference biases. It is easily verified that:

$$\mathbf{P}^* = \begin{pmatrix} 0 & 0.408 & 0.272 & 0.272 & 0.272 & 0.272 & 0.272 \\ 0.408 & 0 & 0.272 & 0.272 & 0.272 & 0.272 & 0.272 \\ 0.292 & 0.292 & 0 & 0.461\,05 & 0.461 & 0.461 & 0.461 \\ 0.292 & 0.292 & 0.461 & 0 & 0.461 & 0.461 & 0.461 \\ 0.292 & 0.292 & 0.461 & 0.461 & 0 & 0.461 & 0.461 \\ 0.292 & 0.292 & 0.461 & 0.461 & 0.461 & 0 & 0.461 \\ 0.292 & 0.292 & 0.461 & 0.461 & 0.461 & 0.461 & 0 \end{pmatrix},$$

$$
\widetilde{\mathbf{P}}^* = \begin{pmatrix} 0 & 0.231 & 0.154 & 0.154 & 0.154 & 0.154 & 0.154 \\ 0.231 & 0 & 0.154 & 0.154 & 0.154 & 0.154 & 0.154 \\ 0.120 & 0.120 & 0 & 0.190 & 0.190 & 0.190 & 0.190 \\ 0.120 & 0.120 & 0.190 & 0 & 0.190 & 0.190 & 0.190 \\ 0.120 & 0.120 & 0.190 & 0.190 & 0 & 0.190 & 0.190 \\ 0.120 & 0.120 & 0.190 & 0.190 & 0.190 & 0 & 0.190 \\ 0.120 & 0.120 & 0.190 & 0.190 & 0.190 & 0.190 & 0 \end{pmatrix}
$$

and

$$
\mathbf{H}^* = \begin{pmatrix} 0.231 \\ 0.231 \\ 0.77 \\ 0.77 \\ 0.77 \\ 0.77 \\ 0.77 \end{pmatrix}
$$

Compared to Example 1, we see that the probability of forming a link for all students is much higher. This is because students socialize more due to higher synergies (i.e. higher ϕ). Despite this, the homophily behavior of all students is relatively similar to that of Example 1 because type$-L$ children are still very homophilous while type$-H$ children are still very heterophilous.

Example 3 Let us go back to the previous example (Example 1 with $\phi = 0.0011$) but let us introduce some heterogeneities in the geographical space. As in our data (see Section 6.1), assume that $\rho^{HH} = 0.09$, $\rho^{LL} = 0.12$ and $\rho^{HL} = \rho^{LH} = 0.10$. To calculate this value, we use: $\rho_{ij} = \exp(-d_{ij})$, where d_{ij} is the average geographical distance between two students i and j (see Section 6.1 below). For example, $\rho^{HH} = 0.09$ means that the average distance between two students of type H is 2.4 km. As a result,

$$
\mathbf{R} = \begin{pmatrix} 0 & 0.09 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 \\ 0.09 & 0 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 \\ 0.10 & 0.10 & 0 & 0.12 & 0.12 & 0.12 & 0.12 \\ 0.10 & 0.10 & 0.12 & 0 & 0.12 & 0.12 & 0.12 \\ 0.10 & 0.10 & 0.12 & 0.12 & 0 & 0.12 & 0.12 \\ 0.10 & 0.10 & 0.12 & 0.12 & 0.12 & 0 & 0.12 \\ 0.10 & 0.10 & 0.12 & 0.12 & 0.12 & 0.12 & 0 \end{pmatrix}
$$

and thus

$$\mathbf{D} \equiv \widetilde{\mathbf{D}} \circ \mathbf{R} = \begin{pmatrix} 0 & 0.09 & 0.06284 & 0.06284 & 0.06284 & 0.06284 & 0.06284 \\ 0.09 & 0 & 0.06284 & 0.06284 & 0.06284 & 0.06284 & 0.06284 \\ 0.06741 & 0.06741 & 0 & 0.12 & 0.12 & 0.12 & 0.12 \\ 0.06741 & 0.06741 & 0.12 & 0 & 0.12 & 0.12 & 0.12 \\ 0.06741 & 0.06741 & 0.12 & 0.12 & 0 & 0.12 & 0.12 \\ 0.06741 & 0.06741 & 0.12 & 0.12 & 0.12 & 0 & 0.12 \\ 0.06741 & 0.06741 & 0.12 & 0.12 & 0.12 & 0.12 & 0 \end{pmatrix}$$

where $\widetilde{\mathbf{D}}$ is still given by (16). We obtain:

$$\mathbf{s}^* = \begin{pmatrix} 1.0001 \\ 1.0001 \\ 1.0002 \\ 1.0002 \\ 1.0002 \\ 1.0002 \\ 1.0002 \end{pmatrix}$$

As before, because ϕ is quite small, there is no much difference in socialization efforts. We also obtain:

$$\mathbf{P}^* = \begin{pmatrix} 0 & 0.0225 & 0.0157 & 0.0157 & 0.0157 & 0.0157 & 0.0157 \\ 0.0225 & 0 & 0.0157 & 0.0157 & 0.0157 & 0.0157 & 0.0157 \\ 0.0169 & 0.0169 & 0 & 0.03 & 0.03 & 0.03 & 0.03 \\ 0.0169 & 0.0169 & 0.03 & 0 & 0.03 & 0.03 & 0.03 \\ 0.0169 & 0.0169 & 0.03 & 0.03 & 0 & 0.03 & 0.03 \\ 0.0169 & 0.0169 & 0.03 & 0.03 & 0.03 & 0 & 0.03 \\ 0.0169 & 0.0169 & 0.03 & 0.03 & 0.03 & 0.03 & 0 \end{pmatrix},$$

$$\widetilde{\mathbf{P}}^* = \begin{pmatrix} 0 & 0.223 & 0.155 & 0.155 & 0.155 & 0.155 & 0.155 \\ 0.223 & 0 & 0.155 & 0.155 & 0.155 & 0.155 & 0.155 \\ 0.110 & 0.110 & 0 & 0.195 & 0.195 & 0.195 & 0.195 \\ 0.110 & 0.110 & 0.195 & 0 & 0.195 & 0.195 & 0.195 \\ 0.110 & 0.110 & 0.195 & 0.195 & 0 & 0.195 & 0.195 \\ 0.110 & 0.110 & 0.195 & 0.195 & 0.195 & 0 & 0.195 \\ 0.110 & 0.110 & 0.195 & 0.03 & 0.195 & 0.195 & 0 \end{pmatrix}$$

$$\mathbf{H}^* = \begin{pmatrix} 0.223 \\ 0.223 \\ 0.775 \\ 0.775 \\ 0.775 \\ 0.775 \\ 0.775 \end{pmatrix}$$

The probabilities of forming links are quite small because they are now weighted a proximity index ρ_{ij}, which is relatively small. However, the homophily behaviors are relatively similar.

5.2 Parents' Socialization Effort

We now turn to the analysis of the parents' effort choice. As stated above, parents do not observe the realization of the network $\mathbf{G}$ created by the probability matrix $\mathbf{P}^*$, but just know the matrix $\mathbf{P}^*$ and thus the expectation over each possible link.

General result Remember from above that we defined by $\mathbf{H}^{t*} = \widetilde{\mathbf{P}}^{tt*}\mathbf{1}$, the $(n^t \times 1)$ vector of (equilibrium) *homophily* behaviors for children of type $t = H, L$. Using (5) and (6), the utility of high-educated parents is given by:

$$U_i^H = V^{HH}\tau_i^H + V^{HH}(1-\tau_i^H)h_i^{H*} + V^{HL}(1-\tau_i^H)\left(1-h_i^{H*}\right) - \frac{1}{2}\left(\tau_i^H\right)^2 + \alpha^H h_i^{H*}\tau_i^H$$

Similarly, using (8) and (9), the utility of low-educated parents can be written as:

$$U_i^L = \tau_i^L V^{LH} + (1-\tau_i^L)h_i^{L*}V^{LL} + V^{LH}(1-\tau_i^L)(1-h_i^{L*}) - \frac{1}{2}\left(\tau_i^L\right)^2 + \alpha^L(1-h_i^{L*})\tau_i^L$$

Proposition 2 *Assume that*

$$\alpha^t \leq 1, \text{ and } \left|\Delta V^t\right| \leq 1, \forall t \in \{H, L\} \tag{18}$$

(i) The optimal education effort of each parent i of type $t \in \{H, L\}$ is given by:

$$\tau_i^{H*} = \Delta V^H + \left(\alpha^H - \Delta V^H\right) h_i^{H*} \tag{19}$$

$$\tau_i^{L*} = \alpha^L - \left(\alpha^L + \Delta V^L\right) h_i^{L*} \tag{20}$$

where $\Delta V^H = V^{HH} - V^{HL} > 0$ and $\Delta V^L = V^{LL} - V^{LH} < 0$.

(ii) The behavior of type$-H$ parents exhibits cultural substitution (cultural complementarity) if

and only if $\alpha^H < \Delta V^H$ ($\alpha^H > \Delta V^H$) *since*

$$\frac{\partial \tau_i^{H*}}{\partial h_i^{H*}} \gtreqqless 0 \Leftrightarrow \alpha^H \gtreqqless \Delta V^H$$

(iii) *The behavior of type$-L$ parents exhibits cultural complementarity (cultural substitution) if and only if* $\alpha^L < -\Delta V^L$ ($\alpha^L > -\Delta V^L$) *since*

$$\frac{\partial \tau_i^{L*}}{\partial h_i^{L*}} \gtreqqless 0 \Leftrightarrow \alpha^L \lesseqqgtr -\Delta V^L$$

First, observe that, since $0 \leq h_i^{t*} \leq 1$, $\forall t = H, L$, $\tau_i^{H*} > 0$ since $\Delta V^H > \Delta V^H h_i^{H*}$ and $\tau_i^{L*} > 0$ since $\alpha^L > \alpha^L h_i^{L*}$. Moreover, conditions (18) are sufficient conditions that guarantee that $\tau_i^{H*} \leq 1$ and $\tau_i^{L*} \leq 1$. Second, each parent of each type t has a different socialization effort, which depends on the specific (expected) network her children belong to and thus the homophily behavior of the children. Third, an interesting result here, is that parents' efforts may show either cultural complementarity or substitution depending on the gain of having a child educated (ΔV^H or $-\Delta V^L$) and the marginal benefit of having an *homophilic* (for type$-H$ parents) or *heterophilic* child (for type$-L$ parents) since $\frac{\partial^2 U_i^H}{\partial h_i^{H*} \partial \tau_i^H} = \alpha^H$ and $\frac{\partial^2 U_i^L}{\partial h_i^{H*} \partial \tau_i^L} = \alpha^L$. To understand this result, let us rewrite the equilibrium efforts (19) and (20) as follows:

$$\tau_i^{H*} = \left(1 - h_i^{H*}\right) \Delta V^H + h_i^{H*} \alpha^H$$

$$\tau_i^{L*} = \left(1 - h_i^{L*}\right) \alpha^L + h_i^{L*} \left(-\Delta V^L\right)$$

One can see that the optimal parental's effort is a convex combination between the utility gain of having an educated child and α^t, where the weights are the degree of homophily of children. When α^H or α^L are large enough, then all the weights are on h_i^{H*} and $1 - h_i^{L*}$ and thus there is *cultural complementarity* between h_i^{H*} and τ_i^{H*}, i.e. the more the child is prone to have type$-H$ relationships (which means homophily for type$-H$ children but heterophily for type$-L$ children), the more parents of both types exert effort in influencing their child of becoming educated. On the contrary, when ΔV^H or $-\Delta V^L$ are large enough, then all the weights are on h_i^{L*} and $1 - h_i^{H*}$, and, as a result, there is *cultural substitutability*, since parents exert more effort, the less children are likely to have type$-H$ relationships.

Example 1 Let us go back to our example 1 of the previous section where homophily behaviors were given by (17). Assume that $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.3212$ and $\alpha^L = 0.2221$. These are the values we obtained when we calibrated our model to our data (see Table 2 below). These values mean that there is *cultural complementarity* for type$-H$ parents (Proposition 2, part (i)) since $\alpha^H > \Delta V^H$ but *cultural substituability* for type$-L$ parents (Proposition 2, part (iii))

since $\alpha^L > -\Delta V^L$. Then, it is easily verified that, using (19), (20) as well as (17), the optimal socialization choices of parents are given by:

$$\tau^{H*} = 0.289 \text{ and } \tau^{L*} = 0.190$$

This result is easy to understand. First, $\Delta V^H = 0.2794 > 0.1787 = -\Delta V^L$, which means parents of type H value much more than parents of type L that their offsprings become educated. Second, $\alpha^H = 0.3212 > 0.2221 = \alpha^L$, which means that the marginal benefit of having an *homophilic* child for a type$-H$ parent is much more higher than for a type$-L$ parent. Third, type$-L$ children are much more homophilous than type$-H$ children ($h^{L*} = 0.748 > 0.240 = h^{H*}$). Since there is cultural substituability for parents of type L, higher homophily means less educational effort. As a result, for all these reasons, type$-H$ parents exert more educational effort than type$-L$ parents.

We can now calculate the equilibrium transition probabilities. They are given by:

$$\pi^{HH*} = 0.460,\ \pi_i^{HL*} = 0.540,\ \pi_i^{LH*} = 0.394 \text{ and } \pi_i^{LL*} = 0.606$$

Children of high-educated parents are more likely to become highly educated than children of low-educated parents ($\pi^{HH*} > \pi_i^{LH*}$). Indeed, on the one hand, type$-H$ parents put more educational effort ($\tau^{H*} > \tau^{L*}$) so their children have more chance to become directly (*vertical transmission*) highly educated. On the other hand, both types of children have nearly the same level of social interactions with type$-H$ children ($h_i^{H*} = 0.252$ for i of type H and $h_i^{H*} = 1 - h_i^{L*} = 0.240$ for i of type L), i.e. their probability of becoming highly educated through *horizontal transmission* is roughly 25%. As a result, $\pi^{HH*} > \pi_i^{LH*}$. Moreover, social mobility for type$-L$ children is relatively low ($\pi_i^{LH*} = 0.394$) because their parents do not exert enough effort in education, they are surrounded by mainly low-educated children and they are too homophilous.

6 Structural estimation

We would now like to structurally estimate the model and perform some comparative statics results as well as some policy experiments.

6.1 Data

For that, we use a (relatively) well-known database on friendship networks from the National Longitudinal Survey of Adolescent Health (AddHealth).[13] The AddHealth survey has been designed

[13]This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain

to study the impact of the social environment (i.e. friends, family, neighborhood and school) on adolescents' behavior in the United States by collecting data on students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in years 1994-95. A subset of these adolescents, about 20,000 individuals, is also asked to compile a longer questionnaire containing sensitive individual and household information, and includes the geographical coordinates of their residential address.

From a network perspective, the most interesting aspect of the AddHealth data is the friendship information, which is based upon actual friends nominations. Indeed, pupils were asked to identify their best friends from a school roster (up to five males and five females).[14] As a result, one can reconstruct the whole geometric structure of the friendship networks (each network is denoted by g), summarized in the adjacency matrix $\mathbf{G} = (g_{ij})$, where $g_{ij} = 1$ if i and j are friends, and $g_{ij} = 0$, otherwise. Quite naturally, $g_{ii} = 0$. Consider a student i of type $t = H, L$. Then, we denote by $\sum_j g_{ij}^t$ the number of friends of student i of type t. Such a detailed information on social interaction patterns allows us to measure the peer group precisely by knowing exactly who nominates whom in a network (i.e. who interacts with whom in a social group). To define a type of a parent or a child, we say a child is of type H if any one of his/her parent (father or mother) is a college graduate. Otherwise, he/she is of type L.

We removed private and vocational public schools from the sample, in order to remove an additional potential source of self-selection and to ensure that distance between students is exogenous conditional on school choice. The final sample size, after removing unconnected individuals and members of networks with less than four members is of $5,281$ students. Among these $5,281$ students, 28% are from high-educated families so that we have: $q = n^H/n = 0.28$ and $1 - q = n^L/n = 0.72$. In our sample we have 66 school and 113 networks. The definition of all variables can be found in Table A1 in Appendix B and a complete summary statistics can be found in Table 2.

[*Insert Table* 2 *here*]

From the geographical coordinates, we are able to build our proximity index ρ_{ij} in the following way. For each pair of individuals i and j in school r, we first compute the Euclidean distance d_{ij} between their homes. The proximity index is then given by: $\rho_{ij,r} = \exp(-d_{ij})$. The index has two desirable properties: it is always in $[0, 1)$ and decreases exponentially as distance increases. In our data, the average distance between two students of type H is 2.41 km (i.e. $\rho^{HH} = 0.09$), the average distance between two students of type L is 2.12 km (i.e. $\rho^{LL} = 0.12$) and the average distance between a student of type H and a student of type L is 2.3 km (i.e. $\rho^{HL} = \rho^{LH} = 0.10$).

To measure socializing activity levels of the children (the s_is in the model), we use a composite

the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

[14] The limit in the number of nominations is not binding. Less than 1% of the students in our sample show a list of ten best friends.

index variable so that s_i is defined as the sum of after-school activities of student i that requires interactions with others. Each of these after-school activities takes a value of 1 if the student perform the activity, and zero otherwise. The activities are: dance, music, any kind of sports, writing or editing the school newspaper, honors club, foreign language clubs, participating in the school council, other clubs. We also add whether it is common for them to just hang out with friends, and whether they know people in their neighborhood and if during the month before the interview they stopped on the street to talk to someone they knew. Finally, if they use physical fitness or recreation center in the neighborhood. To sum it up, our proxy for s is a sum of all possible after school activities that could spur social interactions with other fellow students. The sum is then normalized between zero and one. Summary statistics for all these variables can be found in Table A2 in Appendix B. The average of such socialization activity index for type$-H$ students is 0.54, while it is 0.49 for students from low educated families (type$-L$ students). This indicates that students from high-educated families socialize more than those from low-educated families.

As a measure for parental effort τ, we also use a composite indicator, which is calculated by averaging all the non-zero answers to the questions asked to students regarding the relationship with their parents. In particular, we use the three following questions asked to the parents whether: (i) they talked to the child about his/her grades, (ii) they helped the child on a school project, (iii) talked to the child about other things he/she did at school. The index is then normalized to one by averaging the answers over the number of parents. The average value of τ is 0.24, with 0.23 for low-educated parents (type L) and 0.28 for high-educated parents (type H). This again indicates that high-educated parents put more effort in education-related activities than low-educated parents.

If we use the *observed* network $\mathbf{G}$, then we can calculate homophily as in the model (see (4)). On average, it is equal to 0.66 overall, with $\overline{h}_i^{H*} = 0.48$ (for type$-H$ students) and $\overline{h}_i^{L*} = 0.74$ (for type$-L$ students).[15] This suggests that low-educated students are more homophilic than high-educated students.

6.2 Empirical strategy

We now describe the strategy we employ in order to obtain point estimates of the parameters and how we make inference on them. First of all, recall that our model is based on a definition of a link probability (1), i.e.

$$p_{ij} = \frac{1}{4}\delta_{ij}\rho_{ij}s_i s_j$$

and the three following equilibrium equations (11), (19) and (20) that determine, respectively, the equilibrium socialization effort of children, the socialization effort of parents of type H and L,

[15] A variable with a bar indicates an average value of this variable.

respectively:[16]

$$s_i = b_i + \frac{1}{4}\phi \sum_j \delta_{ij}\rho_{ij}s_j$$

$$\tau_i^H = A^H + B^H \sum_j \widehat{p}_{ij}^H \tag{21}$$

$$\tau_i^L = A^L + B^L \sum_j \widehat{p}_{ij}^L \tag{22}$$

where $A^H \equiv \Delta V^H$, $B^H \equiv \alpha^H - A^H$, $A^L \equiv \alpha^L$ and $B^L \equiv -\Delta V^L - A^L$ and where homophily for type$-t$ students is defined as $h_i^t \equiv \sum_j \widetilde{p}_{ij}^t = \sum_j p_{ij}^t/d_i$. It is easily verified that equations (19) and (20), and equations (21) and (22), are the same.

Estimation of the preference bias parameter Let us start with the estimation of the preference-bias parameter. In the model, we assume that $\delta^{HH} = \delta^{LL} = 1$ so that there is, quite naturally, no preference bias against own group. Now, we need to estimate δ^{HL} and δ^{LH}.

First, let us assume that $\delta^{HL} = \delta^{LH} = \delta$ (*homogenous case*), which implies that $p_{ij} = p_{ji}$ (this is, of course, not necessary true when $\delta^{HL} \neq \delta^{LH}$, which we study below). We would like to structurally estimate δ. Remember that $\mathbf{P} = (p_{ij})$ is a matrix whose element p_{ij} is defined by (1), where $\delta_{ij} = \delta$ if i and j are of different types and 1, otherwise. In fact, $\mathbf{P} = \frac{1}{4}\mathbf{D} \circ \mathbf{S}$ (see (15)), where $\mathbf{D} = \widetilde{\mathbf{D}} \circ \mathbf{R}$ ($\widetilde{\mathbf{D}}$ keeps tract of the preference-bias parameters so in our case it is made of δ and 1 with zero diagonal while $\mathbf{R}$ keeps tract of the geographical distance ρ_{ij} between the residences all students). Moreover, $\mathbf{S} = \mathbf{s}\,\mathbf{s}^{\mathrm{T}}$. Remember also that $\widetilde{\mathbf{P}} = (\widetilde{p}_{ij})$, where $\widetilde{p}_{ij} = p_{ij}/d_i$ and $d_i = \sum_j p_{ij}$, is a matrix of relative link probability and the vector $\mathbf{H} = \begin{pmatrix} \mathbf{H}^H \\ \mathbf{H}^L \end{pmatrix}$ (with $\mathbf{H} = \widetilde{\mathbf{P}}\,\mathbf{1}$) keeps tract of homophily behavior for all students in the network. As discussed above, in our dataset, we observe $\mathbf{s}$ (the socialization effort of each student i), $\mathbf{R}$ (the geographical distances ρ_{ij}) and the network g of friendship links between all students but not $\widetilde{\mathbf{D}}$ since δ is unknown. As a result, we choose a δ, denoted by $\widehat{\delta}$, that minimizes the distance between the observed homophily in the network g and the expected one predicted by our model. We define the vector of *observed* homophily in our dataset by the vector $\mathbf{\Omega} = \begin{pmatrix} \mathbf{\Omega}^H \\ \mathbf{\Omega}^L \end{pmatrix}$, where $\omega_i^t = \sum_j g_{ij}^t / \sum_j g_{ij}$, for $t = H, L$.

We now determine the δ that minimizes the quadratic distance between observed homophily $\mathbf{\Omega}$ and predicted homophily $\mathbf{H}$, i.e.

$$\widehat{\delta} = \arg\min_{\delta} \sum_{i=1}^{n} (\omega_i^t - h_i^t(\delta))^2$$

where n is the number of individuals in our sample. Note that, to determine $\widehat{\delta}$ we are using basically

[16] To ease the presentation of this section, we skip the superscript star for the equilibrium variables.

a nonlinear least squares estimator that poses $\mathbf{\Omega} = \mathbf{H}(\delta)$. Since analytical derivatives are difficult to calculate, to make inference on $\widehat{\delta}$, we use a boostrap routine, where we resample $\mathbf{\Omega}$ within each school/network.

We can now relax the assumption that $\delta^{HL} = \delta^{LH} = \delta$ so that $\delta^{HL} \neq \delta^{LH}$ (*heterogenous case*). We proceed exactly as before, but to calculate $\widehat{\delta}^{HL}$ and $\widehat{\delta}^{LH}$, we split the sample into two as follows:

$$\widehat{\delta}^{HL} = \arg\min_{\delta^{HL}} \sum_{i=1}^{n_H} (\omega_i^H - h_i^H(\delta^{HL}))^2$$

$$\widehat{\delta}^{LH} = \arg\min_{\delta^{LH}} \sum_{i=1}^{n_L} (\omega_i^L - h_i^L(\delta))^2.$$

Estimation of the socialization spillover parameter We know from equation (11) that

$$s_i^*(\phi) = b_i + \frac{1}{4}\phi \sum_j \delta_{ij}\rho_{ij}s_j^*$$

which, in matrix form, can be written as (14), i.e.

$$\mathbf{s}^*(\phi) = \left(\mathbf{I} - \frac{\phi}{4}\mathbf{D}\right)^{-1} \mathbf{b}$$

In our dataset, we have: s_i^* (students' socialization efforts defined in Section 6.1), $\rho_{ij} = \exp(-d_{ij})$ (index of proximity based on geographical distances) and, from above, we have an estimation of the δs. For b_i, we use $\widetilde{b}_i$, the predicted values of a regression of s_i^* on a set of individual level characteristics and school fixed effects. So, the only unknown is ϕ. Denote by $s_{i,r}^o$ the *observed* socialization value of individual i in school r.

We can now determine the $\widehat{\phi}$ that minimizes the quadratic distance between observed socialization $\mathbf{s}^o$ and predicted socialization $\mathbf{s}^*(\phi)$ (defined in (11)), i.e.

$$\widehat{\phi} = \arg\min_{\phi} \sum_{i=1}^{n_r} (s_{i,r}^o - s_{i,r}^*(\phi))^2$$

For inference, as above, we use a boostrap routine, where we resample $s_{i,r}^o$ within each school.

Estimation of parental parameters For all the other parameters, we can estimate (21) and (22) by OLS. Their econometric equivalent can be written as:

$$\tau_{i,r}^H = A^H + B^H \sum_j \widehat{p}_{ij,r}^H + \beta' \mathbf{x}_{i,r}^H + \eta_r^H + \epsilon_{i,r}^H \tag{23}$$

$$\tau_{i,r}^L = A^L + B^L \sum_j \widehat{p}_{ij,r}^L + \beta' \mathbf{x}_{i,r}^L + \eta_r^L + \epsilon_{i,r}^L \tag{24}$$

where the vectors $\mathbf{x}_{i,r}^H$ and $\mathbf{x}_{i,r}^L$ are vectors of individual-level characteristics. Observe that, from the model, $\widetilde{p}_{ij}^t = p_{ij}^t / d_i$ and $d_i = \sum_j p_{ij}$, is a matrix of relative link probability, so that $h_i^t = \sum_j \widetilde{p}_{ij}^t$. In matrix form, $\mathbf{P} = \frac{1}{4}\mathbf{D} \circ \mathbf{S}$. Denote, now, by $\widehat{p}_{ij}^t$ the estimated value of $\widetilde{p}_{ij}^t$, which is defined in the same way with one difference: the matrix $\mathbf{D}$ is replaced by the matrix $\widehat{\mathbf{D}}$, where estimated $\widehat{\delta}$ are used. In equations (23) and (24), we use these values in $\sum\limits_j \widehat{p}_{ij,r}^H$ and $\sum\limits_j \widehat{p}_{ij,r}^L$. Finally, in order to account for the generated regressor $\sum\limits_j \widehat{p}_{ij,r}^L$ and cluster heteroscedasticity, standard errors are bootstrapped at the cluster (school) level. We also include a school fixed effect η_r to control for school selection.

6.3 Empirical results

The estimated parameters are reported in Table 3 for both the homogeneous case ($\delta^{HL} = \delta^{LH} = \delta$) in column 2 and the heterogeneous case ($\delta^{HL} \neq \delta^{LH}$) in column 3. We find that, in the homogenous case, the estimated preference bias is $\widehat{\delta} = 0.6546$, which means that groups do tend to prefer friends of their own type. When we consider heterogeneous preferences across groups, students from high-educated families have a stronger bias against the other type ($\widehat{\delta}^{HL} = 0.6284$) than those from a low-education background ($\widehat{\delta}^{LH} = 0.6741$).

[*Insert Table* 3 *here*]

We also find that the socialization spillover parameter ϕ is positive and significant but very small in magnitude ($\widehat{\phi} = 0.0011$).[17] This means that socialization spillovers are not that strong. Observe that its values does not change a lot from the homogeneous to the heterogeneous case.

Finally, the parameters governing parental effort are shown at the bottom of Table 3. First, we see that all variables are very significant and with the right signs.[18] Second, we observe that the utility of a high-educated parent having an educated child is higher than that of a low-educated parent having an educated child, i.e. $\Delta V^H > -\Delta V^L$. Finally, when looking at the estimations of α^H and α^L (in the heterogeneous case), we find that $0.3212 = \widehat{\alpha}^H > \widehat{\alpha}^L = 0.2221$. This means that the (marginal) benefit for type$-H$ parents of having an homophilic child is higher than the (marginal) benefit for type$-L$ parents of having an heterophilic child. More importantly, since $\alpha^H > \Delta V^H$ and since $\alpha^L > -\Delta V^L$ (for both the homogeneous and heterogeneous cases), we have *cultural complementarity* for high-educated parents and cultural substitution for low-educated parents (see Proposition 2). This means that, the higher is the homophily behavior of children from

[17] We verified that the condition is Proposition 1 is satisfied.

[18] We performed an F test for $A^H = 0$, $B^H + A^H = 0$, $A^L = 0$ and $B^L + A^L = 0$.

high-educated parents, the more these parents exert effort (here spend time doing homeworks with their kids) in educating their offsprings. The converse is true for low-educated parents. This means, in particular, that, for type$-L$ parents, the more their children are heterophilous, i.e. the more they spend time with educated kids, the more they spend effort in educating their offsprings.

7 Comparative statics results

Let us we perform some comparative statics in order to evaluate the properties of the equilibrium. It is, however, very difficult to derive analytical results.[19] As a result, we will resort to simulations based on the calibrated model. We calibrate the model as follows. Apart from the exercise when we vary δ, we will always consider the heterogeneous case when $\delta^{HL} \neq \delta^{LH}$. For the parameters of the model, we will take the values given in Table 3 so that $\delta^{HL} = 0.6284$, $\delta^{LH} = 0.6741$, $\phi = 0.0011$, $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.3212$ and $\alpha^L = 0.2221$.

As stated above, among the 5,281 students in our dataset, 28% are from high-educated families so that we have: $q = n^H/n = 0.28$ and $1 - q = n^L/n = 0.72$. For the geographical proximities, ρ_{ij}, recall we defined it as $\rho_{ij} = exp(-d_{ij})$, where d_{ij} is the euclidean distance between i and j residential locations. From the data, we take only three values: ρ^{HH}, ρ^{LL}, $\rho^{HL} = \rho^{LH}$, where $\rho^{tt'}$ is the average (over all schools) ρ between a child with a type$-t$ parent and another child with a type$-t'$ parent. In our dataset, we obtain: $\rho^{HH} = 0.09$, $\rho^{LL} = 0.12$ and $\rho^{HL} = 0.10$. Finally, for b_i, the ex ante characteristics of each student i, we take two values: b^H and b^L, one for type$-H$ students and one for type$-L$ students. To obtain these values, we take $\widetilde{b}_i$ as described before, which is the predicted value of a regression of s_i on all observable characteristics described in Table A1 and school fixed effects. We then take a linear combination of all these estimated characteristics. We obtain: $b^H = 0.54$ and $b^L = 0.49$.

7.1 Equilibrium behavior and preference bias

Let us study the effect the preference bias on equilibrium behaviors. For simplicity and tractability, we consider the homogeneous case so that $\delta^{HL} = \delta^{LH} = \delta$. Figure 1 displays the results of the simulations. In panel (a), we see that, when δ increases, children's socialization efforts of both types increase while children's homophily of both types decrease. Furthermore, when δ increases, parents' education effort increases for the low-educated parents while it decreases for the high-educated ones. Indeed, where people of a group become less prejudiced against the other group (i.e. δ increases), they tend to socialize more and thus are more likely to form links with each other. This implies that, for each pair i and j, p_{ij}^* increases but also the expected degree d_i^* increases. Since, homophily h_i^{t*} of child i of type t is defined as the ratio between $\sum_j p_{ij}^{t*}$ and $d_i^* = \sum_j p_{ij}^*$, then it should be

[19] We have some analytical results that we derive in Appendix C.

clear that h_i^{t*} is reduced following an increase of δ because d_i^* decreases faster than $\sum_j p_{ij}^{t*}$ increases. Since we have *cultural complementarity* for type$-H$ parents and *cultural substituability* for type$-L$ parents, parents of type H decrease their education efforts while the opposite is true for type$-L$ parents. Observe, however, that, even if the parental effort (τ_L^*) of type$-L$ parents is increasing, it is still much lower than that of type $-H$ parents but the homophily behaviors of their children is roughly the same.

$[Insert\ Figure\ 1\ here]$

In panel (b), we look at the impact of δ on the transition probabilities π^{HH} (probability that a high-educated parent has a child who becomes high-educated) π^{LH} (probability that a low-educated parent has a child who becomes high-educated) and $\pi^{H} = q\pi^{HH} + (1-q)\,\pi^{HL}$ (average transition probability). Remember that $\pi_i^{HH} = \tau_i^H + (1-\tau_i^H)h_i^H$ and $\pi_i^{LH} = \tau_i^L + (1-\tau_i^L)(1-h_i^L)$. Interestingly, Figure 1 shows that, when kids of one type become less prejudiced against kids of the other type (higher δ), the chance of becoming educated decreases for children from high-educated families while we have the opposite result for children from low-educated families. As stated above, when δ increases, type$-H$ parents decrease their educational effort τ_i^H and their kids become less homophilous (interact less with educated kids). As a result, π_i^{HH} decreases. On the contrary, when δ increases, type$-L$ parents increase their educational effort τ_i^L and therefore π_i^{LH} increases. This is interesting because it shows that less preference bias benefits the low-educated kids but harms the high-educated ones. When we look at the average effect π^H, we see that it is slightly increasing with δ meaning that, on average, children, whatever their origin, tend to be more educated with less preference bias. Observe finally that, at the extreme when $\delta = 1$ so that there is no prejudice bias at all, π_i^{LH} tends to converge towards π_i^{HH} so that the education status of the parent does not matter anymore. This is due to the fact that, in our data, the majority of students have low-educated parents (72 %) and thus, when students do not take into account the education status of of their friends' parents when forming friendship links, they all have nearly the same probability of becoming educated.

7.2 Equilibrium behavior and peer effect

Consider now the effect of an increase in the peer effect parameter ϕ. Figure 2 displays the simulation results. Consider first panel (a). When ϕ increases, there are more synergies and thus more returns from socialization between children (see (11)). As a result, because of these higher complementarities in efforts, children's socialization efforts increase for both types. This increases the probability p_{ij} of forming links between any pair of students (see (1)) and thus the expected degrees d_i of all students increases. In terms of homophily, both types of students interact less with type$-H$ students (less homophily for type $H-$students and less heterophily for type $L-$students). This is due to the meeting bias $q = 0.28$, which means that students tend to meet more often type$-L$

students than type$-H$ students. As a result, when both types of students exert more socialization efforts, they tend to socialize more with type$-L$ students than with type$-H$ students. Because there is *cultural complementarity* for type$-H$ parents and cultural substitution for type$-L$ parents, this, in turn, leads to the fact that both types of parents *decrease* their educational effort. Both parents do so because their children spend more and more time with low-educated children.

$[Insert\ Figure\ 2\ here]$

Let us comment the effects on the transition probabilities, as displayed by panel (b). Contrary to Figure 1, when ϕ increases, there is very little effect since the curves are relatively flat. Take, for example, parents and children of type H. The curve of π^{HH} is slightly decreasing when ϕ rises because parental educational effort decreases (less vertical transmission) but only slightly and homophily increases only slightly.

8 Policy experiments

So far, we have studied how the variation of two interesting parameters (δ, the preference bias, and ϕ, the synergy of social interactions) of the model affects equilibrium outcomes. However, these two parameters are not policy relevant since it is difficult for a policy maker to manipulate either of them. We would now like to study three different policy parameters: q, the fraction of high-educated children (measuring the meeting bias between high and low-educated children), ρ, the geographical proximity between students of different types (measuring the geographical segregation between high and low-educated children) and b, the degree of ex ante heterogeneity of the children (measuring the differences in observable characteristics between high and low-educated children). These three parameters are highly relevant for policy and can clearly be manipulated by the policy maker. The variation of these different parameters correspond to different policies that we would like to investigate now.

8.1 Social-mixing policies

We would like to study two *social-mixing* policies, which both reduce the meeting bias between students of different types. First, we examine a social-mixing policy where we increase q, the fraction of high-educated children, keeping constant the total number of children in the population. Second, we investigate another type of social-mixing policy where we reduce the geographical distance between children from high-educated and low-educated families.

8.1.1 Social-mixing policy 1: Increasing the fraction of high-educated children in the population

It is well-documented that children from high-educated families and low-educated families do not interact much with each other. For example, they tend not to attend the same school, the same after-school activities, etc. In this section, we would like to see what will happen to the equilibrium outcomes if we "integrate" more these two types of children by increasing q, the fraction of type$-H$ children in the economy. Observe that a higher q affects outcomes only through the matrices $\widetilde{\mathbf{D}}^H$ and $\widetilde{\mathbf{D}}^L$, located in $\widetilde{\mathbf{D}}$ (see (12)), since they change their dimension. Indeed, $\widetilde{\mathbf{D}}^H$ and $\widetilde{\mathbf{D}}^L$ are $(n^H \times n^L)$ and $(n^L \times n^H)$ matrices, respectively. So when q augments, so that, for example, n^H increases to $n^{H'} > n^H$ while n stays constant so that n^L decreases to $n^{L'}$, then the dimension of these matrices change to $(n^{H'} \times n^{L'})$ and $(n^{L'} \times n^{H'})$. These changes have an impact on socialization effort since the preference bias parameter $\delta^{HL}(= \delta^{LH})$ have now more weights than the δ^{LH}.

Figure 3 displays the results. In panel (a), we see that the impact of q on outcomes are quite different between the children. First, when q increases, type$-H$ students tend to socialize (slightly) more while type$-L$ students tend to socialize (slightly) less. Indeed, when q rises, both types of students are more likely to meet type$-H$ students. Since there is preference bias only for type$-L$ students, we obtain the effect on socialization efforts. Not surprisingly, this implies that type$-H$ students will become more homophilous while type$-L$ will become more heterophilous since both types will interact more with high-educated students. This implies that, in more segregated environment with more type$-H$ students, children from high-educated families will tend to have more friends while children from low-educated families will have less friends. Interestingly, both parents will exert more educational effort because, for type$-H$ parents, higher homophily induces them to invest more in their children's education (cultural complementarity) while, for type$-L$ parents, higher heterophily have a similar effect (cultural substituability).

Consider now panel (b). We find that, when q increases, both π^{HH} and π^{LH} increase, which means that *both* types of parents have *more* chance to have a highly educated kid. Indeed, for type$-H$ parents, we have seen (panel (a)) that, when q increases, their educational effort is increased. As a result, the vertical transmission of education (within the family) is higher. We have also seen that their offsprings become more homophilous so that the horizontal transmission is also higher. Therefore, π_i^{HH} increases. For type$-L$ parents, they reduce their educational effort (lower vertical transmission) but their kids become more heterophilous (higher horizontal transmission). The latter effect is much stronger than the former and, as a result, π^{LH} increases. When we look at the average effect of π^H, it is clearly positive and even above the 45 degree line, where the probability of being educated (y axis) is equal to the population share (x axis), $\pi^H = q$.

[*Insert Figure* 3 *here*]

8.1.2 Social-mixing policy 2: Reducing the geographical segregation between children from high and low-educated families

Another related social-mixing policy instrument consists in manipulating ρ, the geographical proximity (inverse of the distance) between students of different types. This instrument is relatively similar to that of q even it has a different impact: q was acting directly on the meeting bias while ρ, also has an impact of the meeting bias but in a more indirect way. We now implement a policy where we only increase ρ^{HL}, which means that we decreases the geographical distance between children of high-educated and low-educated families. Indeed, since $\rho_{ij,r} = \exp(-d_{ij})$, then when $\rho_{ij,r}$ increases, d_{ij} decreases. For example, in our dataset, $\rho^{HL} = \rho^{LH} = 0.12$, which implies an average distance of 2.12 km. If one increases $\rho^{HL} = \rho^{LH}$ to 0.5, then the average distance between a type$-L$ and a type$-H$ child decreases to 0.69 km. Figure 4 displays the results of the simulations of our calibrated model.

In panel (a), we see that, quite naturally, children of both types interact more with each other (higher socialization efforts) and thus have higher (expected) number of friends (degrees) but they both tend to be less homophilous. Indeed, since children of both types become closer in the geographical space, type$-H$ students have more type$-L$ friends while type$-L$ students have more type$-H$ friends. These two effects lead to the fact that both types of students are less homophilous. Because of cultural complementarity for type$-H$ parents and cultural substituability for type$-L$ parents, the former increase their educational effort while the latter decrease it when ρ^{HL} increases.

If we now consider panel (b), we see that this policy *favors* low-educated families while *harming* high-educated families. Indeed, for families of type H (type L), the direct vertical socialization is decreased (increased) since parents reduce (increase) their educational effort. Because kids of type H (type L) become less homophilous (more heterophilous), the horizontal transmission is also reduced (increased). As a result, when one spatially integrates children from low and high-educated families, the probability of becoming high-educated decreases for children from the high-educated families (π^{HH} decreases) while the probability of becoming high-educated increases for children from the low-educated families (π^{LH} increases). When we look at the average effect of π^{H}, we see that it is generally positive. This is also because there are much lower kids from high-educated families ($q = 28\%$) compared to kids from low-educated families.

This implies that the two social-mixing policies have very different impact on the probability of becoming educated for the children of different types. This is because they are quite different. Indeed, the first one, by increasing q without affecting the geographical distance between the different children, mechanically decreases $1 - q$, the fraction of kids from low-educated families, which, in turn, puts more weight on the preference bias of children from high-educated families. The second one, by increasing $\rho^{HL} = \rho^{LH}$ but keeping $q = 0.28$, changes the distribution of geographical distances between children of different types in the population, and, as a result, changes both the behavior of children (in terms of homophily) and of parents (in terms of educational efforts). In terms of educational outcomes, this favors type$-L$ children while harming type$-H$ children.

[*Insert Figure* 4 *here*]

8.2 Voucher policies

Let us now consider another policy, which consists in helping one type of children by giving vouchers to them. In our estimation, b_i is a linear combination of individual-level characteristics (described in Table A1). A way this policy can work is to increase the b_i of each student. For example, with this voucher policy the value of b_i increases from $b_i^t = b^t$ to $b_i^t = b^t + v^t$, where b^t is the idiosyncratic individual. Remember that, in our model, b_i is the direct marginal benefits from socialization for children of both types. In reality, such a policy could be implemented by giving some monetary or non-monetary incentives to children for socializing more. For example, one could induce children to participate more to after-school activities (such as the ones described in our data, i.e. dance, music, any kind of sports, writing or editing the school newspaper, honors club, foreign language clubs, participating in the school council, other clubs). In other words, any incentives that could be given to students for after school activities that could spur social interactions with other fellow students would be part of this policy. We call this policy a *voucher policy*, even though it is not the way researchers have designed voucher policies in the education literature since it usually means giving money or incentives to go to school and not necessary to participate to after-school activities.

8.2.1 Voucher policy 1: Helping only children from high-educated families

In our first policy experiment (Figure 5), we give *a voucher only to children from high-educated families*, that is we only increase v^H, so that b_i^H increases from $b^H = 0.54$ to $0.54 + v^H$ while b_i^L remains constant and equal to $b^L = 0.49$. In panel (a), we see that this has a very large positive impact on the socialization effort of type$-H$ students and a small positive impact on the socialization effort of type$-L$ students. Moreover, both types of students tend to interact more with type$-H$ students and have higher (expected) degrees. Indeed, by giving incentives to socialize more to only type$-H$ students spurs socialization efforts from this group who, because of preference bias, interact even more with type$-H$ students (higher homophily). This, however, generates positive externalities to type$-L$ students because their socialization effort is a function of the socialization effort of both types of children (see (11)). Therefore, type$-L$ students increase their socialization effort and also their degree of heterophily. These outcomes and the fact that we have cultural complementarity for type$-H$ parents and cultural substituability for type$-L$ parents, imply that both parents increase their educational effort.

If we now look at panel (b), we see that this policy benefits everybody since both types of children have a higher chance of becoming educated. Indeed, because parental's educational efforts increase, there is more vertical transmission for both types of children. Because there is more homophily for type$-H$ kids and more heterophily for type$-L$ kids, the horizontal transmission is also increased. As a result, all children have now a higher probability of becoming educated. For example, if we increase the voucher v^H from 0 to 0.45 (so that b^H increases from 0.54 to 0.99, i.e. an increase of 83%), then the probability of becoming educated for a type$-H$ child increases from 46% to 58% while, for a type$-L$ children, it increases from 27.5% to 37.5%.

[*Insert Figure* 5 *here*]

8.2.2 Voucher policy 2: Helping only children from low-educated families

In our second policy experiment (Figure 6), we give *a voucher only to children from low-educated families*, that is we only increase v^L so that b_i^L increases from $b^L = 0.49$ to $0.49 + v^L$ while b_i^H remains constant and equal to $b^H = 0.54$. In panel (a), in terms of socialization effort, we obtain the opposite result to Figure 5: type$-L$ children increase a lot their s_i^L while type$-H$ children increase slightly their s_i^H. The key difference is in terms of homophily. Because there is more benefits for type$-L$ students to socialize and because there is more type$-L$ children around $(1 - q = 0.72)$, both types of students socialize more with students from low-educated families. This implies that both parents reduce their educational efforts. As can be seen in panel (b), this has dramatic consequences on the (expected) level of education of children. Because both vertical and horizontal transmissions decrease, children from low and high-educated families are less likely to be educated. For example, if we increase the voucher v^L from 0 to 0.41 (so that b^L increases from 0.49 to 0.90, i.e. an increase of 83%), then the probability of becoming educated for a type$-H$ child decreases from 45% to 35% while, for a type$-L$ children, it decreases from 27.5% to 20%. We have seen above that an increase of the voucher v^H of also 83% has exactly the opposite effect with an increase of these two probabilities from 46% to 58% (type$-H$ children) and from 27.5% to 37.5% (type$-L$ children). Starting from the same probability, the difference in outcome is very large between the two policies, i.e. it is equal to 23% for type$-H$ children and 17.5% for type$-L$ children.

[*Insert Figure* 6 *here*]

9 Concluding remarks

In this paper, we presented a model of intergenerational transmission of preference for high education, in the presence of an endogenously-determined social context. To be more precise, we study the formation of a network of students as an equilibrium outcomes of of socializing activities between students. Parents observe the (expected) homophily of their children (i.e. share of own-type friends) and decide accordingly how much educational effort to exert. We structurally estimated all parameters of the model using adolescent friendship networks in the United States. We find that children's homophily acts as a complement to the educational effort of high-educated parents and as a substitute to the educational effort of low-educated parents

With the goal of increasing the probability of being educated of all students, we use the estimated parameters to run some policy experiments. We find that *place-based policies* that facilitate *social mixing* lead to higher probabilities of being educated, even though children from high-educated families may be harmed by a policy aiming at reducing geographical distance between the two groups. We also find that *school-based policies* aiming at increasing the socialization levels of children by giving them vouchers have a positive effect on educational outcomes only if these vouchers are given to children from the high-educated families.

Interestingly, social-mixing or place-based policies have been implemented in the United States. The most prominent social-mixing policy is the so-called Moving to Opportunity (MTO) programs (Katz et al., 2001; Kling et al., 2005, 2007; Ludwig et al., 2013), which give housing assistance (i.e. vouchers and certificates) to low-income families if they relocate to better and richer neighborhoods. The main idea of these programs is to reduce the geographical segregation between poor low-educated families and richer better-educated families. In some sense, our second social mixing policy (that increases ρ^{HL}) is very similar to the MTO programs. The main effect of the Moving to Opportunity (MTO) programs is to show no significant effects of neighborhoods on educational outcomes (see Katz et. al., 2001; Kling et al., 2001, 2007; Ludwig et al., 2001, 2013; Sanbonmatsu et al., 2006). However, it also shows a significant and positive long-term effect of neighborhoods on education (college attendance) and earnings only for kids who move when they were younger than 13 years old (Chetty et al., 2016). The main explanation of this result is that, when the MTO programs move people (mostly ethnic minorities) from poor areas to richer areas, they do not much interact with their "new neighbors" but instead with their "old neighbors" who are their real peers. For example, de Souza Briggs et al. (2010) document the fact that many African American families who moved to richer areas thanks to the MTO programs did not interact with their new neighbors because they felt rejected. In particular, on Sundays they were still going to the church in their previous neighborhood, even though it was located very far away from their current residence. However, when young children (below 13) move to a new area, they have time to build a new network of friends and therefore their new neighbors can become their peers and have a positive impact on their education outcomes. In our policy experiments where we increased ρ^{HL}, we have shown that this policy can be successful for the education outcomes of children from low-educated families because it helps them socializing more with children from high-educated families. We have also shown (Figure 4) that this positive effect was at this expense of children from high-educated families since they probability of being educated was decreasing with less segregation.

Our results in terms of voucher policies suggest that *school-based policies* could also be successful in terms of educational outcomes if they target the right students. Fryer and Katz (2013) and Katz (2015) suggest that school-based policies could be more effective than place-based policies. In other words, it could be better to improve the quality of schools in a bad neighborhood rather than moving people to better neighborhoods unless they are very young (Fryer and Katz, 2013).

Which policy is best for improving the educational outcomes of children is a very difficult question. In this paper, we have proposed an answer based on the role of children's socialization and homophily behavior and of parents effort in education outcomes. We believe that a successful education policy should therefore take into account its impact on children's social networks and on parents' education transmission.

References

[1] Albornoz, F., Cabrales, A. and E. Hauk (2014), "Which club should I attend, Dad? Targeted socialization and production," Barcelona GSE Working Paper nº 797.

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

[3] Bisin, A. and Patacchini, E. and Verdier, T. and Zenou, Y., (2011). "Formation and persistence of oppositional identities," *European Economic Review*, vol. 55(8), 1046-1071.

[4] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2016), "Bend it like Beckham. Ethnic identity and integration," *European Economic Review*, forthcoming.

[5] Bisin, A., and Verdier, T. (1998), "On the cultural transmission of preferences for social status," *Journal of Public Economics*, 70(1), 75-97.

[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 Verdier, T. (2001). "The economics of cultural transmission and the dynamics of preferences." *Journal of Economic Theory*, 97(2), 298-319.

[8] Bisin, A. and T. Verdier (2011), "The economics of cultural transmission and socialization," In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics*, Amsterdam: Elsevier Science, pp. 339-416.

[9] Black, S.E. and P.J. Devereux (2011), "Recent developments in intergenerational mobility," In: D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics, Vol. 4B*, Amsterdam: Elsevier Science, pp. 1487-1541.

[10] Blume, L.E., Brock, W.A., Durlauf, S.N. and Y.M. Ioannides (2011), "Identification of social interactions," In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics, Vol. 1B*, Amsterdam: Elsevier Science, pp. 853-964.

[11] Boucher, V. and B. Fortin (2016), "Some challenges in the empirics of the effects of networks," In: Y. Bramoullé, B.W. Rogers and A. Galeotti (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, pp. 277–302.

[12] Botticini, M. and Z. Eckstein (2005), "Jewish occupational selection: education, restrictions, or minorities?" *Journal of Economic History*, 65(04), 922-948.

[13] Botticini, M. and Z. Eckstein (2007), "From farmers to merchants, conversions and diaspora: Human capital and Jewish history," *Journal of the European Economic Association*, 5(5), 885-926.

[14] Buechel, B., Hellmann, T. and M.M. Pichler (2014), "The dynamics of continuous cultural traits in social networks," *Journal of Economic Theory* 154, 274-309.

[15] Cabrales, A., Calvó-Armengol, A. and Y. Zenou (2011), "Social interactions and spillovers," *Games and Economic Behavior* 72, 339-360.

[16] Calvó-Armengol, A., Patacchini, E. and Y. Zenou (2009), "Peer effects and social networks in education," *Review of Economic Studies* 76(4), 1239-1267.

[17] Card, D. and A.B. Krueger (1992), "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics* 107(1), 151-200.

[18] Card, D. and A.B. Krueger (1996), "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina," *Journal of Economic Perspectives* 10(4), 31-50.

[19] Chetty, R., Hendren, N. and L.F. Katz (2016), "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment," *American Economic Review* 106(4), 855–902.

[20] Currarini, S., Jackson, M.O., and P. Pin (2009), "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica* 77, 1003-1045.

[21] Currarini, S., Jackson, M.O., and P. Pin (2010), "Identifying the roles of race-based choice and chance in high school friendship network formation," *Proceedings of the National Academy of Sciences of the USA* 107, 4857-4861.

[22] de Souza Briggs, X., Popkin, S.J. and J. Goering (2010), *Moving to Opportunity: The Story of an American Experiment to Fight Ghetto Poverty*, Oxford: Oxford University Press.

[23] Ermisch, J. and M. Francesconi (2001), "Family matters: Impacts of family background on educational attainments," *Economica*, 68(270), 137-156.

[24] Francois, P. (2002), *Social Capital and Economic Development*, London: Routledge.

[25] Fryer, R.G. and L.F. Katz (2013), "Achieving escape velocity: Neighborhood and school interventions to reduce persistent inequality," *American Economic Review: Papers & Proceedings* 103, 232-237.

[26] Golub, B. and M.O. Jackson (2010), "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics* 2, 112-149.

[27] Golub, B. and M.O. Jackson (2012a), "How homophily affects the speed of learning and best-response dynamics," *Quarterly Journal of Economics* 127, 1287-1338.

[28] Golub, B. and M.O. Jackson (2012b), "Does homophily predict consensus times? Testing a model of network structure via a dynamic process," *Review of Network Economics* 11(3), Article 9.

[29] Graham, B.S. (2015), "Methods of identification in social networks," *Annual Review of Economics* 7, 465-485.

[30] Guryan, J., Hurst, E. and M. Kearney (2008), "Parental Education and Parental Time with Children,"*Journal of Economic Perspectives* 22(3), 23-46

[31] Hanushek, E.A. (2002), "Publicly provided education," In: A. J. Auerbach and M. Feldstein (Eds.), *Handbook of Public Economics*, Vol. 4, Amsterdam: Elsevier Science, pp. 2045-2141.

[32] Hauk, E. and M. Sáez-Martí (2002), "On the cultural transmission of corruption," *Journal of Economic Theory* 107(2), 311-335.

[33] Ioannides, Y.M. (2012), *From Neighborhoods to Nations: The Economics of Social Interactions*, Princeton: Princeton University Press.

[34] Jackson, M.O. (2008), *Social and Economic Networks*, Princeton: Princeton University Press.

[35] Jackson, M.O. and Y. Zenou (2015), "Games on networks," In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier, pp. 91-157.

[36] Jackson, M.O., Rogers, B.W. and Y. Zenou (2016), "The economic consequences of social network structure," *Journal of Economic Literature*, forthcoming.

[37] Jackson, M.O. and Y. Zenou (2015), "Games on networks," In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier, pp. 91-157.

[38] Jellal, M. and F.C. Wolff (2002), "Cultural evolutionary altruism: Theory and evidence," *European Journal of Political Economy* 18(2), 241-262.

[39] Jencks, C. and M. Phillips (1998), *The Black-White Test Score Gap*, Washington, D.C.: Brookings Institution.

[40] Katz, L.F. (2015), "Reducing inequality: Neighborhood and school interventions," *Focus* 31, 12-17.

[41] 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.

[42] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment," *Quarterly Journal of Economics* 120, 87-130.

[43] Kling, J.R., Liebman, J.B. and L.F. Katz (2007), "Experimental analysis of neighborhood effects," *Econometrica* 75, 83-119.

[44] Ludwig, J., Duncan, G.J., Gennetian, L.A., Katz, L.F., Kessler, R.C., Kling, J.R. and L. Sanbonmatsu (2013), "Long-term neighborhood effects on low-income families: Evidence from Moving to Opportunity," *American Economic Review: Papers & Proceedings* 103, 226-231.

[45] McPherson, M., Smith-Lovin, L. and J.M. Cook (2001), "Birds of a feather: Homophily in social networks," *Annual Review of Sociology* 27, 415-444.

[46] Olcina, G. and C. Pearrubia (2004), "Hold up and intergenerational transmission of preferences," *Journal of Economic Behavior & Organization* 54(1), 111-132.

[47] Olivier, J., Thoenig, M. and T. Verdier (2008), "Globalization and the dynamics of cultural identity," *Journal of International Economics* 76, 356-70.

[48] Panebianco, F. (2014), "Socialization networks and the transmission of interethnic attitudes," *Journal of Economic Theory* 150, 583-610.

[49] Panebianco, F., and Verdier, T. (2015), "Paternalism, cultural transmission and diffusion on complex networks," CEPR Discussion Papers No. 10722.

[50] Patacchini, E. and Y. Zenou (2009), "On the sources of the black-white test score gap in Europe," *Economics Letters* 102, 49–52.

[51] Patacchini, E. and Y. Zenou (2011), "Neighborhood effects and parental involvement in the intergenerational transmission of education," *Journal of Regional Science* 51(5), 987-1013

[52] Plug, E., and W. Vijverberg (2003), "Schooling, family background, and adoption: Is it nature or is it nurture?" *Journal of political Economy* 111(3), 611-641.

[53] Sacerdote, B. (2001), "Peer effects with random assignment: Results from Dartmouth roomates," *Quarterly Journal of Economics* 116, 681-704.

[54] Sacerdote, B. (2002), "The Nature and Nurture of Economic Outcomes," *American Economic Review* 92(2), 344-348.

[55] Sacerdote, B. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?", In: E.A. Hanushek, S. Machin and L. Woessmann (Eds.), *Handbook of Economics of Education*, *Vol. 3*, Amsterdam: Elevier Science, pp. 249-277.

[56] Sacerdote, B. (2014). "Experimental and quasi-experimental analysis of peer effects: two steps forward? " *Annual Review of Economics*, 6(1), 253–272.

[57] Saez-Marti, M. and Y. Zenou (2012), "Cultural transmission and discrimination," *Journal of Urban Economics* 72(2), 137–146.

[58] Sanbonmatsu, L., Kling, J.R., Duncan, G.J. and J. Brooks-Gunn (2006), "Neighborhoods and academic achievement: Results from the Moving to Opportunity experiment," *Journal of Human Resources* 41, 649–691.

[59] Vaughan, D. (2013), "Nurture vs. nurture: Endogenous parental and peer effects and the transmission of culture," Bank of Mexico Working Paper Series No. 2013-04.

[60] Verdier, T. and Zenou, Y. (2015), "The role of cultural leaders in the transmission of preferences," *Economics Letters*, 136, 158–161.

[61] Verdier, T. and Zenou, Y. (2016), "Cultural leader and the dynamics of assimilation," CEPR Discussion Paper No. 11174.

Tables

Table 1: Determinants of college attendance.

	(1) High	(2) Low	(3) High	(4) Low	(5) High	(6) Low
Parental effort	0.0621**	0.105***	0.0296	0.0938**	0.187	0.186
	(0.0318)	(0.0242)	(0.0569)	(0.0400)	(0.134)	(0.157)
Homophily			0.0916	-0.0568	0.185**	-0.0291
			(0.0886)	(0.0771)	(0.0922)	(0.100)
Effort × homophily					-0.341	-0.114
					(0.262)	(0.194)
Female	0.137***	0.0737***	0.172***	0.0734***	0.174***	0.0734***
	(0.0158)	(0.00869)	(0.0305)	(0.0141)	(0.0304)	(0.0141)
Black	-0.0658***	-0.0146	-0.0346	-0.0116	-0.0318	-0.0114
	(0.0250)	(0.0132)	(0.0531)	(0.0364)	(0.0532)	(0.0364)
Other Races	0.0350	-0.0190*	0.00723	-0.0224	0.00915	-0.0216
	(0.0270)	(0.0114)	(0.0432)	(0.0318)	(0.0430)	(0.0318)
Religion practice	0.0403***	0.0276***	0.0373***	0.0378***	0.0370***	0.0377***
	(0.00599)	(0.00366)	(0.0116)	(0.00669)	(0.0117)	(0.00669)
Physical devel	-0.00255	-0.000269	-0.0193	0.00117	-0.0190	0.00110
	(0.00682)	(0.00432)	(0.0155)	(0.00789)	(0.0153)	(0.00789)
Family size	-0.0284***	-0.0124***	-0.0277**	-0.0154**	-0.0274**	-0.0154**
	(0.00586)	(0.00312)	(0.0122)	(0.00703)	(0.0122)	(0.00702)
Two parents family	0.170***	0.0491***	0.138***	0.0370*	0.138***	0.0373*
	(0.0224)	(0.0111)	(0.0521)	(0.0211)	(0.0520)	(0.0210)
Parental education	0.120***	0.0595***	0.106***	0.0583***	0.106***	0.0584***
	(0.0181)	(0.00587)	(0.0328)	(0.0115)	(0.0329)	(0.0116)
Hours worked (parent)	-0.000529	0.000750**	-0.000954	0.00172***	-0.000927	0.00172***
	(0.000514)	(0.000317)	(0.00111)	(0.000534)	(0.00111)	(0.000532)
Parent unemployed	-0.0690	-0.0697***	-0.0615	-0.120***	-0.0607	-0.119***
	(0.0524)	(0.0228)	(0.0921)	(0.0412)	(0.0927)	(0.0411)
Family income	0.000542***	0.000729***	0.00197***	0.000782*	0.00198***	0.000782*
	(0.000197)	(0.000197)	(0.000520)	(0.000395)	(0.000519)	(0.000395)
Family income refused	0.0230	0.0225	0.129*	-0.00643	0.127*	-0.00651
	(0.0298)	(0.0148)	(0.0744)	(0.0269)	(0.0749)	(0.0269)
Age	0.00403	-0.00488	0.00346	0.00549	0.00378	0.00555
	(0.00447)	(0.00336)	(0.0101)	(0.00586)	(0.0101)	(0.00583)
Constant	-0.304**	0.0471	-0.231	-0.217	-0.288	-0.241
	(0.150)	(0.107)	(0.330)	(0.175)	(0.328)	(0.171)
Observations	3,596	8,990	1,036	2,777	1,036	2,777
R-squared	0.097	0.048	0.105	0.048	0.107	0.049
Number of schools	139	145	89	93	89	93

Linear probability model with school fixed effect. Dependent variable is college attendance.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Summary statistics

	Mean	St. dev.	Min	Max
Parental effort	0.2469	0.2410	0	1
Socialization	0.5112	0.1698	0.0714	1
Female	0.5266	0.4993	0	1
Other races	0.1322	0.3387	0	1
Two parent family	0.7232	0.4475	0	1
Parent unemployed	0.0718	0.2581	0	1
Family income refused	0.1110	0.3141	0	1
Black	0.2115	0.4084	0	1
Parental education	3.1515	0.9327	0	5
Family size	3.5111	1.4565	1	14
Religion practice	2.5737	1.433	0	4
Family income	42.6965	55.9145	0	999
Physical developement	3.2566	1.1001	1	5
Parent worked hours	40.2340	19.2336	0	84
Student grade	9.5755	1.5916	7	12

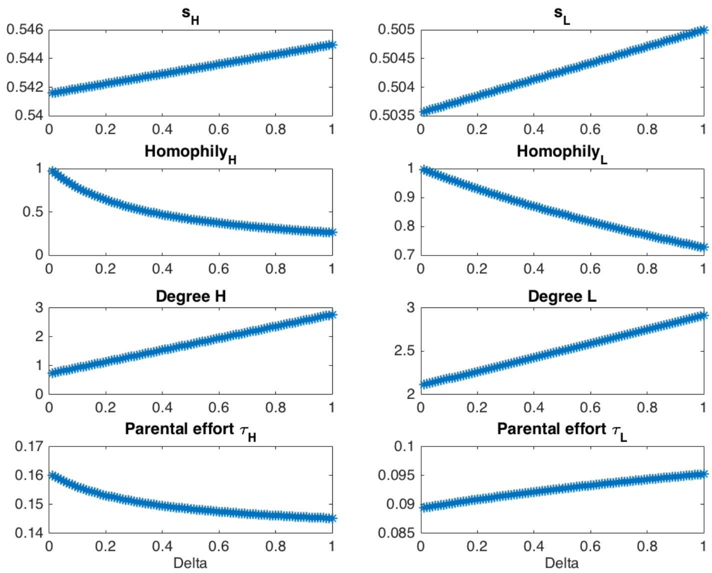
Sample size is 6,060

Table 3: Estimation results

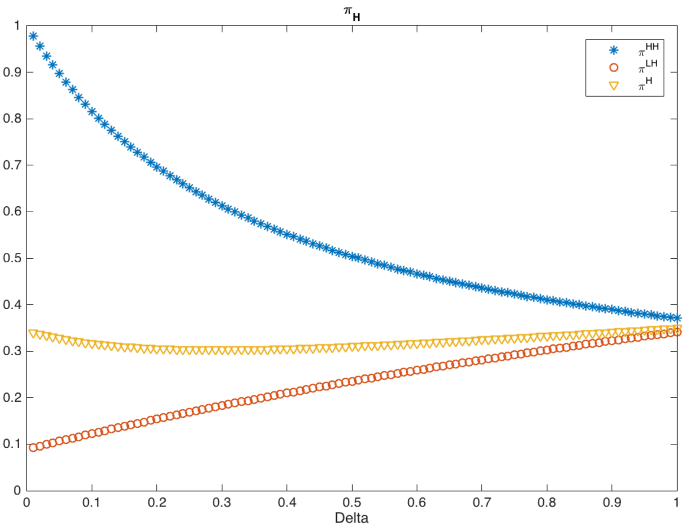
Parameter	Homogeneous	Heterogeneous
	Preference bias	
δ	0.6546***	-
δ^{HL}	-	0.6284***
δ^{LH}	-	0.6741***
	Spillover in socialization	
ϕ	0.0010***	0.0011***
	(0.0001)	(0.0001)
	Parental effort	
ΔV^H	0.2796***	0.2794***
$-\Delta V^L$	0.1788***	0.1787**
α^H	0.3211***	0.3212***
α^L	0.2227***	0.2221***

Bootstrapped standard errors clustered at the school level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Figures

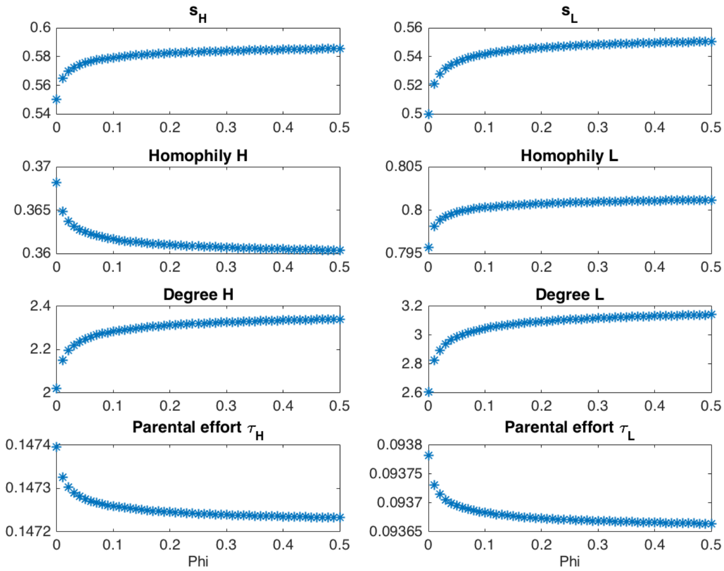


(a)

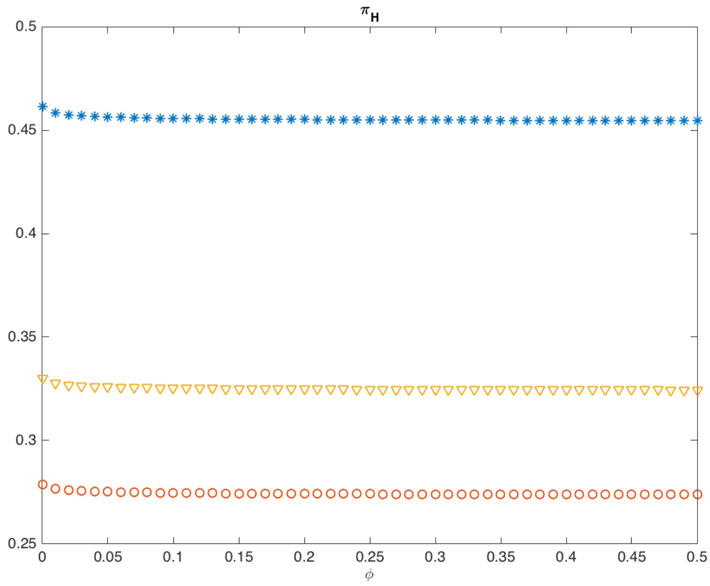


(b)

Figure 1: Simulation of equilibrium values for different levels of preference bias' parameter δ. Parameter values: $\phi = 0.0011$, $q = 0.28$, $b^H = 0.54$, $b^L = 0.49$, $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.32$, $\alpha^L = 0.22$, $\rho^{HH} = 0.09$, $\rho^{LL} = 0.12$, $\rho^{HL} = 0.10$

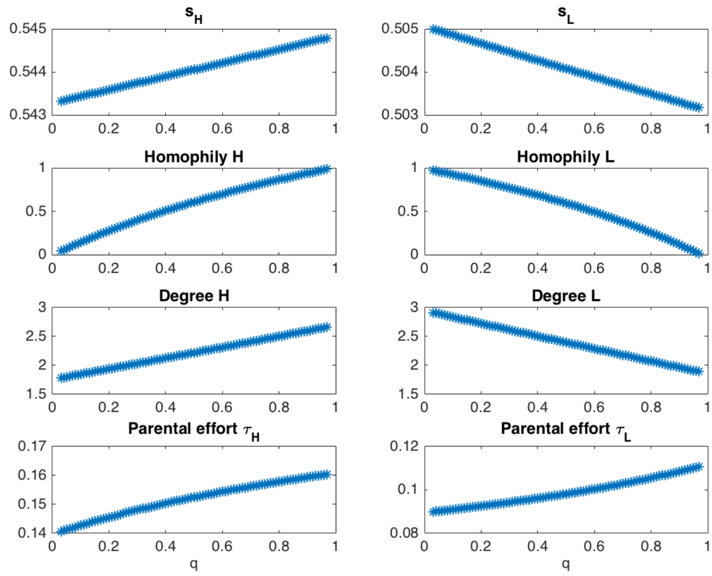


(a)

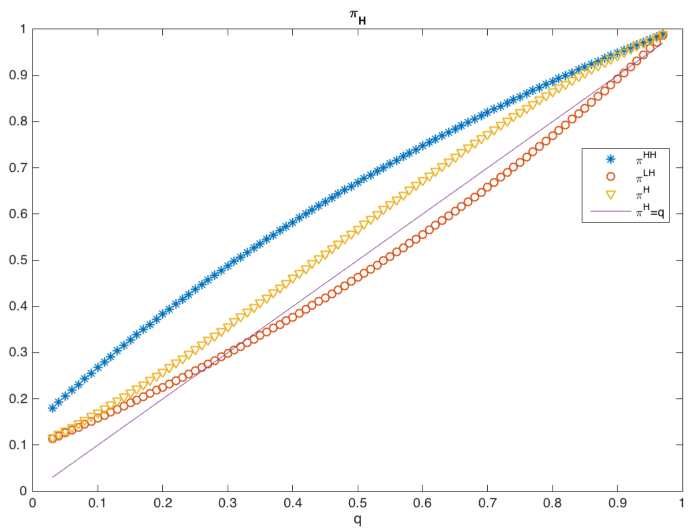


(b)

Figure 2: Simulation of equilibrium values for different levels of socialization spillover parameter ϕ. Parameter values: $\delta^{HL} = 0.6284$, $\delta^{LH} = 0.6741$, $q = 0.28$, $b^H = 0.54$, $b^L = 0.49$, $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.32$, $\alpha^L = 0.22$, $\rho^{HH} = 0.09$, $\rho^{LL} = 0.12$, $\rho^{HL} = 0.10$

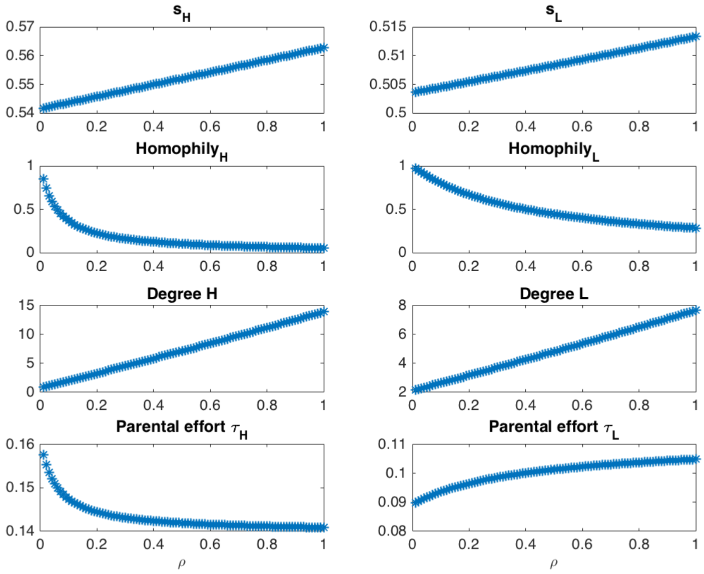


(a)

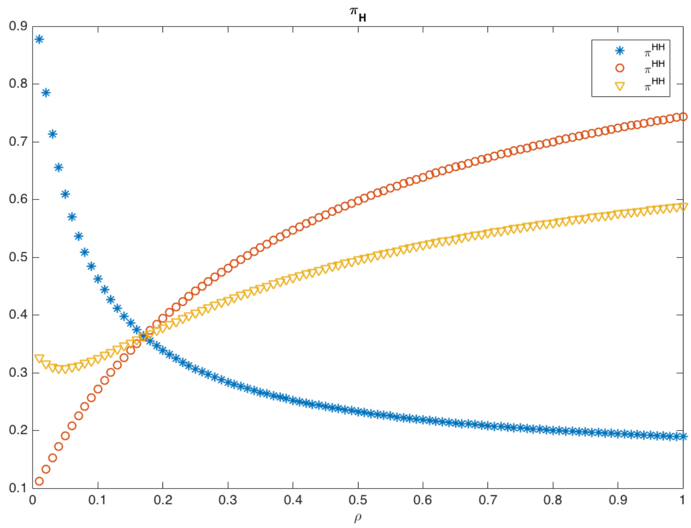


(b)

Figure 3: Simulation of equilibrium values for different levels of population share q. Parameter values: $\phi = 0.0011$, $\delta^{HL} = 0.6284$, $\delta^{LH} = 0.6741$, $b^H = 0.54$, $b^L = 0.49$, $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.32$, $\alpha^L = 0.22$, $\rho^{HH} = 0.09$, $\rho^{LL} = 0.12$, $\rho^{HL} = 0.10$

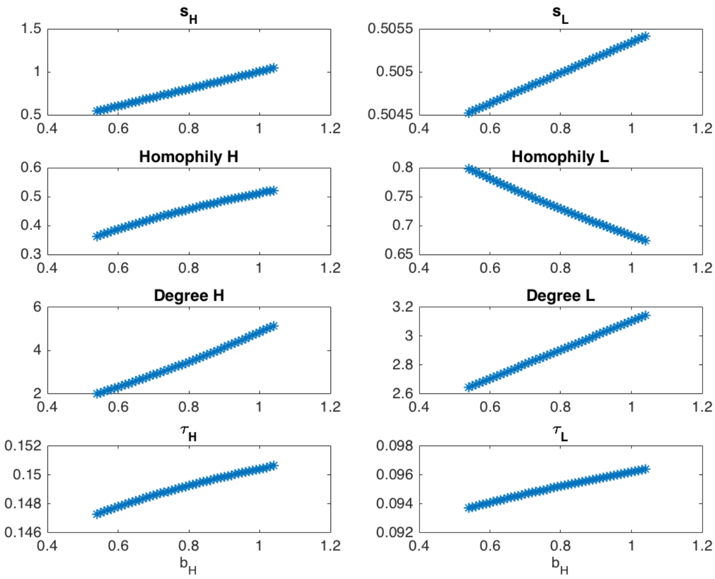


(a)

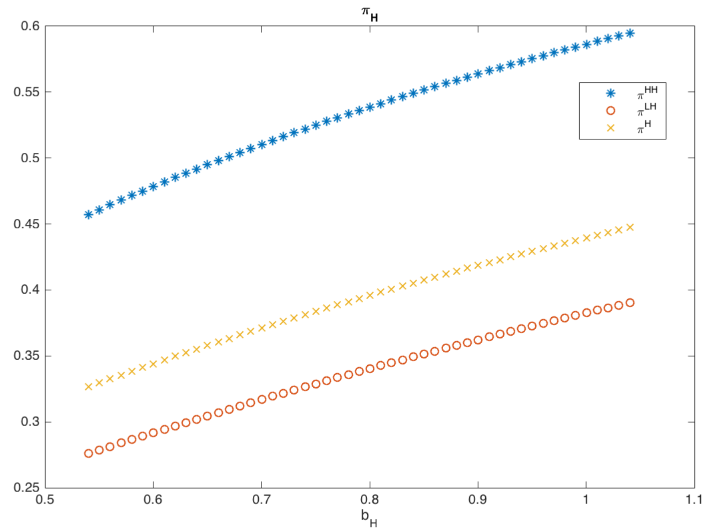


(b)

Figure 4: Simulation of equilibrium values for different levels of ρ^{HH}. Parameter values: $\phi = 0.0011$, $\delta^{HL} = 0.6284$, $\delta^{LH} = 0.6741$, $q = 0.28$, $b^H = 0.54$, $b^L = 0.49$, $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.32$, $\alpha^L = 0.22$, $\rho^{LL} = 0.12$, $\rho^{HL} = 0.10$

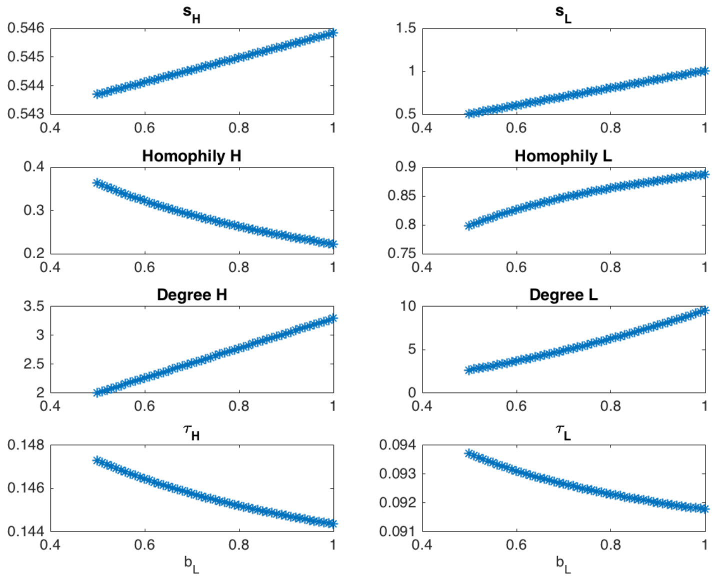


(a)

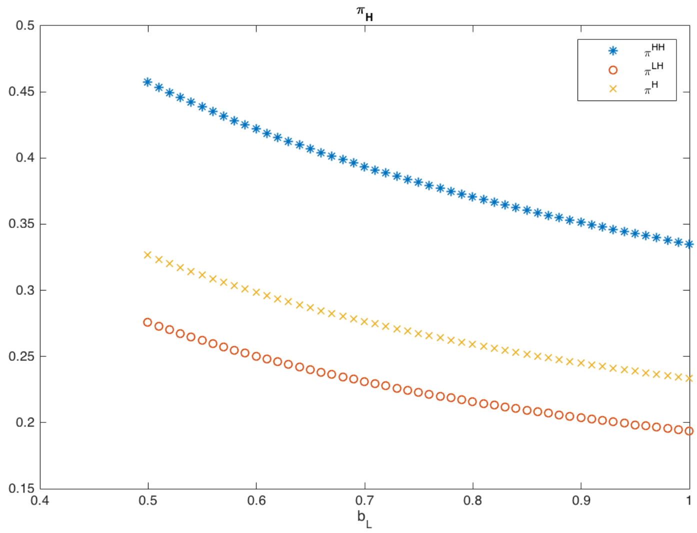


(b)

Figure 5: Simulation of equilibrium values for different levels of b^H. Parameter values: $\phi = 0.0011$, $\delta^{HL} = 0.6284$, $\delta^{LH} = 0.6741$, $q = 0.28$, $b^L = 0.49$, $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.32$, $\alpha^L = 0.22$, $\rho^{HH} = 0.09$, $\rho^{LL} = 0.12$, $\rho^{HL} = 0.10$



(a)



(b)

Figure 6: Simulation of equilibrium values for different levels of b^L. Parameter values: $\phi = 0.0011$, $\delta^{HL} = 0.6284$, $\delta^{LH} = 0.6741$, $q = 0.28$, $b^H = 0.54$, $\Delta V^H = 0.2794$, $-\Delta V^L = 0.1787$, $\alpha^H = 0.32$, $\alpha^L = 0.22$, $\rho^{HH} = 0.09$, $\rho^{LL} = 0.12$, $\rho^{HL} = 0.10$

Appendix A: Normalization Constant

Consider our link formation probability (1) for a general constant $c > 0$:

$$p_{ij} = \frac{1}{c^2}\delta_{ij}\rho_{ij}s_i s_j$$

where p_{ij} is the probability of forming the link ij, $\delta_{ij} = 1$ if i and j are of the same type and $\delta_{ij} = \delta_{ij} < 1$ if they are of different types. In this formulation, c is a normalizing constant representing the highest possible value of $\mathbf{s}$ independent of the network structure. We will show below how c is computed so that it guarantees that $p_{ij} \in [0, 1]$. Plugging this link probability into the utility function (3) leads to:

$$u_i = b_i s_i + \frac{\phi}{c^2}\sum_j \delta_{ij}\rho_{ij}s_i s_j - \frac{1}{2}s_i^2$$

First-order condition for individual i is given by:

$$s_i^* = b_i + \frac{\phi}{c^2}\sum_j \delta_{ij}\rho_{ij}s_j^*$$

In matrix form:

$$\mathbf{s} = \left(\mathbf{I} - \frac{\phi}{c^2}\mathbf{D}\right)^{-1}\mathbf{b}$$

where, as in the text, $\mathbf{D} = \widetilde{\mathbf{D}} \circ \mathbf{R}$, where $\mathbf{R}$ is the matrix of proximity indexes and $\widetilde{\mathbf{D}}$ is the matrix of δs. Notice that $\mathbf{D}$ have the same entries for agents of the same type and at the same distance. Thus heterogeneity in the link formation comes from heterogeneity in the vector of idiosyncratic characteristics and in the distance. Observe that

$$\left(\mathbf{I} - \frac{\phi}{c^2}\mathbf{D}\right)^{-1} = \sum_k^{\infty}\left(\frac{\phi}{c^2}\right)^k \mathbf{D}^k$$

Thus if $\phi < \frac{c^2}{\mu_1(\mathbf{D})}$, where $\mu_1(\mathbf{D})$ is the largest eigenvalue of $\mathbf{D}$, there exists a unique equilibrium. Notice that, in the most extreme case (where all preference-bias parameters as well as all the proximity indexes are equal to 1), $\mathbf{D} = \mathbf{U}$ is the matrix of ones with a zero diagonal. This is the highest value it can take. In that case, $\mu_1(\mathbf{D}) = n-1$, so that a sufficient condition for the existence of an equilibrium is $\phi < \frac{c^2}{n-1}$.

Let us now consider how c is computed. This is an upper bound for s_i. This upper-bound can be found by computing

$$\mathbf{s} = \left(\mathbf{I} - \frac{\phi}{c^2}\mathbf{D}\right)^{-1}\mathbf{b}$$

in the most favorable case for s, that is $\delta_{ij} = 1$ for all i and j, $b_i = 1$ for all i and $\rho_{ij} = 1$ for all i and j. Then it is immediate to see that c is defined as any of the identical entries of the vector

$$\mathbf{c} = \left(\mathbf{I} - \frac{\phi}{c^2}\left(\mathbf{U} - \mathbf{I}\right)\right)^{-1} \mathbf{1}$$

This is well defined when $\phi < \frac{c^2}{n-1}$. It is easily shown that

$$c = \frac{1}{1 - \frac{\phi}{c^2}(n-1)}$$

This equation has three solutions. In order to bound p_{ij} away from 1, we choose the highest solutions among these three solutions, which is:

$$c = \frac{1}{2}\left(1 + \sqrt{1 + 4\phi(n-1)}\right)$$

Notice that this value makes $\phi < \frac{c^2}{n-1}$ always satisfied. Finally, observe that $c \in \left[1, \left(1 + \sqrt{5}\right)/2\right]$ so that using $c = \left(1 + \sqrt{5}\right)/2 \simeq 1.618$, or c equal to any other larger scalar, makes $p_{ij} \in [0, 1]$. In the model, we take $c = 2 > \left(1 + \sqrt{5}\right)/2$ so that

$$p_{ij} = \frac{1}{4}\delta_{ij}\rho_{ij}s_i s_j$$

Appendix B: Description of Variables and Additional Tables

Table A1: Variable Definitions

Variable	Description
Individual socio-demographic variables	
Female	Dummy variable taking value one if the respondent is female.
Black or African American	Race dummies. "White" is the reference group
Other races	Race dummies. "White" is the reference group
Student grade	Grade of student at the time of Wave I interview
Religion practice	Response to the question: "In the past 12 months, how often did you attend religious services?", coded as 0= never, 1= less than once a month, 2= once a month or more, but less than once a week, 3= once a week or more.
Physical development	Response to the question: "How advanced is your physical development compared to other boys/girls your age?", coded as 1= I look younger than most, 2= I look younger than some, 3= I look about average, 4= I look older than some, 5= I look older than most
Two parent family	Dummy that takes the value 1 if respondent lives with two parents
Family size	Number of element in respondent's household
Parent unemployed	Dummy that takes the value 1 if respondent's parents are both unemployed
Parent worked hours	Hours worked by parents. Mean is considered if both parents are present.
Religion practice	Response to the question: "How often have you attended religious services in the past 12 months?", coded as 0= never, 1= a few times, 2= once a month, 3= 2 or 3 times a month, 4=once a week, 5=more than once a week.
Parent education	Schooling level of the (biological or non-biological) parent who is living with the child, distinguishing between "never went to school", "not graduated from high school", "high school graduate", "graduated from college or a university", "professional training beyond a four-year college", coded as 1 to 5. We consider only the education of the father if both parents are in the household.
Family income	Yearly family income in thousands USD, as asked to the parents. Missing answers are treated as zeros
Family income refused	To account for possible reporting bias, this dummy takes the value 1 when the respondent didn't answer.
Socialization variables	Defined in the text.
Parental effort	Defined in the text.

Table A2: Summary statistics for socialization variables

	Mean	St. dev.	Min.	Max
Extra activities at school				
School clubs	0.3991	0.8244	0	6
Dance	0.0718	0.2582	0	1
Music	0.1820	0.4447	0	3
Sport	0.7350	1.1077	0	4
Newspaper	0.0352	0.1842	0	1
Honors	0.0741	0.2619	0	1
School council	0.0613	0.2399	0	1
Yearbook	0.0654	0.2473	0	1
Extra activities outside school				
Rollerblades	0.2076	0.3119	0	1
Sport (out of school)	0.4622	0.3817	0	1
Exercise	0.5454	0.3518	0	1
Hangout with friends	0.6582	0.3362	0	1
Socialization in the neighborhood				
Know people	0.7053	0.4559	0	1
Stopped to talk	0.7843	0.4113	0	1
Fitness	0.2040	0.4030	0	1

Sample size is 5,281

Appendix C: Equilibrium Behavior and Transition Probabilities

We have the following proposition:

Proposition 3 *For any parameter $x \in \{q, b_i, \phi, \delta_{ij}, \rho_{ij}\}$, we have:*

$$Sign|\frac{\partial \pi^{HH}}{\partial x}| = Sign|\frac{\partial h_i^H}{\partial x}|$$

and

$$Sign|\frac{\partial \pi^{LH}}{\partial x}| \neq Sign|\frac{\partial h_i^L}{\partial x}|$$

Proof. The proof stems from simple algebra. Consider

$$\boldsymbol{\tau}^{H*} = \frac{1}{2}\Delta V^H \mathbf{1} + \frac{1}{2}[\alpha^H - \Delta V^H]\mathbf{h}^H$$

and

$$\pi_i^{HH} = \tau_i^H + (1 - \tau_i^H)h_i^{H*}$$

Then it is immediate to have

$$\frac{\partial \pi_i^{HH}}{\partial x} = \frac{\partial \tau_i^{H*}}{\partial x} + (1 - \tau_i^{H*})\frac{\partial h^{Hi*}}{\partial x} - h^{Hi}\frac{\partial \tau_i^{H*}}{\partial x}$$

where

$$\frac{\partial \tau_i^{H*}}{\partial x} = \frac{1}{2}(\alpha_H - \Delta V_H)\frac{\partial h^{Hi*}}{\partial x}$$

By substituting equilibrium efforts we get

$$\frac{\partial \pi_i^{HH}}{\partial x} = \frac{\partial h^{Hi*}}{\partial x}\left[\frac{1}{2}(\alpha_H - \Delta V_H) + 1 - \frac{1}{2}\Delta V_H - (\alpha_H - \Delta V_H)h^{H^i}\right]$$

then,

$$Sign|\frac{\partial \pi^{HH}}{\partial x}| = Sign|\frac{\partial h_i^H}{\partial x}|$$

if and only if the term in square brackets is positive. Solving for h^{H^i} we get that

- If $\alpha_H > \Delta V_H$ then $Sign|\frac{\partial \pi^{HH}}{\partial x}| = Sign|\frac{\partial h_i^H}{\partial x}|$ if and only if $h_i^H < \frac{1 - \Delta V_H + \frac{\alpha_H}{2}}{\alpha_H - \Delta V_H}$;
- if If $\alpha_H < \Delta V_H$, then $Sign|\frac{\partial \pi^{HH}}{\partial x}| = Sign|\frac{\partial h_i^H}{\partial x}|$ always.

Notice now that

$$\frac{1 - \Delta V_H + \frac{\alpha_H}{2}}{\alpha_H - \Delta V_H} > 1$$

if $\alpha_H < 2$ which is always the case. Then it is always true that If $\alpha_H > \Delta V_H$ then

$$Sign|\frac{\partial \pi^{HH}}{\partial x}| = Sign|\frac{\partial h_i^H}{\partial x}|$$

In a very similar way consider

$$\boldsymbol{\tau}_L^* = \frac{1}{2}\alpha_L \cdot \mathbf{1} + \frac{1}{2}[\Delta V_L - \alpha_L]\mathbf{h}^L \tag{25}$$

and

$$\pi_i^{LH} = \tau_i^L + (1 - \tau_i^L)(1 - h_i^{L*})$$

Then

$$\frac{\partial \pi_i^{LH}}{\partial x} = \frac{\partial \tau_L^i *}{\partial x} - (1 - h^{L^i})\frac{\partial \tau_H^i *}{\partial x} - (1 - \tau_L^i *)\frac{\partial h^{L^i} *}{\partial x} \tag{26}$$

where

$$\frac{\partial \tau_L^i *}{\partial x} = \frac{1}{2}(\Delta V_L - \alpha_L)\frac{\partial h^{H^i} *}{\partial x} \tag{27}$$

by substituting equilibrium efforts we get

$$\frac{\partial \pi_i^{LH}}{\partial x} = \frac{\partial h^{L^i} *}{\partial x}\left[\frac{1}{2}(\Delta V_L - \alpha_L)h^{L^i} - 1 + \frac{1}{2}\alpha_H + \frac{1}{2}(\Delta V_L - \alpha_L)h^{L^i}\right] \tag{28}$$

then $Sign|\frac{\partial \pi^{LH}}{\partial x}| = Sign|\frac{\partial h_i^L}{\partial x}|$ if and only if the term in square brackets is positive. Solving for h^{L^i} we get

- If $\alpha_H > \Delta V_H$ then $Sign|\frac{\partial \pi^{LH}}{\partial x}| \neq Sign|\frac{\partial h_i^L}{\partial x}|$ always;
- if If $\alpha_H < \Delta V_H$, then $Sign|\frac{\partial \pi^{LH}}{\partial x}| = Sign|\frac{\partial h_i^L}{\partial x}|$ if only if $h_i^H > \frac{1 - \frac{\alpha_L}{2}}{\Delta V_L - \alpha_L}$

Notice now that $\frac{1 - \frac{\alpha_L}{2}}{\Delta V_L - \alpha_L} > 1$ whenever $\alpha_L > 2(\Delta V_L - 1)$ which is always the case, then it is impossible to satisfy the condition $h^{H^i} > \frac{1 - \frac{\alpha_L}{2}}{\Delta V_L - \alpha_L}$, and the result in the proposition holds. Q.E.D.

This proposition shows how parental effort and homophily balance each other to shape the transition probabilities of becoming educated. In particular, Proposition 3 shows that the sign of the effect is totally independent of the ordering among α and ΔV and independent of τ as well. Notice that this does not mean that parental effort does not play any role. It just states that parental efforts either go in the same direction as homophily or it is not strong enough to revert the direction imposed by homophily. However, the magnitude is strongly affected by the effort τ.