Targeted socialization and production*

Facundo Albornoz† Antonio Cabrales‡ Esther Hauk§

June 7, 2016

Abstract

We study a model that integrates productive and socialization efforts with network choice and parental investments. We characterize the unique symmetric equilibrium of this game. Individuals underinvest in productive and social effort. However, solving only the investment problem can exacerbate the misallocations due to network choice, to the point that in the presence of congestion effects the intervention may generate an even lower social welfare than no intervention at all. We also study the interaction of parental investment with network choice. In many scenarios, intergenerational transmission of abilities leads to a tendency towards to conformism, which aggravates potential problems of network overpopulation. We relate our equilibrium results with the existing evidence on parental occupational transmission.

JEL-Classification: D85, I20, I28, J15, J24. J61, J62.

Key-words: peer effects, network formation, parental involvement, intergenerational mobility, cultural identity, immigrant sorting.

*Hauk acknowledges financial support from the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Program for Centers of Excellence in R&D and through CICYT project number ECO2015-67171-P and from the government of Catalonia (2014 SGR 1064). We thank participants at the Barcelona Graduate School of Economics Summer Forum on Economics of Socialization, the Vienna Workshop on Networks and Social Norms, the PEPA-CEMMAP Workshop on social networks, the Murat Sertel lecture at the ASSET conference, and seminar audiences at the universities of Bath, Bonn, Bristol, Leicester, Louvain-la-Neuve, San Andrés and York for helpful comments. Pablo Warnes and Ezequiel García Lembergman provided outstanding assistance.

†University of Nottingham and CONICET; facundo.albornoz@nottingham.ac.uk

‡University College London; email: a.cabrales@ucl.ac.uk

§Instituto de Análisis Económico (IAE-CSIC) and Barcelona Graduate School of Economics; email: esther.hauk@iae.csic.es

1 Introduction

Many productive processes are mediated by social interaction. The accumulation of human capital (Moretti, 2004), innovation activities (Cassiman and Veugelers, 2002), and crime (Glaeser, Sacerdote, and Scheinkman, 2003), are all affected by the actions and abilities of others around us. When social interactions have productive consequences, economic agents do not only devote a considerable effort to develop social interactions but also to interact with the "right" individuals. Until now, the literature has explored these two efforts separately. On the one hand, Benabou (1993) sought to understand the process of selecting the best neighborhood to profit from spillovers.[1] On the other hand, Cabrales, Calvó-Armengol, and Zenou (2011) studied the interaction between (undirected) socialization and production efforts, but did so within the confines of a single network.

This paper examines how individuals make optimal decisions about matching with the best possible group in terms of enhancing their productive ability, as well as about the intensity of socialization and their productive effort within the chosen group. We model these choices in a tractable framework that allows for a complete equilibrium and welfare analysis of individual decisions and generates novel results with implications for policy interventions. Although the model has a variety of potential applications, we focus on human capital acquisition in environments where individuals have diverse backgrounds and abilities.

Our model has three main components. First, it recognizes that there are complementarities in productive investments (direct human capital acquisition) within networks. We assume multiplicative spillovers between an agent's effort and those of other members of his network. One can view this spillover as the result of information sharing between the learners, which implies that the individual marginal productivity with respect to one's own stock of human capital increases linearly in the knowledge stock of other network members. An individual's return on productive effort is idiosyncratic and can vary across

[1]This initial study has been expanded upon by a recent economic literature on social networks. See e.g. books by Goyal (2012), Jackson (2010) or Vega-Redondo (2007).

networks. Second, the amount of spillover also depends positively on the costly socialization efforts of the individual and his network members. That is, taking advantage of others' information requires socialization and relies on the socializing intensity of all network members. One's incentives to socialize increase with the information everyone has, which implies complementarity between socialization and productive investments. Finally, the agents also decide which network to join. This choice depends on their relative abilities to gain surplus in the different networks. Since these relative abilities are important, we also devote a section to explore the allocation of parental effort to improve/expand network-specific abilities.

The decision makers in the model can be seen as young people building human capital. They can choose to invest in the Mainstream (M) or the Alternative (F) network. Within each group, the young person can exert two types of costly effort: *productive effort* (devoting time to learning the skills necessary to the main activity of the network) and *socialization effort* (going to bars, libraries, sport clubs, or any activity that involves other young people who are also developing skills). Within a network, each youngster does not decide with whom he interacts, i.e. meetings are random, but the venue (say a social club or a bar) determines the people with whom he is likely to interact. The activities serve to share information which improves future production, be it through shared knowledge or trust relations that are indispensable in any productive activities. The fact that socialization is random within groups makes our analysis more tractable than other models of social network creation. This allows us to use standard Nash equilibrium analysis. However, the fact that agents also choose their network enhances our ability to evaluate realistic implications and connect them to the scarce available empirical evidence on occupational mobility.

We now summarize our results. First, we fully characterize the unique symmetric equilibrium,[2] both in terms of socialization and production effort as well as for network choice. From the equilibrium characterization, a corollary follows: the average socialization in a group is increasing in the average type

[2]The equilibrium is symmetric in the sense that agents with the same type choose the same options.

of the group.[3]

Second, we compare the equilibrium outcomes with those that would be chosen by a utilitarian social planner. As expected in a model with positive complementarities, the decentralized outcomes exhibit under-investment in both socialization and production (Proposition 2). The results on network choice are more subtle. When individual productivities are uniformly distributed, there are more people than the socially desirable number in the network whose distribution of types have a larger mean (Proposition 3). This is noteworthy because it is the *a priori* more productive network that is overpopulated with decentralized sorting and a uniform distribution of individual productivities. If we think of networks as different labor markets, our result implies that, contrary to conventional wisdom, there could be too much integration into a mainstream labor market that has a productivity advantage over the alternative labor market. The reason for this result is that the more productive mainstream network creates stronger positive spillovers than the less productive alternative network. The mainstream network therefore also attracts types whose relative network-specific productivity in the mainstream network is fairly low and who have a relative productivity advantage in the alternative network. Moving these types to the alternative labor market improves average welfare of everybody remaining in the mainstream labor market. Notice, however, that despite their relative productivity advantage in the alternative market, the relative productivity of those choosing to join the alternative market is even higher. Hence individuals who voluntarily sorted themselves into the alternative market might be adversely affected when these lower types are moved into the alternative market. Under a uniform distribution this second effect is close to zero, therefore the mainstream network is overpopulated.

The overpopulation result depends on the distribution of talents. For alternative distributions this result still occurs but does not hold over the full

[3]This empirical implication of the model is consistent with Currarini, Jackson, and Pin (2009) where the number of interactions within friendship networks are increasing in size, and Albornoz, Cabrales, and Hauk (2014) who find that more productive economics fields are characterized by higher levels of co-authorship.

range of parameter values.[4] For all possible distribution functions the general message is that individuals fail to sort themselves into networks in a socially optimal way. This imposes novel constraints to the optimality of policy intervention. To explore this further, we show that a government operating on one margin only, inducing efficient socialization and production effort within a network, may harm global efficiency by exacerbating misallocation due to network choice. In the presence of within network congestion effects, this policy intervention may reduce social welfare with respect to a situation with no policy intervention at all. This somewhat extreme outcome reveals an important novel point of our paper. The fact that individuals do not only choose their efforts within networks, but also the networks to which they belong (in a sense the intensive and extensive margins of socialization) makes policy design more challenging: there needs to be a coordination between the local within-network choice and the overall process of network selection.

The implications of our result about sub-optimal one-margin policy interventions are far-reaching. For example, consider the case of public funding of science. Government subsidies for scientific endeavors are sometimes done by disciplinary bodies. They incentivize both production (by subsidizing partially or totally the inputs into the research process) and socialization (e.g. funding for conferences and workshops) within broad scientific fields. Our results show that different funding bodies need to coordinate their activities in order to avoid excessive congestion in particularly well endowed fields. In this sense, our results suggest that the UK situation, with seven funding bodies for science[5] may lead to less efficient outcomes than the US which only has the NSF.[6]

Third, we explore an extension of our model where parents invest in de-

[4]In Appendix A.4, we provide conditions for overpopulation of the mainstream network when productivities are distributed according to a Pareto distribution.

[5]EPSRC, Engineering and Physical Sciences Research Council; MRC, Medical Research Council; BBSRC, Biotechnology and Biological Sciences Research Council; NERC, Natural Environment Research Council; STFC, Science and Technology Facilities Council; ESRC, Economic and Social Research Council; AHRC, Arts and Humanities Research Council.

[6]A recent report to the UK government was "believed to be considering various options, including consolidating several of the councils or appointing a single official to oversee the budgets of all of them. See Brumfiel (2013).

veloping their children' network-specific abilities (e.g. training of children for specific occupations before they go into the labor market). Without parental investment a child's ability in a network is a random draw from an initially given uniform distribution: the upper bound of this distribution might be different for the parental and non-parental network. Parental investment towards training of their child in a given network will raise the upper bound of the distribution from which the child's network-specific ability is drawn, hence generating a first order stochastically dominating shift in the distribution of the child's ability. The first result is again a full characterization of the symmetric equilibrium of the parental investment game. We establish existence and uniqueness of this equilibrium and show that parents will invest in enhancing their child's ability in only one of the networks. We then examine the conditions under which parents from both networks choose to enhance their child's ability in the same network regardless of their origin. We show that if the initial distribution of abilities across networks is the same in both networks, parents will put all of their effort to enhance their child's ability in the *a priori* most productive network (Proposition 8). This tendency of "conforming" to the more productive network can also arise under asymmetries in the initial distribution of abilities across networks. In particular, if the initial endowment of ability is higher in the family network than in the alternative one, parents will invest in their child's ability in the more productive network if they had sufficient investment resources to revert the initial influence of the parental network. This is an additional force towards overpopulation of the more productive mainstream network M and can thereby increase the cost of congestion, which as we have argued earlier, may also have negative feedback effects if policies are not well coordinated between networks. Since the direction of the intergenerational transmission of network-specific abilities depends on parents' time endowments and relative network influence, our model points to different forms of intervention should governments want to affect the tendencies towards or against conformism.

We also examine a situation in which there is a dominant network with a higher initial endowment for all groups. This may happen even if the dominant network is not the most productive one. We refer to this case as a "cultural trap" which can occur as a result of inculturation via education or mass media. If the pattern of choices is inefficient, our model highlights the importance of providing extra resources to the inefficiently underpopulated network.[7]

Our model predicts that intergenerational network persistence is more likely when the initial endowment of ability is higher in the family network than in the alternative network. Moreover, this intergenerational network persistence should be higher in the more profitable networks even if initial endowments of network abilities are similar. We end this paper by showing that these theoretical predictions are consistent with the evidence provided by the literature on intergenerational occupational mobility, where it seems reasonable to assume an initial ability advantage in the parental network.[8] This literature finds a high persistence of occupational categories within the family across all countries studied. Moreover, the probability for an individual to fall within the same occupational category as her/his parent is increasing in the size of this occupational category (more network externalities). Also, intergenerational job persistence is higher in the more profitable jobs (network) and switchers tend to move into more profitable jobs.[9]

This paper is organized as follows. Section 2 describes the model. Section 3 contains the equilibrium and welfare analysis. Section 4 describes the effects of parental effort and relates our results to evidence from the occupational mobility literature. Section 5 discusses some additional relevant literature. Section 6 concludes. Most proofs are gathered in the Appendix.

2 The model - payoffs

We consider an economy with a continuum of heterogeneous individuals that choose their productive group/network. There are two different networks with local complementarities in productive investment, M and F. Each individual i has a network specific individual productivity b_i^n for $n \in \{M, F\}$ which

[7] We also analyze for the sake of completeness the less plausible case where the parental networks is in all cases less dominant than the non-parental network.

[8] See Long and Ferrie (2013) for the U.S. and U.K., Azam (2013) for India, Binzel and Carvalho (2013) for Egypt and Knoll, Riedel, and Schlenker (2013) for Germany.

[9] Details are given in Section 4.3.

is randomly and independently drawn for each network. These abilities are distributed uniformly and independently in $[0, B^n]$ for $n \in \{M, F\}$.[10]

After agents have chosen their network all agents within the same network simultaneously decide their direct productive effort k_i^n and their socialization effort s_i^n. Socialization activities allow them to take advantage of productive efforts made by the other members of the network. Consequently, the payoff within a particular network n is the sum of two components, a private component P_i^n, and a synergistic component S_i^n derived from the interactions. The private component P_i^n has a linear-quadratic cost-benefit structure and is given by

$$P_i^n = b_i^n k_i^n - 1/2 \left(k_i^n\right)^2 .$$

The synergistic component in our model, S_i^n has the feature that socialization is required to take advantage of the network externality, which is due to the complementarity in productive efforts. In addition, the socialization within each network is undirected.[11] Specifically, this means that within networks agents only choose the amount of interaction s_i, but not the identity of the individuals with whom they interact. However, we allow individuals to choose the group of people with whom to socialize (the network). This is the way in which socialization often occurs in reality: individuals choose the neighborhoods where to live, the schools or colleges to attend, and the social ties therein are mostly the result of random events. Researchers go to conferences, businesspeople go to fairs, and synergistic effort is mostly generic within the conference or fair attended; but clearly both researchers and businesspeople carefully choose which conference / fair to attend.

Denoting by $\mathcal{N}_i$ be the network to which individual i belongs, synergistic

[10]We also study the case of abilities distributed according to a Pareto distribution. See Appendix A.4.

[11]Both features of the synergistic payoffs are shared by the model in Cabrales, Calvó-Armengol, and Zenou (2011). However, we propose a different functional form for the benefits from synergistic returns. We will show that using our synergistic component S_i^n leads to a game with a *unique* symmetric equilibrium within a network, while the game in Cabrales, Calvó-Armengol, and Zenou (2011) has multiple equilibria. Equilibrium uniqueness in socialization and productive efforts facilitate our analysis of directed network choice.

returns are given by

$$S_i^n = ab_i^n \left(k_i^n\right)^{1/2} \int_{j\in\mathcal{N}_i} \left(b_j^n \left(k_j^n\right)^{1/2} g_{ij}^n(\mathbf{s})\right) dj - \frac{1}{2}\left(s_i^n\right)^2,$$

where the parameter a captures the overall strength of synergies, $\mathbf{s}$ is the profile of all socialization efforts and $g_{ij}^n(\mathbf{s})$ is the link intensity of individual i and j which we will define below. Each network is composed by a continuum of individuals $\mathcal{N}^n \subset \mathbb{R}$ for $n \in \{M, F\}$, where the measure of the set $\mathcal{N}^n$ is N^n.

Observe that synergistic returns are multiplicative in individual productivity parameters and in the square root of productive efforts additively separable by pairs, hence productive efforts are complementary.[12] The specific functional form implies that synergistic returns are symmetric in pairwise productive efforts and that the synergistic returns exhibit constant returns to scale to overall productive efforts. Similar assumptions are imposed on the link intensity which captures to which extent individuals take advantage of these productive network externalities. These assumptions are:[13]

(A1) Symmetry: $g_{ij}^n(s_i^n, s_j^n) = g_{ji}^n(s_j^n, s_i^n)$, for all i, j, n;

(A2) The total interaction intensity of individual i in network n exhibits constant returns to scale to overall inputs in socialization efforts and symmetry: $\int_{j\in\mathcal{N}_i} g_{ij}^n(s_i^n, s_j^n)dj = \frac{1}{N^n}\int_{j\in\mathcal{N}_i} \left(s_i^n\right)^{1/2}\left(s_j^n\right)^{1/2} dj$;

(A3) Anonymous socialization: $g_{ij}^n(s_i^n, s_j^n)/\left(s_j^n\right)^{1/2} = g_{ki}^n(s_k^n, s_i^n)/\left(s_k^n\right)^{1/2}$, for all i, j, k;

These assumptions imply a specific functional form of $g_{ij}^n(s_i^n, s_j^n)$, which we state in the following result:

[12] Complementarity in productive returns in Cabrales, Calvó-Armengol, and Zenou (2011) is modeled by synergistic returns being multiplicative in productive efforts and additively separable by pairs.

[13] While Cabrales, Calvó-Armengol, and Zenou (2011) also model symmetric and anonymous socialization, which is the key for generic socialization, they assume that link intensity satisfies aggregate constant returns to scale.

Lemma 1. *Suppose that, for all* $\mathbf{s} \neq \mathbf{0}$*, the link intensity satisfies assumptions* $(A1)$*,* $(A2)$ *and* $(A3)$*. Then, the link intensity is given by*

$$g_{ij}^{n}(s_i^n, s_j^n) = \frac{1}{N^n} \left(s_i^n\right)^{1/2} \left(s_j^n\right)^{1/2} \tag{1}$$

Proof of Lemma 1: Fix $\mathbf{s}$. Combining $(A1)$ and $(A3)$ gives

$$\left(s_k^n\right)^{1/2} g_{ij}^{n}(s_i^n, s_j^n) = \left(s_j^n\right)^{1/2} g_{ij}^{n}(s_i^n, s_k^n)$$

Integrating across all j's and using $(A2)$ gives $g_{ij}^{n}(s_i^n, s_k^n) = \frac{1}{N^n} \left(s_i^n\right)^{1/2} \left(s_k^n\right)^{1/2}$. ▮

Notice that given $(A2)$ and given a level of socialization effort for all members of the network, total socialization of an individual in a network $\int_{j \in \mathcal{N}_i} g_{ij}^{n}(s_i^n, s_j^n) dj$ is independent of the size of the network. In other words, individuals will not have more contacts in larger networks if all their members choose the same s_i^n independent of size. One could easily accommodate other assumptions, where socialization is either easier or more difficult in larger networks by using $1/\left(N^n\right)^{\beta}$ for some β different from 1.

Combining the private returns and the network externality yields individual payoffs in network i as:

$$\begin{aligned} u_i^n &= P_i^n + S_i^n \\ &= b_i^n k_i^n + a b_i^n \left(k_i^n\right)^{1/2} \int_{j \in \mathcal{N}_i} \left(b_j^n \left(k_j^n\right)^{1/2} g_{ij}^{n}(\mathbf{s})\right) dj - \frac{1}{2}\left(k_i^n\right)^2 - \frac{1}{2}\left(s_i^n\right)^2 \end{aligned} \tag{2}$$

We assume that individuals can only belong to one single network. This assumption is consistent with a number of potential applications: most people have only one profession to which they dedicate themselves; academics generally do not work simultaneously in very distinct fields; top athletes generally only excel in one sport; and in spite of "Ingres' violin" the same thing generally holds for artists. It can also be justified formally within the model in a variety of ways. For example, by adding a sufficiently large fixed cost to join a network which could arise from training costs.

Finally, the timing of events is as follows: each individual i first chooses in which network to participate, and then takes the decisions over k_i and s_i simultaneously.

3 The equilibrium

We solve the game by backward induction. We compare the individual optimum with the social optimum in which a social planner maximizes the sum of individual utilities. We first solve for the optimal efforts within a network and then let individuals sort themselves (or be sorted by a social planner) into the networks.

3.1 Choice of production and socialization efforts

For each individual we have to find the optimal productive and socialization effort within each network (we suppress the superindex referring to the network when there is no ambiguity). For the individual choice problem - the decentralized problem - this is the choice of k_i and s_i that maximizes (2). The social planner, on the other hand, chooses k_i^s and s_i^s to maximize the sum of individual utilities given by

$$\int_{i\in\mathcal{N}^M\cup\mathcal{N}^F} u_i(b_i)di = \int_{i\in\mathcal{N}^M\cup\mathcal{N}^F}\left(b_ik_i + ab_i\sqrt{k_is_i}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj - \frac{1}{2}k_i^2 - \frac{1}{2}s_i^2\right)di \tag{3}$$

Proposition 1. *Let $a^2\overline{b^2}^2 < 1$. Then both the individual choice problem and the social planner choice problem have a unique (interior)*[14] *solution which for each individual depends on the individual's own productivity and is multiplicative in a parameter common to all individuals in the network. Hence*

$$k_i = b_ik \text{ and } s_i = b_is \text{ for all } i \tag{4}$$

$$k_i^s = b_ik^s \text{ and } s_i^s = b_is^s \text{ for all } i \tag{5}$$

[14]The individual choice problem also has a trivial partial corner solution where $s_i = 0$. If nobody socializes, socialization is not profitable. However, this equilibrium is not stable, since the marginal utility of s_i is positive for any (even infinitesimally small) average level of socialization in the network. We therefore ignore this solution in our analysis.

where the optimal common network parameters for productive and socialization effort are given by

$$k = \frac{4}{4 - a^2\overline{b^2}^2} \tag{6}$$

$$s = \frac{2a\overline{b^2}}{4 - a^2\overline{b^2}^2} \tag{7}$$

for the individual choice problem and by

$$k^s = \frac{1}{1 - a^2\overline{b^2}^2} \tag{8}$$

$$s^s = \frac{a\overline{b^2}}{1 - a^2\overline{b^2}^2} \tag{9}$$

for the social planner.

The resulting individual utilities are

$$u_i(b_i) = 2b_i^2 \frac{\left(4 + a^2\overline{b^2}^2\right)}{\left(4 - a^2\overline{b^2}^2\right)^2} \tag{10}$$

in the individual choice problem and

$$u_i^s(b_i) = \frac{b_i^2}{2}\left(\frac{1}{1 - a^2\overline{b^2}^2}\right). \tag{11}$$

for the social planner solution.

Proof. See Appendix □

From Proposition 1, it is easy to see that individuals fail to internalize the positive externality of their investment decisions on the other members of their network. Therefore, the individual utility resulting from the decentralized solution (10) is lower than the individual utility resulting from the social planner solution (11):

Proposition 2. *Individuals underinvest in both productive and socialization effort ($k^s > k$ and $s^s > s$)*

The common network parameters are increasing in the network parameter a and average network squared productivity $\overline{b^2}$ and hence in average network productivity $\overline{b}$. Since individual socialization is $s_i = b_i s$, average socialization is $\overline{b}s$. As a corollary of Proposition 1 we then have

Corollary 1. *Average socialization, $\overline{b}s$, is increasing in $\overline{b}$.*

Corollary 1 implies that individuals within more productive networks socialize more on average, an empirical implication of our model which is consistent with the evidence presented in Currarini, Jackson, and Pin (2009) showing that the number of interactions within friendship networks are increasing in size. In Albornoz, Cabrales, and Hauk (2014) we provide some further empirical evidence for this prediction based on analysis of co-authorships within economics fields. Academic life is clearly an example of a situation in which an individual's productive outcomes are affected by the abilities and activities of other researchers involved in the same production process. Hence socialization decisions become key productive choices. Moreover academics choose their field of research: their network. Using data scrapped from the IDEAS-RePEc website Albornoz, Cabrales, and Hauk (2014) establish that economic researchers who work in more productive fields[15] tend to have more co-authors.[16]

3.2 Choice of network

Having found the second-stage utilities, we can now solve the first-stage in which individuals sort themselves into one of the two networks M and F. We will show now that independently of whether productive or socialization efforts within the network are chosen by individuals (decentralized solution) or by the social planner, there is a unique dividing line $b_i^M = Cb_i^F$ such that

[15] Albornoz, Cabrales, and Hauk (2014) use JEL identifiers at the uppermost level to associate an author with a field. For every individual author, they construct a vector with the sum of all of the JEL information contained in her papers, divided by field. An author's field is defined as the one in which she has the maximum value in this vector. Field productivity is measured by the share of top 10% of the IDEAS-RePEc authors in the field.

[16] The measure for co-authorship is constructed scrapping data from CollEc, a RePEc service of rankings by co-authorship centrality for authors registered in the RePEc Author Service.

individuals who fall below the line will choose network F, while individuals above the line will choose network M. Such a dividing line implies that

$$\begin{aligned} \overline{b^{M^2}} &= E\left(b_i^{M^2} \left| b_i^M > C b_i^F \right.\right) \\ \overline{b^{F^2}} &= E\left(b_i^{F^2} \left| b_i^M < C b_i^F \right.\right) \end{aligned}$$

We will denote the slope of the dividing line by C^P if effort choices in the networks are decentralized and by C^E if the social planner implements efficient effort choices in the networks.

When deciding which network to join, individuals take the network choices of others as given and choose the network that grants them the maximal utility given optimal investment choices within the network, which could result from the decentralized or the centralized solution derived in the previous section.

Under the decentralized solution, individuals choose network M if and only if $u_i(b_i^M) \geq u_i(b_i^F)$ hence whenever

$$2b_i^{M^2} \frac{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2} \geq 2b_i^{F^2} \frac{\left(4 + a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4 - a^{F^2}\overline{b^{F^2}}^2\right)^2} \tag{12}$$

If the dividing line exists it is defined when both terms of (12) are equal or equivalently when

$$b_i^M = b_i^F \sqrt{\frac{\left(4 + a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)} \frac{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2}{\left(4 - a^{F^2}\overline{b^{F^2}}^2\right)^2}} = b_i^F C_P \tag{13}$$

Hence C_P is given as the fixed point of

$$C_P = \sqrt{\frac{\left(4 + a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)} \frac{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2}{\left(4 - a^{F^2}\overline{b^{F^2}}^2\right)^2}} \tag{14}$$

If s^s and k^s are induced (say via subsidies) by the social planner, people would choose network M if and only if $u_i^s(b_i^M) \geq u_i^s(b_i^F)$ and the dividing line should it exists would solve

$$C_E = \sqrt{\frac{1 - a^{M2}\overline{b^{M2}}^2}{1 - a^{F2}\overline{b^{F2}}^2}} \tag{15}$$

Lemma 2. *Both C_P defined by (14) and C_E defined by (15) exist and are unique.*

Proof. See Appendix A.2. □

We now check whether the decentralized and centralized networks reach a social optimum. The following results show that in both cases the social planner would choose a cutoff that lies to the right of the cutoff chosen by the individuals, i.e. $C_E^* > C_E$ and $C_P^* > C_P$ where C_E^* is the cutoff a social planner would choose when effort choices in the networks are centralized while C_P^* is the cutoff the social planner would choose when effort choices in the networks are decentralized.

Proposition 3. *If $B^M > C_E B^F$, social welfare is increasing in C for all $C \leq C_E$ and $C \leq C_P$.*

Proof. See Appendix A.3. □

Proposition 3 implies that with a uniform distribution of individual talent, too few people join the F network, independently of whether there is a social planner or not.[17] Interestingly, when people freely sort themselves into networks, it is the more efficient network that becomes overpopulated. For example, consider an immigrant who has to decide between integrating into the mainstream labor market or remaining within the immigrant labor network, which is less efficient overall. Contrary to popular wisdom, Proposition 3 implies excessive integration into the mainstream labor market. A similar claim could be made for the scientific community that has to sort themselves into theoretical and applied research activities. In the case of uniformly distributed

[17] The result was derived under the assumption that when $B^M > C_E B^F$ or $B^M > C_P B^F$ respectively. If the corresponding assumptions were violated we would get the opposite, namely an underpopulated M network.

individual productivities and assuming that productivity in applied research has a higher upper bound than in applied, the applied research community will be overpopulated and too few people will voluntarily classify themselves as theorists.

One can explain the result of proposition 3 as follows. At either C_E or C_P, individuals at the margin are indifferent between both networks. For this reason, moving them from one network to the other does not affect their welfare. However, it affects average welfare in these networks by affecting the average network type. For the F network this effect is almost non-existing because average type does not depend on C under the uniform distribution (see equation (41) in Appendix A.2). In the M network the average type improves with C (see equation (43) in Appendix A.2) and hence average welfare improves when the indifferent and close to indifferent M-types in the network are moved to the F-network. Society would be better off had they joined the less efficient F network. This occurs independently of whether productive and socialization efforts are generated in a socially optimal way, or in a decentralized way.

However, the stark result that the efficient network is overpopulated hinges on the assumption that individual productivities are uniformly distributed. Appendix A.4 illustrates that social welfare might be increasing or decreasing in C at $C = C_E$ when individual productivities follow a Pareto distribution. This happens because with a Pareto distribution the average type in the less efficient F network decreases with C while the average type in the more efficient M network increases with C.[18] Hence when moving the close to indifferent M-types to the F network, average welfare in the F network falls because average type decreases while average welfare in the M network increases because average type increases. The overall effect on social welfare is therefore ambiguous. It also depends on the parameters a^F and a^M that capture the overall strength of the synergies in the networks. Overpopulation of the more efficient network though does occur in many cases.[19]

[18] This can be seen by differentiating equations (58) and (59) in Appendix A.4 with respect to C.

[19] In the appendix A.4 we give a sufficient condition for the M network to be overpopulated with a Pareto distribution, when the distribution's tail is sufficiently thin. We also show that for constant $\frac{a^{M^2}}{a^{F^2}} < 1$ underpopulation of the more efficient M network occurs if synergistic

Although the direction of overpopulation is conditional on the distribution of abilities, the general message of our analysis is relevant in its own right: decentralized network selection involves sub-optimal network composition. Thus, the social optimum is achieved when the central planner intervenes at both margins, i.e. she/he induces optimal efforts within the networks and chooses the optimal dividing line C_E^*. By Proposition 3 $C_E^* > C_E$, however, the proposition is silent towards the position of C_E^* with respect to C_P. This is an interesting question, since C_E requires intervention by the social planner when choosing productive and socialization efforts within a network, while C_P is the cutoff chosen by individuals in the absence of any intervention. Can no intervention be better than intervening at one margin only? We turn to this question in the following section.

3.3 Global efficiency

In this section we first show that inducing the optimal socialization and production effort can induce over-congestion. Then we show that if congestion is costly, local efficiency for a given network would reduce global efficiency.

3.3.1 Optimal socialization and production effort can induce excessive congestion.

We want to show that there are parameter values for which a decentralized choice of network, together with an optimal choice of socialization and production efforts can lead to over-congestion in the mainstream network M.

Proposition 4. *Suppose $a^M = a^F$ and $B^M = B^F + \varepsilon$. For ε small enough we have that $C_E < C_P < 1 < C_E^*$.*

Proof. Suppose first that $a^M = a^F$ and $B^M = B^F$. It is easy to see that in the absence of any asymmetries $C_E^* = C_E = C_P = 1$. It is then optimal, both socially and individually, for individuals to sort into the network where they have the higher productivity draw b_i^j. Now, suppose we give a small advantage to network M, by increasing B^M. Since the F parameters are left unchanged

returns are sufficiently low (a^F sufficiently small).

also the utility for an individual from joining the F network is unaffected (since average type in the F network is independent of C). From the definitions of C_E and C_P it is easy to see that increasing B^M will reduce C_E and C_P: for a fixed C, the right hand side of both defining equations (14) and (15) decrease if B^M increases. Hence to preserve the equality, the variable C has to fall. An increase in B^M increases average type in the M network, and therefore it draws more people into this network. However, individuals with lower b_i^M than before the increase in B^M are now drawn into the M network, and those individuals lower the average type in the M network, which eventually stops the inflow. This effect is stronger when efforts are induced optimally since the optimal effort choices allow individuals to take more advantage of the improved parameters in the M network, hence $C_E < C_P < 1$ when $a^M = a^F$ and $B^M = B^F + \varepsilon$, for ε small enough.

From the point of view of the social planner, when $a^M = a^F$ and $B^M > B^F$ she wants a more restrictive M network, because the marginal type she pushes into the F network does not affect the average type in the F network, while it improves the average type in the B network. From the above discussion it is immediate that $C_E < C_P < 1 < C_E^*$ and the result follows. □

$C_E < C_P < 1 < C_E^*$ corresponds to situations in which reaching within network efficiency induces a larger number of individuals joining the mainstream network. Under these circumstances, the regulating government operating only on one margin is "wasting" part of the effort because it generates a counter reaction on the other margin it does not control and induces an even more severe overpopulation of the more efficient network than in the absence of any intervention. We next show that this can lead to a reduction in global efficiency if congestion is costly.

3.3.2 Congestion can reduce global efficiency

We model congestion as follows. The utility of agents in the more crowded network M is multiplied by the following function $f\left(C, b_i^M\right)$

$$f\left(C, b_i^M\right) = \begin{cases} 1 \text{ if } b_i^M < C^* B^F \\ \frac{\left(C^* B^F\right)^2}{b_i^{M2}} + \left(1 - v\left(C\right)\right)\left(1 - \frac{\left(C^* B^F\right)^2}{b_i^{M2}}\right) \text{ if } b_i^M \geq C^* B^F \end{cases} \tag{16}$$

with $f_C'\left(C, b_i^M\right) \geq 0$ which captures that congestion is more harmful the larger the population in M since there are fewer people in M the larger is C. Observe that (16) takes away part of the welfare of b_i^M types above $C^* B^F$ and make it closer to the welfare of type $b_i^M = C^* B^F$ when C is progressively smaller.[20] This way of modeling congestion has the advantage that it does not alter our equilibrium analysis. The reason is that it takes welfare away only from agents that are "supramarginal", i.e., they will choose to go to the M network anyway. This is admittedly artificial, but the point of this proposition is only to highlight a theoretical possibility, and this particular modeling device is the simplest one that delivers the conclusion in a transparent way.

Proposition 5. *Suppose congestion costs are given by $f\left(C, b_i^M\right)$ defined in equation (16), and also that there is no intervention in network choice. An intervention designed to optimize the s_i, k_i choice within network, taking as given the equilibrium network choice might lead to a lower welfare than no intervention. That is, there are parameter values for which the welfare under socially optimal socialization and productive efforts within networks is lower than the welfare with individually optimal choice of both socialization and productive effort.*

To prove Proposition 5 we first derive an expression for welfare.

Lemma 3. *With congestion welfare when the government induces efficient*

[20] Other ways that take surplus away from high b_i^M and are also related to C would also work.

We can also have for symmetry congestion in the other network but that would not change things. So for notational simplicity we apply congestion only in the large network.

productive and socialization efforts within networks is given by

$$\begin{aligned} w_E(C) &= \frac{CB^{F^3}}{8B^M}\left[\left(\frac{1}{1-a^{F2}\overline{b^{F^2}}^2}\right)+\left(\frac{1}{1-a^{M2}\overline{b^{M^2}}^2}\right)C^2\right] \\ &\quad +\left(1-C\frac{B^F}{B^M}\right)\frac{\left(CB^F\right)^2}{2}\left(\frac{1}{1-a^{M2}\overline{b^{M^2}}^2}\right)+\left(1-v\left(C\right)\right)G_E\left(C\right) \end{aligned}$$

where

$$G_E\left(C\right)=\frac{1}{B^FB^M}\left(\int_0^{B^F}\int_{CB^F}^{B^M}\frac{b_i^{M^2}}{2}\left(\frac{1}{1-a^{M2}\overline{b^{M^2}}^2}\right)db_i^Mdb_i^F\right)- \frac{1}{B^FB^M}\left(\frac{\left(B^M-CB^F\right)C^2B^{F^3}}{2}\left(\frac{1}{1-a^{M2}\overline{b^{M^2}}^2}\right)\right) \tag{17}$$

while welfare in the absence of government intervention is given by

$$\begin{aligned} w_P(C) &= \frac{CB^{F^3}}{2B^M}\left[\frac{\left(4+a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4-a^{F^2}\overline{b^{F^2}}^2\right)^2}+\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}C^2\right] \\ &\quad +\left(1-C\frac{B^F}{B^M}\right)\left(CB^F\right)^2\frac{2\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}+\left(1-v\left(C\right)\right)G_P\left(C\right) \end{aligned}$$

where

$$\begin{aligned} G_P\left(C\right) &= \int_0^{B^F}\int_{CB^F}^{B^M}\frac{2b_i^{M^2}}{B^FB^M}\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}db_i^Mdb_i^F- \\ &\quad \left(1-C\frac{B^F}{B^M}\right)\frac{\left(CB^F\right)^2}{2}\frac{4\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2} \end{aligned} \tag{18}$$

Proof. See Appendix B □

To complete the proof of Proposition 5 we only need to find a function $v(C)$ and some parameter values such that no intervention gives a higher welfare than intervention on one margin only inducing optimal productive and socialization efforts within a network. This is done in Lemma 4. It assumes that congestion only has a bite in extremely crowded networks.

Lemma 4. *Let $v(C) = 1$ for C close to zero, and for C bounded away from zero $v(C) = 0$. Then if $1 - a^{M^2} \frac{B^{M^4}}{9} \approx 0$ (sufficiently small) and B^M big enough, $w(C_E) < w(C_P)$*

Proof. See Appendix B □

Proposition 5 points to the possibility that intervention at one margin might be worse than no intervention. Returning to our immigrant example, if the immigrant network has clearly a productivity disadvantage and there are some congestion costs, the absence of any government intervention can be socially better than the existence of local intervention via transfers and taxes that induce the efficient effort levels within the networks.[21]

4 Parental influence on the child's private return to the network

The theory developed so far can be extended to study the evolution of different occupational or cultural groups through intergenerational transmission of group-specific abilities. If parents can induce their children to choose a particular network, which one will they choose? The answer is not obvious, even for purely altruistic parents who have no direct preferences to induce their children to follow in their footsteps. These altruistic parents internalize the welfare of their children in different networks and would like to induce their child to join the most beneficial network. If the non-parental network is more

[21] A parallel result is found in education literature in models in which overall student effort is influenced both by parental effort and the school environment. In this context Albornoz, Berlinski, and Cabrales (2014) have shown that a reduction in class size leads to lower parental effort and hence little (or no) improvement in overall educational performance.

productive than the parental network parents may prefer their children to embrace a network different from their own. However, whether or not they can do so also depends on how easy it is to enhance their child's ability in the different networks. In which network-specific abilities parents invest for their children is an interesting question because if all parents induce their children to join the same group, say the most productive one, then social cultural tendencies may result in further overpopulation of the mainstream network and exacerbate its potential cost of congestion. In this section, we show that in many plausible scenarios, intergenerational transmission of abilities leads to a tendency towards conformism, which pushes towards overpopulation of the mainstream network M, with the potentially negative consequences on welfare we have discussed in section 3.3. In section 4.1, we introduce parental influence and prove existence and uniqueness of the new sub-game. In section 4.2, we study the different tendencies in society that emerge in our framework. Finally, in section 4.3, we discuss some evidence on occupational intergenerational mobility that is consistent with our theoretical results.

4.1 The parental influence subgame

Parents invest in improving/expanding the abilities of their children in one or both of the networks.[22] More precisely, they can change the support of the distribution of productivities b^ls from which b_i^l, the child's productivity in network l, is drawn. As children choose networks based on their own productivity in each network, one can view parental influence as a process through which relative productivities are affected in the direction parents decide. A realistic consequence of our modeling choice is that parents cannot fully determine which network their children eventually select.

We study the case in which network-specific productivity (b^l) are distributed according to a uniform distribution described by:

$$b_i^l \sim U\left[0, e_{p_i}^l\right],$$

[22] With this assumption we depart from direct socialization efforts aiming at affecting children preferences, a possibility explored in Bisin and Verdier (2001).

where

$$e_{p_i}^l = \begin{cases} \overline{A^k} + x_{p_i}^l, \text{ if the parental network } k = l \\ \underline{A}^k + x_{p_i}^l \text{ if the parental network } k \neq l \end{cases} \tag{19}$$

and $x_{p_i}^l$ summarizes the productivity-enhancing parental effort towards child i in network l. Positive parental effort in any network induces a first order stochastically dominating shift in the distribution of abilities for that network.

Parameters $(\underline{A}^l, \overline{A^l})$ capture the initial endowment of ability in the non-parental and the parental networks, respectively. Consider a parent belonging to network F. In this case, $\overline{A^F}$ summarizes how the parental network F (e.g. a neighborhood, a religious community, an occupation or any particular cultural identity) influences the distribution of abilities from which his child i will draw his ability in the parental network b_i^F. However, the other network, in this case M, may also favor the acquisition of its specific abilities through, for example, mainstream education or mass media. This is captured by $\underline{A}^F$ which is the upper bound for the ability distribution of an F-parent's child in network M. Both parameters $(\underline{A}^l, \overline{A^l})$ may differ in each network and their different combinations can describe a rich variety of cultural or organizational contexts in society. For example, a society characterized by a dominant culture/group may be described as a case where $\underline{A}^F > \overline{A^F}$ and $\underline{A}^M < \overline{A^M}$. Similarly, a society where the influence of the parental network is specifically strong would have $\underline{A}^F < \overline{A^F}$ and $\underline{A}^M < \overline{A^M}$.

For parents, the decision to influence their children's network selection consists in choosing $x_{p_i}^F$ and $x_{p_i}^M$ to maximize:

$$E\left[\max\left[2b_i^{F^2}\frac{\left(4 + a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4 - a^{F^2}\overline{b^{F^2}}^2\right)^2}, 2b_i^{M^2}\frac{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2}\right]\right] \tag{20}$$

subject to the constraint that total effort cannot exceed an exogenous time endowment (K):[23] such as

$$x_{p_i}^F + x_{p_i}^M = K. \tag{21}$$

Our parents are perfectly altruistic and do not want to impose any particu-

[23]In principle K could vary across networks and parents. However, endogenizing K would complicate the analysis with no additional insights.

lar network *per se*. They only care about their children's welfare, irrespectively of whether they join their same own network or not.[24] As discussed below, altruistic parental involvement may generate intergenerational network mobility within the same family.

We proceed by finding the equilibrium levels of $x_{p_i}^F$ and $x_{p_i}^M$. As the time allocated to enhance the abilities in the mainstream network ($x_{p_i}^M$) may be expressed as $K - x_{p_i}^F$, the exercise boils down to finding the optimal $x_{p_i}^F$. Notice that $x_{p_i}^F$ and $x_{p_i}^M$ are individual decisions and therefore parents take $\overline{b^{F^2}}^2$ and $\overline{b^{M^2}}^2$ as given. This allows us to ease the analysis by defining the parameters F and M that do not depend on the individual's characteristics as:

$$F = \frac{\left(4 + a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4 - a^{F^2}\overline{b^{F^2}}^2\right)^2}, M = \frac{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2} \tag{22}$$

Thus, parents choose $x_{p_i}^F$ and $x_{p_i}^M$ to maximize

$$E\left[\max\left[2b_i^{F^2}F, 2b_i^{M^2}M\right]\right] \tag{23}$$

subject to

$$x_{p_i}^F + x_{p_i}^M = K \text{ and } 0 \leq x_{p_i}^F \leq K \tag{24}$$

In Appendix C we show that equation (23) for F-parents can be expressed as:[25]

[24]This is another difference with the literature on cultural transmission, which typically assumes that parents are imperfectly altruistic and encourage their children to adopt their cultural traits (their own network). See Bisin and Verdier (2010) for a review.

[25]For M-parents the equation becomes

$$g\left(x_{p_i}^F\right) = \begin{cases} 2F\frac{\left(\underline{A}+x_{p_i}^F\right)^3}{6\left(\overline{A}+K-x_{p_i}^F\right)}\sqrt{\frac{F}{M}} + 2M\frac{\left(\overline{A}+K-x_{p_i}^F\right)^2}{3} & \text{if } x_{p_i}^F < \frac{K+\left(\overline{A}-\underline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}} \\ 2M\frac{\left(\overline{A}+K-x_{p_i}^F\right)^3}{6\left(\underline{A}+x_{p_i}^F\right)}\sqrt{\frac{M}{F}} + 2F\frac{\left(\underline{A}+x_{p_i}^F\right)^2}{3} & \text{if } x_{p_i}^F > \frac{K+\left(\overline{A}-\underline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}} \end{cases} \tag{25}$$

$$
g\left(x_{p_i}^F\right)=\begin{cases} 2F\frac{\left(\overline{A}+x_{p_i}^F\right)^3}{6\left(\underline{A}+K-x_{p_i}^F\right)}\sqrt{\frac{F}{M}}+2M\frac{\left(\underline{A}+K-x_{p_i}^F\right)^2}{3} & \text{if } x_{p_i}^F<\frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}} \\ 2M\frac{\left(\underline{A}+K-x_{p_i}^F\right)^3}{6\left(\overline{A}+x_{p_i}^F\right)}\sqrt{\frac{M}{F}}+2F\frac{\left(\overline{A}+x_{p_i}^F\right)^2}{3} & \text{if } x_{p_i}^F>\frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}} \end{cases} \tag{26}
$$

In the Appendix C.1 we prove that $g\left(x_{p_i}^F\right)$ is convex in both branches (Lemma 14). This implies that obtaining the maximum simply requires to compare the value of this function at the extreme points of its two branches. However, we need to take into account that parental investment cannot be negative, which implies that for some parameter values only one of the two branches exists. In these cases we only need to compare $g(0)$ with $g(K)$ in the only existing branch to find the optimal solution and parents will invest their entire time K in one network only. In the appendix, we show that both branches of $g\left(x_{p_i}^l\right)$ exist if the switching value from one branch to the other is neither negative nor bigger than the time endowment K which leads to the conditions

$$
\frac{K+\underline{A}}{\overline{A}}>\sqrt{\frac{F}{M}}>\frac{\underline{A}}{K+\overline{A}} \text{ for } F \text{ parents} \tag{27}
$$

and

$$
\frac{K+\underline{A}}{\overline{A}}>\sqrt{\frac{M}{F}}>\frac{\underline{A}}{K+\overline{A}} \text{ for } M \text{ parents} \tag{28}
$$

Recall that F and M capture the common utility parameters for all individuals in network F and M respectively and individual utility is multiplicative in these terms. Hence, F and M capture the network's productivity independently of individual productivity. The conditions that both branches of the parental optimization problem exist put restrictions on the relative network productivities with respect to the relative upper bounds from which the child's productive abilities in each network are drawn. $K+\underline{A}$ describes the maximum possible upper bound for a child's ability in the non-parental network which is achieved when parents invest all their resources in the non-parental network. Similarly, $K+\overline{A}$ captures this maximum possible upper bound for a child's ability in the parental network. Conditions (27) and (28) say that in order for both branches of the parental objective function to exist, the (square-root of the) relative common network returns (parental network/ non-parental network) have to be bounded from above by the relative upper bounds of the talent distribution if parents invested only in the non-parental network and from below by the relative upper bound of the talent distribution if parents invested only in their parental network. In other words network productivity (F and M) has to be fairly similar in both networks. It can be more dissimilar, the higher the available investment resources K.

When these conditions are satisfied, there exists the possibility of an interior solution in which parents invest in both networks. However, the discussion in Appendix C.2 states that this is not an equilibrium result.[26] This implies that parents will fully invest K in one network only. However, for some parameter values it is optimal for parents to always invest in the parental network and for other values in the non-parental network. Proposition 6 states these conditions. The conditions are derived by comparing the utility from investing in the parental network to the utility of investing in the non-parental network and we need to distinguish the cases where both branches of the objective function exists (network productivity across networks is fairly similar) and where only one branch exists.

Proposition 6. *The decisions on cultural parental involvement are as follows:*

1. *F-parents will fully invest in their own network F if condition (27) is not satisfied and*

$$\frac{\underline{A}^3}{(\overline{A}+K)} + 2\frac{F}{M}\sqrt{\frac{F}{M}}\left(\overline{A}+K\right)^2 > \frac{(\underline{A}+K)^3}{\overline{A}} + 2\frac{F}{M}\sqrt{\frac{F}{M}}\overline{A}^2 \tag{29}$$

   or if condition(27) is satisfied and

$$\frac{\underline{A}^3}{(\overline{A}+K)} + 2\frac{F}{M}\sqrt{\frac{F}{M}}\left(\overline{A}+K\right)^2 > \left(\frac{F}{M}\right)^2\frac{\overline{A}^3}{(\underline{A}+K)} + 2\sqrt{\frac{F}{M}}\left(\underline{A}+K\right)^2 \tag{30}$$

[26] The lack of interior equilibria is probably not an essential result of our framework and seems to be driven by the specific functional forms we choose for parents to enhance their children's network abilities and by the absence of externalities across networks.

while they will fully invest in the other network M in the remaining cases.

2. *M-parents will fully invest in their own network M if condition (28) is not satisfied and*

$$\frac{\underline{A}^3}{\left(\overline{A}+K\right)} + 2\frac{M}{F}\sqrt{\frac{M}{F}}\left(\overline{A}+K\right)^2 > \frac{\left(\underline{A}+K\right)^3}{\overline{A}} + 2\frac{M}{F}\sqrt{\frac{M}{F}}\overline{A}^2 \tag{31}$$

or if condition(28) is satisfied and

$$\frac{\underline{A}^3}{\left(\overline{A}+K\right)} + 2\frac{M}{F}\sqrt{\frac{M}{F}}\left(\overline{A}+K\right)^2 > \left(\frac{M}{F}\right)^2 \frac{\overline{A}^3}{\left(\underline{A}+K\right)} + 2\sqrt{\frac{M}{F}}\left(\underline{A}+K\right)^2 \tag{32}$$

while they will fully invest in the other network F in the remaining cases.

Proof. Condition (30) says that investing fully in F is better than investing fully in M for F-parents, while condition (32) says that investing fully in M is better than investing fully in F for M-parents when both branches of the parental objective function as described by (75) exist. When only one branch exists, it is the first branch for M-parents and the second branch for F-parents. Comparing the corners in each branch gives inequalities (29) and (31). □

The conditions for the different type of parents are in reality identical: they only differ in the way relative network productivity enters in the expression. Relative productivity always enters as productivity of parental network/ productivity of non-parental network. While these conditions look complicated, they behave as one expects intuitively. When network productivity differs significantly across networks the condition to invest in their parental network ((29) for F parents and (31) for M parents) is more easily satisfied the bigger the relative parental/non-parental network productivity, the bigger the initial parental network productivity upper bound $\overline{A}$ and the lower the initial non-parental network productivity bound $\underline{A}$.[27] The effect of expanding the

[27] To see this notice condition (29) for F-parents can be rewritten as

$$2\frac{F}{M}\sqrt{\frac{F}{M}}\left(\left(\overline{A}+K\right)^2 - \overline{A}^2\right) > \left(\frac{\left(\underline{A}+K\right)^3}{\overline{A}} - \frac{\underline{A}^3}{\left(\overline{A}+K\right)}\right) \tag{33}$$

investment resources K is likely to tighten the condition unless $\overline{A} \lll \underline{A}$, i.e. the non-parental network has a huge initial productivity advantage, in which case the condition is unlikely to be satisfied in the first place.[28]

Higher K makes it more likely that both branches of the parental objective function exist in which case parents will invest in their own network if conditions (30) for F-parents and (32) for M-parents hold. Analytically, it cannot be shown how these conditions change with the underlying parameters $\underline{A}$, $\overline{A}$ and K without making stark assumptions on how these parameters differ. Differences in these observable parameters will give rise to different cases of cultural transmission as we will show in subsection 4.2 where we will give an intuitive explanation for the different outcomes. Before doing so we need to prove the existence of an equilibrium with endogenous parental influence formally.

Finally, we prove existence. This requires distinguishing between different cases depending on which corner is chosen: (i) everybody invests in the same network or (ii) parents invest in their own network only.[29] In both cases, we have to show that the defining equation of C_P given by (14) has a fixed point. The only difficulty consists in calculating $\overline{b^{M^2}}$ and $\overline{b^{F^2}}$, since we need to take into account that children coming from different networks may face different uniform distributions. In Appendix C.3, we calculate $\overline{b^{M^2}}$ and $\overline{b^{F^2}}$ taking into account the proportions of F/M children coming from F/M parents, derive an expression of C_P and obtain $G(C)$. Then, we show that $G(C)$ has a fixed

The LHS is clearly increasing in $\frac{F}{M}$, hence the condition is easier to satisfy. LHS increases in $\overline{A}$ while the RHS decreases, hence the condition easier to satisfy for high $\overline{A}^2$. The RHS is increasing in $\underline{A}$, so the condition is harder to satisfy.

[28] To see this for F-parents rewrite (33) as

$$2\frac{F}{M}\sqrt{\frac{F}{M}} > \frac{\left(\frac{(\underline{A}+K)^3}{\overline{A}} - \frac{\underline{A}^3}{(\overline{A}+K)}\right)}{\left(\left(\overline{A}+K\right)^2 - \overline{A}^2\right)}$$

Easy calculus show that how the RHS changes with K depends on the sign of

$$2K^3 + 3K^2\underline{A} + 7K^2\overline{A} + 12K\underline{A}\overline{A} + 4K\overline{A}^2 + 3\underline{A}^2\overline{A} + 6\underline{A}\overline{A}^2 - \underline{A}^3$$

which tends to be positive unless $\overline{A} \lll \underline{A}$.

[29] It is not possible to have both types of parents investing in the opposite network.

point, which proves:

Proposition 7. *An equilibrium with endogenous parental influence exists and is unique.*

4.2 Tendencies in intergenerational cultural transmission

Parents influence the way in which the distribution of society into different groups evolves. According to Proposition 6, different combinations of the overall productivity ratio, the time endowment, and the relative initial ability determine in which network the parents invest effort to enhance their children's abilities.

Symmetric network influence

Consider first a case in which the parental and non-parental networks are symmetric in terms of how they influence the distribution of network-specific abilities ($\overline{A^l} = \underline{A^l}\ \forall l$). As the only difference between networks is overall network efficiency, parents invest in the most *a priori* profitable network irrespectively of their time endowment. This is stated in the following proposition:

Proposition 8. *If* $\overline{A^l} = \underline{A^l}$ *then*

$$\begin{aligned} x^F_{p_i} &= 0 \textit{ and } x^M_{p_i} = K \quad \textit{if } M > F \\ x^F_{p_i} &= K \textit{ and } x^M_{p_i} = 0 \quad \textit{if } M < F \end{aligned}$$

Proof. See Appendix C.4. □

Thus, in absence of asymmetries in network influence, there exists a general tendency toward conformism where parents spend all their influence effort in generating abilities for the most profitable network. Notice that the reasons for anyone joining the "alternative" network are purely idiosyncratic. This tendency towards conformism is important in the light of our previous result about overpopulation in the mainstream network, as it suggests that the intergenerational transmission of ability may exacerbate potential problems of congestion. Importantly, conformism is not confined to symmetric network influence.

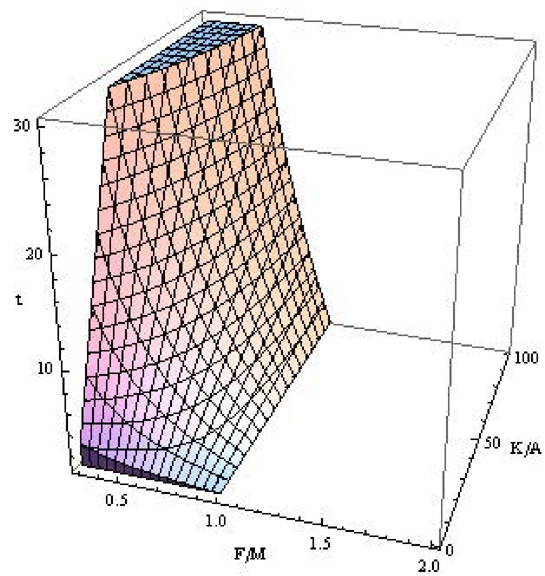
Asymmetric network influence

We allow now for differences in network influence. Given the large number of parameters we make our point through numerical simulations. For simplicity, we reduce the spectrum of parameters by defining $y = F/M \in (0, \infty)$ as the ratio of exogenous efficiency of the networks, $t^l = \overline{A^l}/\underline{A^l} \in (0, \infty)$ as the ratio of the influence of parents' network to "the others"' network; and $z = K/\underline{A^l} \in (0, \infty)$ as a parameter that captures time endowment restrictions.

From the perspective of the child, there are two possibilities. Either their parents' network dominates or he receives more influence from the other network. To capture the case of parents' network dominance, it suffices to assume $\overline{A}^l/\underline{A}^l > 1 \forall l$. To explore the resulting social inculturation pattern, notice that unconstrained parents want their children to join the most productive network. For this reason, K becomes a key parameter in the presence of network influence asymmetry.

If the time endowment (K) is large enough as to revert the neighborhood's influence (t), then parents in the disadvantaged network will invest in improving their children's skills to succeed in the most productive network. This is illustrated in Figure 1 where we depict F-parents' decision of investing in M network. While the x-axis represents the variable F/M, the y-axis represents $K/\underline{A}$ and the z-axis represents t. Notice that the M-network being preferred by F-parents requires that the overall productivity of the network F must be highly inferior to the M network (i.e. a relatively low value of F/M). This is, as the M-network overall productivity relative to the $F-$network increases, it is more likely to find F-parents investing in M's network education. This condition is relaxed for higher $K/\underline{A}$ and lower values of t. Naturally, whenever F-parents have incentives to induce their child to choose the M-network, M-parents will also do so. Thus, our model generates societies where parents voluntarily promote intergenerational differences within families and

Figure 1: Cultural conformism under asymmetric network influence: F-parents investing in the M-network



cultural conformism emerges in equilibrium even under asymmetric network influence.[30]

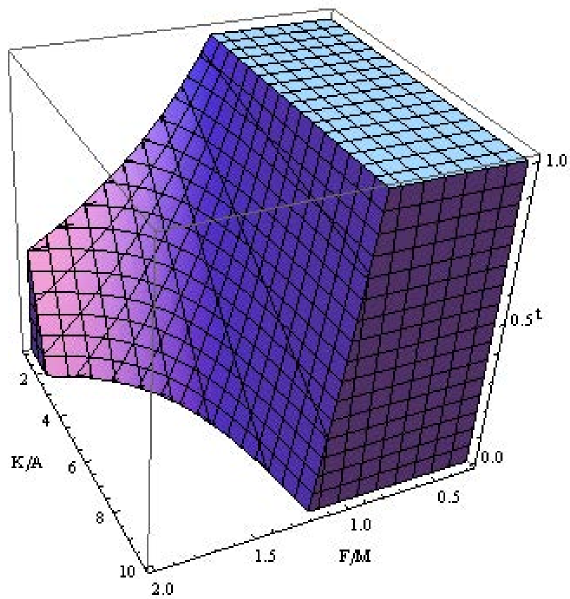
Conformism is not always the outcome. For low levels of the time endowment (K), parents cannot revert the initial influence of the child's environment (t). In this case, the society is characterized by intergenerational transmission of within family abilities and low mobility between networks, resulting in fragmentation into two well defined networks. In these societies, intergenerational differences within families are purely associated with idiosyncratic characteristics of children. For example, switching networks involve children with special abilities in specific activities required in the non-parental network.

We now consider the case of an influential established network (not necessary advantageous in terms of profits). This corresponds to societies where the institutions are designed to promote a dominant, say M, network.[31] This

[30]Note that if $\frac{F}{M} < 1$, and K is high enough, we have the same result but with all parents inducing their children to choose $F - Network$.

[31]In many modern societies schools enhance the skills and abilities required by mainstream activities (the M-network in our model).

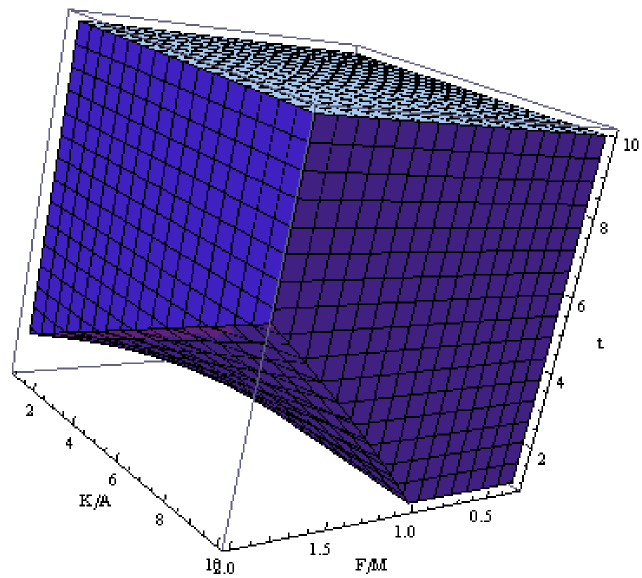
Figure 2: Cultural conformism under an hegemonic network: F-parents investing in the M-network



is captured by assuming $\underline{A}^F > \overline{A}^F$ and $\underline{A}^M < \overline{A}^M$. If network M is not only dominant but also the most efficient one, parents from both networks trivially will spend all their education effort to improve their child's productivity in network M. The case where the M-network is not the most productive one, (i.e.: $F > M$) is more interesting: parents in both networks may be trapped into investing in the preponderant, yet least productive, network M if the time endowment (K) is not large enough. While the former case is similar (but stronger) to the cases associated with conformism, the latter case corresponds to societies where rational parents tend to invest in the less profitable but dominant network in society. Figures 2 and 3 display the graphic examination of these two cases:

Figure 2 illustrates the case in which parents from the F-network invest in capacities to be part of the M-network for $0 < t^F < 1$, while Figure 3 displays the case in which parents from the M-network invest their education effort in their own network abilities for $t^M > 1$. Naturally, both parents invest

Figure 3: Cultural trap under with an hegemonic network: M-parents investing in the M-network



in network M-abilities for $F/M < 1$. Surprisingly this might happen even if $F/M \geq 1$. As Figure 2 shows F-parents are willing to invest in the inferior network M when $K/\underline{A}$ is sufficiently low or the dominant network's influence is relatively too strong (i.e.: $t^F \to 0$). Similar results hold for M-parents (Figure 3) but in this case, the strong influence of the predominant network is captured by $t^M \to \infty$. Notice that if the M-network is deeply rooted in the society's culture (this is, if $t^M. \to \infty, t^F. \to 0$), our model can explain societies where parents might choose not to invest in educating their children in the most profitable network. Investment in low productivity networks is imposed by cultural factors, trapping the society in a low productivity cultural dynamics.

4.3 Occupational mobility

Our analysis shows that parents matter to explain the distribution of individuals across networks. But crucially, they matter more in some networks than in others. Key observable variables are the relative productivity and the ability advantage in the parental network. The occupational mobility literature provides an interesting ground to study the implications of our analysis. In this setup, where the parental network is clearly dominant, our model yields the following predictions: there is high persistence of occupational categories within families; the probability for an individual to fall within the same occupational category as her/his parent is increasing in the size of this occupational category (more network externalities); intergenerational occupational persistence is higher in more profitable occupations (network); and switchers (children joining a different occupation than their parents') tend to move into more profitable jobs. We discuss now how these predictions are consistent with available evidence on occupational mobility across four different countries.

We base our analysis on the evidence provided by the following sources of information about occupational mobility in different countries: Long and Ferrie (2013), from which we draw evidence on Britain and the U.S.; Azam (2013) for the case of India; and Binzel and Carvalho (2013) for Egypt. We restrict ourselves to these papers because they either calculate or at least provide data to calculate the unconditional probability to belong to a certain occupational category (w_i), and the conditional probability to belong to this category provided that it is the father's category (H_i). The occupational categories studied across these countries are very similar: Long and Ferrie (2013) classify professions into two categories of white collar workers (high white collar, HWC, and low white collar, LWC), farmers, skilled/semiskilled and unskilled. Azam (2013) uses the same categorization but with a single white collar category. Binzel and Carvalho (2013) look at farmers, unskilled/semi manual, skilled manual, white collar and professionals. However, they do not necessarily reflect the profitability of each occupation. When the white collar category is not split into high white collar and low white collar the profitability is no longer clear. Also some countries group skilled and semiskilled together while others group semi skilled and unskilled together.[32]

Table 1 summarizes the unconditional w_i and conditional probability H_i of belonging to the different occupational categories. Following Currarini,

[32] Also the profitability of farming is clearly higher in the U.S and U.K than in Egypt which most likely is more productive in farming than India.

Jackson, and Pin (2009) we also calculate and report a measure of inbreeding homophily:

$$IH_i = \frac{H_i - w_i}{1 - w_i},$$

which measures the extent of the bias with respect to baseline since IH_i relates to the maximum possible bias (the term $1 - w_i$). The results for Britain and the U.S. are calculations based on data from the online appendix of Long and Ferrie (2013) for intergenerational occupational mobility in Britain and the U.S. 1949-55 to 1972-73. The data uses males age 31-37 in 1972 from the Oxford Mobility Study and white, native-born males age 33-39 in 1973 from the Occupational Change survey. The occupation of the father is the one he had when the respondent was age 14 in Britain and age 16 in the U.S. The total number of respondents (son-father pairs) were 1123 for Britain and 2988 for the U.S. The data on India is taken from table 1 in Azam (2013) and based on the Indian Human Development Survey 2005. We arbitrarily took the 1965-1974 birth cohort which is based on 11557 father-son pairs.[33] The data on Egypt stems from Binzel and Carvalho (2013) web appendix based on the 2006 cross-section of the Egypt Labor Market Panel Survey and we report data on men born in 1968-1977.

The conditional probability to work in a certain occupational category is always higher than the unconditional probability, i.e. $H_i > w_i$. In other words, there is a high persistence of occupational categories within the family across all countries studied as expected in a setup where the parental network is dominant. This intergenerational persistence is more pronounced in the occupational categories with more people, a fact which points to the presence of higher network externalities. Figure 4 clearly shows that H_i increases in w_i and that this is true even when mixing different countries. The difference between w_i and H_i varies across occupational categories. The IH_i index reveals that inbreeding homophily is highest in the most profitable occupations (high white collar in Britain and the U.S., skilled/semi skilled in India and professional in Egypt) and the bias is fairly low in the less profitable categories.

[33] Azam (2013) does not find any differences in mobility in successive ten year birth cohorts.

Table 1: Occupational mobility across job categories and countries

Country	Category	w_i	H_i	IH_i
Britain	high white collar (HWC)	0,259	0,628	0,498
	skilled/semi skilled	0,542	0,622	0,175
	farmer	0.013	0,209	0,199
	low white collar (LWC)	0,123	0,209	0,098
	unskilled	0,062	0,09	0,03
U.S.	high white collar (HWC)	0,372	0,617	0,39
	skilled/semi skilled	0,398	0,466	0,112
	farmer	0,025	0,135	0,113
	low white collar (LWC)	0,111	0,161	0,07
	unskilled	0,093	0,133	0,043
India	skilled/semi skilled	0,400	0,716	0,527
	white collar (WC)	0,120	0,444	0,369
	unskilled	0,337	0,552	0,325
	farmer	0,144	0,265	0,141
Egypt	professional	0,276	0,556	0,386
	skilled manual	0,235	0,43	0,255
	white collar (WC)	0,16	0,254	0,111
	semi skilled/unskilled manual	0,155	0,294	0,1639
	farmer	0,174	0,403	0,276

Figure 4: Intergenerational occupational mobility: relationship between H_i and w_i

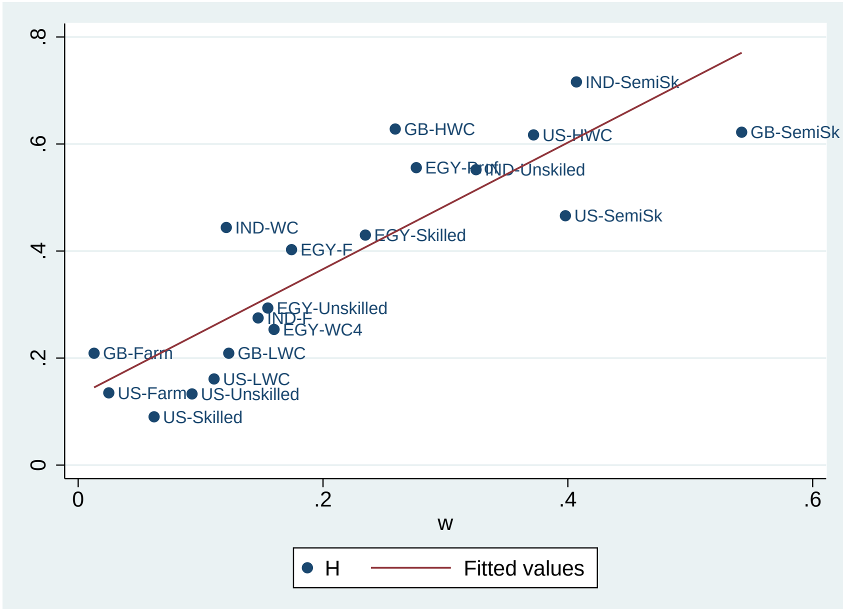


Table 2 displays the probabilities for a son to end up in a occupational category j that is different from that of her/his father i. Each panel represents a different country (Britain, US, India, Egypt). The general message of these transition matrices is that switching is more likely towards the more profitable occupation. In Britain and the US, switchers are more likely to be HWC (high weight collar) or skilled/semiskilled than to fall into the lower skill categories (farmer, LWC, unskilled), which is consistent with parents investing in their children's skills in the more advantageous occupations. The same is true in the case of Egypt, where the most advantageous occupational categories are professional and skilled manual. India exhibits a general tendency towards switching to skilled jobs too, except for the case of farmers' children who seem to have a harder time to leave low skill occupations.

The literature on the intergenerational transmission of employers (see e.g. Stinson and Wignall (2014)) is also consistent with the above predictions of our model. Employer sharing probabilities between father and sons in the U.S. in 2010 are much higher than the baseline probability that a father shares a

Table 2: Occupational intergenerational transitions

Britain		son				
		HWC	skilled/semi	farmer	LWC	unskilled
Father	HWC	0,63	0,16	0,01	0,16	0,05
	skilled/semi	0,20	0,62	0,004	0,1	0,07
	farmer	0,19	0,44	0,21	0,07	0,09
	LWC	0,38	0,39	0,07	0,21	0,015
	unskilled	0,13	0,66	0,006	0,12	0,09
US		son				
		HWC	skilled/semi	farmer	LWC	unskilled
Father	HWC	0,62	0,21	0,004	0,12	0,05
	skilled/semi	0,33	0,47	0,006	0,11	0,09
	farmer	0,24	0,43	0,14	0,08	0,12
	LWC	0,50	0,26	0,004	0,16	0,08
	unskilled	0,25	0,51	0,01	0,1	0,13
India		son				
		skilled/semi	WC	unskilled	farmer	
Father	skilled/semi	0,72	0,15	0,11	0,02	
	WC	0,4	0,44	0,1	0,06	
	unskilled	0,33	0,07	0,55	0,05	
	farmer	0,3	0,09	0,34	0,27	
Egypt		son				
		professional	skilled	WC	semi & / unskilled	farmer
Father	professional	0,56	0,15	0,18	0,09	0,03
	skilled manual	0,20	0,43	0,16	0,18	0,03
	WC	0,27	0,24	0,25	0,16	0,08
	semi/unskilled	0,26	0,21	0,17	0,29	0,07
	farmer	0,15	0,22	0,1	0,13	0,4

job with an unrelated male similar to her/his son.[34] The employer sharing probability is found to depend on parental earning: it is less than average for the sons whose fathers are in the lowest earning decile, and higher than average for the sons of the highest earning fathers. Moreover, average log earnings are found to be significantly higher at shared than at unshared jobs except in the case of the sons of fathers whose earnings are in the first and second lowest decile.

5 Additional relevant literature

Our findings on parental investment are relevant to the economic analysis of cultural transmission. An influential strand of this literature uses models where agents have what Bisin and Verdier (2001) called *imperfect empathy* (see also Bisin and Verdier (2010), Hauk and Saez-Marti (2002)). Under *imperfect empathy*, altruistic parents evaluate their children's choices in light of their own preferences and invest in transmitting their own cultural traits. In our model, parents would like their children to join the most profitable network irrespectively to which network the parents themselves belong. As a consequence, social minorities, if unconstrained, would tend to encourage assimilation into the mainstream culture as long as it is more productive. However, social conformism weakens when parents lack enough inculturation resources (time or material goods) or if they are exposed to a strong cultural pressure from their own network.[35] In this sense, our model also identifies forces for cultural segregation and integration, although we emphasize a complementary mechanism.

Selecting networks is similar to the process of choosing friends. In this sense, our paper is related to a growing literature on friendship formation. Currarini, Jackson, and Pin (2009) study individual preferences in friend choice.

[34]The overall job sharing probabilities of father and sons found by Stinson and Wignall (2014) for the U.S., are remarkably similar to the ones found for Canada and Denmark (see Kramarz and Skans (2007) using Swedish data, and Corak and Piraino (2011) based on Canadian data, and Corak (2013), using both Canadian and Danish data).

[35]Bisin, Patacchini, Verdier, and Zenou (2010) provides evidence showing that cultural identity investments of minority groups increase under stronger group pressure.

Their model can explain various empirical patterns of homophily, some of which are consistent with our theoretical observations. For example, they document that larger groups tend to exhibit more homophily, and they also find that individuals in larger groups tend to socialize more (they have more ties). Our model thus generates similar patterns of socialization but in a complementary framework in which we emphasize productive processes. There are other points of connection between our paper and this literature. Currarini and Vega-Redondo (2011) present a model in which individuals draw from either a homophilous network of same-type agents, or from the whole network. The main result is that inbreeding is more likely to happen in large groups because they are the ones for which the extra (fixed) cost of searching in the whole network does not warrant the extra benefit of a wider search. We could easily extend our model to allow agents to form connections in the two networks, and although network size does not matter in our context, we may also find that individuals from the less productive (*a priori*) network would be more willing to pay the fixed cost to enjoy the benefits of a wider interaction.

The literature on academic connections is also relevant to this study. For example, Ductor, Fafchamps, Goyal, and van der Leij (2013) empirically evaluate the predictive power of several network characteristics on individual research outputs in economic research. The productivity of coauthors, closeness centrality, and the number of past coauthors are particularly relevant to infer young researchers future productivity. Given that the network selection in our model can be naturally assumed to take place at an early stage of an individuals' life or career, our model can generate this observation. Moreover, our framework suggests that an individual's network selection could be used to infer her/his unobservable productivity and, once chosen, the productive effect of a network is amplified by endogenous socialization.

6 Conclusion

We have studied a model that integrates productive and socialization efforts with network choice and parental investments. The relative simplicity of our framework allows us to characterize the unique symmetric equilibrium of this game. As expected in a model with complementarities, individuals underinvest in productive and social effort. However, solving only the investment problem can exacerbate misallocations of individuals across groups due to network choice, to the point that in the presence of congestion costs it may generate an even lower social welfare than no intervention at all. This is an important novel conclusion of this paper. Individuals do not only choose their efforts within a network but also the network to which they belong, and this has implications for policy design. We also examine the interaction of parental investment with network choice and obtain two main results: there is preponderant tendency towards conformism and intergenerational network persistence should be higher in more profitable networks. We relate this last equilibrium result to empirical findings on the intergenerational occupational mobility.

One possible avenue for further research would be to explore the dynamic implications of our model. The agents' choices in our framework are static, but the work on homophily shows that some fruitful insights can be obtained from dynamic models of group formation. For example, Bramoullé, Currarini, Jackson, Pin, and Rogers (2012) show that it is only for young individuals that homophily-based contact search biases the type distribution of contacts.[36] Hence long-term networks need not be type-biased. We could extend our model to allow for participation in diverse networks over time and thus ascertain if biases in productive network choice persist over time. Clearly, another important extension would be to allow some spillovers between networks and partial participation of agents in several of them.

As noted in the introduction, this study is primarily focused on the productive reasons for choosing networks, a line of research that is complementary to work on cultural transmission (Bisin and Verdier (2001)). That said, Reich (2012) has shown that it is possible to fruitfully integrate cultural and productive considerations into a network model. This may prove to be another interesting line for extending our model.

[36] Another example of the interaction of homophily and dynamics is Golub and Jackson (2012), which shows that homophily induces a lower speed of social learning (the opinions of others like me are likely to be similar to my own).

References

Albornoz, F., S. Berlinski, and A. Cabrales (2014): "Motivation, resources and the organization of the school system," mimeo.

Albornoz, F., A. Cabrales, and E. Hauk (2014): "Which club should I attend, Dad?: Targeted socialization and production," Discussion paper, CEPR DP10321.

Azam, M. (2013): "Intergenerational occupational mobility in India," Discussion paper, IZA Discussion Paper.

Benabou, R. (1993): "Workings of a City: Location, Education, and Production," *The Quarterly Journal of Economics*, 108(3), 619–652.

Binzel, C., and J. P. Carvalho (2013): "Education, social mobility and religious movements: A theory of the Islamic revival in Egypt," Discussion paper, IZA Discussion Paper.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2010): "Bend it like Beckham: Ethnic identity and integration," Discussion paper, National Bureau of Economic Research.

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

——— (2010): "The Economics of Cultural Transmission and Socialization," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. Jackson. Elsevier.

Bramoullé, Y., S. Currarini, M. O. Jackson, P. Pin, and B. W. Rogers (2012): "Homophily and long-run integration in social networks," *Journal of Economic Theory*, 147(5), 1754–1786.

Brumfiel, G. (2013): "UK research councils could face mergers," *Nature*, 493, 586.

Cabrales, A., A. Calvó-Armengol, and Y. Zenou (2011): “Social interactions and spillovers,” *Games and Economic Behavior*, 72(2), 339–360.

Cassiman, B., and R. Veugelers (2002): “R&D cooperation and spillovers: some empirical evidence from Belgium,” *American Economic Review*, 92(4), 1169–1184.

Corak, M. (2013): “Income inequality, equality of opportunity, and intergenerational mobility,” *The Journal of Economic Perspectives*, pp. 79–102.

Corak, M., and P. Piraino (2011): “The intergenerational transmission of employers,” *Journal of Labor Economics*, 29(1), 37–68.

Currarini, S., M. O. Jackson, and P. Pin (2009): “An economic model of friendship: Homophily, minorities, and segregation,” *Econometrica*, 77(4), 1003–1045.

Currarini, S., and F. Vega-Redondo (2011): “A simple model of homophily in social networks,” Discussion paper, mimeo.

Ductor, L., M. Fafchamps, S. Goyal, and M. J. van der Leij (2013): “Social networks and research output,” *forthcoming Review of Economics and Statistics.*

Glaeser, E. L., B. I. Sacerdote, and J. A. Scheinkman (2003): “The social multiplier,” *Journal of the European Economic Association*, 1(2-3), 345–353.

Golub, B., and M. O. Jackson (2012): “How homophily affects the speed of learning and best-response dynamics,” *The Quarterly Journal of Economics*, 127(3), 1287–1338.

Goyal, S. (2012): *Connections: an introduction to the economics of networks.* Princeton University Press.

Hauk, E., and M. Saez-Marti (2002): “On the cultural transmission of corruption,” *Journal of Economic Theory*, 107(2), 311–335.

Jackson, M. O. (2010): *Social and economic networks.* Princeton University Press.

Knoll, B., N. Riedel, and E. Schlenker (2013): "He's a Chip Off the Old Block-The Persistence of Occupational Choices Across Generations," Discussion paper, CESifo Working Paper.

Kramarz, F., and O. N. Skans (2007): "With a little help from my... parents? family networks and youth labor market entry," *Unpublished paper, CREST.*

Long, J., and J. Ferrie (2013): "Intergenerational occupational mobility in Great Britain and the United States since 1850," *The American Economic Review*, 103(4), 1109–1137.

Moretti, E. (2004): "Workers' education, spillovers, and productivity: evidence from plant-level production functions," *American Economic Review*, 94(3), 656–690.

Reich, B. (2012): "Culture, Community and Segregation," Discussion paper, mimeo.

Stinson, M., and C. Wignall (2014): "Fathers, Children, and the Intergenerational Transmission of Employers," Discussion paper, U.S. Census Bureau.

Vega-Redondo, F. (2007): *Complex social networks.* Cambridge University Press.

A Appendix A: Equilibrium in the network game

A.1 Proof of Proposition 1

The FOC for the decentralized problem are

$$k_i = b_i + \frac{a}{2} b_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \text{ for all } i \tag{34}$$

$$s_i = \frac{a}{2} b_i \sqrt{\frac{k_i}{s_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \text{ for all } i \tag{35}$$

while the FOC for the social planner simplify to

$$k_i^s = b_i + a b_i \sqrt{\frac{s_i^s}{k_i^s}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j^s s_j^s}}{N^i} dj \text{ for all } i \tag{36}$$

$$s_i^s = a b_i \sqrt{\frac{k_i^s}{s_i^s}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j^s s_j^s}}{N^i} dj \text{ for all } i \tag{37}$$

We first prove that $\frac{k_i}{s_i} = \frac{k_j}{s_j}$ for all i and j.

We divide (34) by (35) to get

$$\frac{k_i}{s_i} = \frac{1 + \frac{a}{2}\sqrt{\frac{s_i}{k_i}} K(\mathbf{b}, \mathbf{k}, \mathbf{s})}{\frac{a}{2}\sqrt{\frac{k_i}{s_i}} K(\mathbf{b}, \mathbf{k}, \mathbf{s})} = \frac{\sqrt{\frac{k_i}{s_i}} + \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s})}{\frac{a}{2}\frac{k_i}{s_i} K(\mathbf{b}, \mathbf{k}, \mathbf{s})} \tag{38}$$

where bold face letters denote vectors and

$$K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj$$

Rearranging (38) gives

$$\left(\frac{k_i}{s_i}\right)^2 \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = \sqrt{\frac{k_i}{s_i}} + \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) \tag{39}$$

from which it is immediate that

$$\frac{k_i}{s_i} = F(K(\mathbf{b}, \mathbf{k}, \mathbf{s}))$$

for some $K(.)$ with a unique solution. To see the uniqueness notice that letting

$\sqrt{\frac{k_i}{s_i}} = x_i$ (39) can be written as

$$x_i^4 \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = x_i + \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) \tag{40}$$

the left hand side of (40) is a convex function taking the value 0 when $x_i = 0$ and the right hand side it is a linear and takes the positive value $\frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s})$ when $x_i = 0$. Hence there is a single crossing point at the positive orthant.

Hence

$$\begin{aligned} k_i &= b_i + \frac{a}{2} b_i \frac{K(\mathbf{b}, \mathbf{k}, \mathbf{s})}{\sqrt{F(K(\mathbf{b}, \mathbf{k}, \mathbf{s}))}} \text{ for all } i \\ s_i &= \frac{a}{2} b_i \sqrt{F(K(\mathbf{b}, \mathbf{k}, \mathbf{s}))} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) \text{ for all } i \end{aligned}$$

Thus it is clear we can write

$$\begin{aligned} k_i &= b_i k(\mathbf{b}, \mathbf{k}, \mathbf{s}) \text{ for all } i \\ s_i &= b_i s(\mathbf{b}, \mathbf{k}, \mathbf{s}) \text{ for all } i \end{aligned}$$

An analogous proof establishes that also for the centralized problem

$$\begin{aligned} k_i^s &= b_i k^s(\mathbf{b}, \mathbf{k}^s, \mathbf{s}^s) \text{ for all } i \\ s_i^s &= b_i s^s K^s(\mathbf{b}, \mathbf{k}^s, \mathbf{s}^s) \text{ for all } i \end{aligned}$$

It remains to determine the common optimal network parameters.

Using $k_i = b_i k$ and $s_i = b_i s$ it follows that $K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = \int_{j \in \mathcal{N}_i} \frac{b_j^2 \sqrt{ks}}{N^i} dj = \overline{b^2} \sqrt{ks}$ for the individual problem where

$$\overline{b^2} = \int_{j \in \mathcal{N}_i} \frac{b_j^2}{N^i} dj$$

and using $k_i^s = b_i k^s$ and $s_i^s = b_i s^s$ it follows that $K^s(\mathbf{b}, \mathbf{k}^s, \mathbf{s}^s) = \overline{b^2} \sqrt{k^s s^s}$ for the centralized problem.

Suppressing the dependence on the vectors, we get two simultaneous equations with two unknowns, namely

$$\begin{aligned} k &= 1 + \frac{a}{2}\sqrt{\frac{s}{k}}\overline{b^2}\sqrt{ks} = 1 + \frac{a}{2}\overline{b^2}s \\ s &= \frac{a}{2}\sqrt{\frac{k}{s}}\overline{b^2}\sqrt{ks} = \frac{a}{2}\overline{b^2}k \end{aligned}$$

for the decentralized problem and

$$\begin{aligned} k^s &= 1 + a\sqrt{\frac{s^s}{k^s}}\overline{b^2}\sqrt{k^s s^s} = 1 + a\overline{b^2}s^s \\ s^s &= a\sqrt{\frac{k^s}{s^s}}\overline{b^2}\sqrt{k^s s^s} = a\overline{b^2}k^s \end{aligned}$$

for the social planner. The optimal investments follow immediately from solving this system of linear equations. Assuming $a^2\overline{b^2}^2 < 1$ guarantees positive investment levels.

Introducing the optimal investment levels into the utility functions gives us

$$\begin{aligned} u_i(b_i) &= b_i^2 k + a b_i^2 k s \overline{b^2} - \frac{b_i^2}{2}k^2 - \frac{b_i^2}{2}s^2 \\ &= 2b_i^2 \frac{\left(4 + a^2\overline{b^2}^2\right)}{\left(4 - a^2\overline{b^2}^2\right)^2} \end{aligned}$$

for the decentralized solution and

$$\begin{aligned} u_i^s(b_i) &= b_i^2 k^s + a b_i^2 k^s s^s \overline{b^2} - \frac{b_i^2}{2}\left(k^s\right)^2 - \frac{b_i^2}{2}\left(s^s\right)^2 \\ &= \frac{b_i^2}{2}\left(\frac{1}{1 - a^2\overline{b^2}^2}\right). \end{aligned}$$

for the centralized solution.

A.2 Proof of Lemma 2

Let $B^M \geq CB^F$. Then assuming a uniform distribution on individual productivities between zero and B^l we can calculate $\overline{b^{F^2}}$ and $\overline{b^{M^2}}$.

$$\overline{b^{F^2}} = E\left(b_i^{F^2} \left| b_i^M < Cb_i^F\right.\right) = \frac{\int_0^{B^F} \int_0^{cb_F} b^{F^2} db^M db^F}{\int_0^{B^F} \int_0^{cb_F} db^M db^F} = \frac{CB^{F^4}}{2CB^{F^2}}$$

So

$$\overline{b^{F^2}} = \frac{B^{F^2}}{2} \tag{41}$$

$$\overline{b^{M^2}} = E\left(b_i^{M^2} \left| b_i^M > Cb_i^F\right.\right) = \frac{\int_0^{B^F} \int_{cb_F}^{B^M} b^{M^2} db^M db^F}{\int_0^{B^F} \int_{cb_F}^{B^M} db^M db^F} = \frac{1}{6}\frac{4B^{M^3} - C^3B^{F^3}}{2B^M - CB^F} \tag{42}$$

So

$$\overline{b^{M^2}} = \frac{1}{6}\left(B^{F^2}C^2 + 2B^F B^M C + 4B^{M^2} - \frac{4B^{M^3}}{2B^M - CB^F}\right) \tag{43}$$

We first show when $\overline{b^{M^2}}$ is maximized.

Lemma 5. *$\overline{b^{M^2}}$ is maximized at $C = \frac{B^M}{B^F}$ and obtains the value*

$$\overline{b^{M^2}_{\max}} = \frac{B^{M^2}}{2} \tag{44}$$

Proof. Observe that

$$\frac{\partial \overline{b^{M^2}}}{\partial C} = \frac{1}{6}\left(2CB^{F^2} + 2B^F B^M - \frac{4B^{M^3}B^F}{\left(2B^M - CB^F\right)^2}\right) \tag{45}$$

Since $2CB^{F^2} + 2B^F B^M$ is linear and $\frac{4B^{M^3}B^F}{(2B^M - CB^F)^2}$ is convex then $\frac{\partial \overline{b^{M^2}}}{\partial C} > 0$ provided it is positive for $C = 0$ and for $B^M = CB^F$. But $6\frac{\partial \overline{b^{M^2}}}{\partial C} = B^F B^M$ when $C = 0$ and $6\frac{\partial \overline{b^{M^2}}}{\partial C} = 0$ when $B^M = CB^F$. The solution is $C = \frac{B^M}{B^F}$ and substituting this value into the definition of $\overline{b^{M^2}}$ we obtain (44). Observe that we have additionally shown that

$$\frac{\partial \overline{b^{M^2}}}{\partial C} \geq 0 \text{ for } B^M \geq CB^F \text{ with strict equality when } B^M = CB^F \tag{46}$$

a result we will use later. □

Using the expressions derived for (41) and (42) we can calculate C_P and C_E. In the case of C_P the expression (14) becomes

$$C_P = \sqrt{\frac{\left(4+\left(a^F \frac{B^{F^2}}{2}\right)^2\right)}{\left(4-\left(a^F \frac{B^{F^2}}{2}\right)^2\right)^2}} \sqrt{\frac{\left(4-\left(a^M \frac{1}{6}\frac{4B^{M^3}-C_P^3 B^{F^3}}{2B^M - C_P B^F}\right)^2\right)^2}{\left(4+a^{M^2}\left(\frac{1}{6}\frac{4B^{M^3}-C_P^3 B^{F^3}}{2B^M - C_P B^F}\right)^2\right)}}$$

Rearranging we get

$$\frac{\left(4-\left(a^F \frac{B^{F^2}}{2}\right)^2\right)^2}{\left(4+\left(a^F \frac{B^{F^2}}{2}\right)^2\right)} C_P^2 = \frac{\left(4-\left(a^M \frac{1}{6}\frac{4B^{M^3}-C_P^3 B^{F^3}}{2B^M - C_P B^F}\right)^2\right)^2}{\left(4+a^{M^2}\left(\frac{1}{6}\frac{4B^{M^3}-C_P^3 B^{F^3}}{2B^M - C_P B^F}\right)^2\right)} \tag{47}$$

We define

$$F(C) \equiv \frac{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}$$

and check how it changes with the dividing line C.

$$\frac{\partial F(C)}{\partial C} = -\frac{12+a^{M^2}\overline{b^{M^2}}^2}{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)^2}\left(4-a^{M^2}\overline{b^{M^2}}^2\right) 2\overline{b^{M^2}}\frac{\partial \overline{b^{M^2}}}{\partial C} < 0$$

where the last inequality is true because we know that for equilibrium k and s to be well defined it is necessary that $4-a^{M^2}\overline{b^{M^2}}^2 > 0$ and $\frac{\partial \overline{b^{M^2}}}{\partial C} > 0$ in the relevant range. Hence the LHS of (47) is increasing in C while the RHS is decreasing in the relevant range, namely $B^M > CB^F$, so equilibrium when it exists is unique. Existence requires that for the maximum C, namely $C = \frac{B^M}{B^F}$

the RHS of (47) is smaller than the LHS. Since the value of $\overline{b^{M^2}}(C = \frac{B^M}{B^F}) = \overline{b^{M^2}_{\max}} = \frac{B^{M^2}}{2}$ by Lemma 5 existence requires that

$$\frac{\left(4+\left(a^M\frac{B^{M^2}}{2}\right)^2\right)}{\left(4-\left(a^M\frac{B^{M^2}}{2}\right)^2\right)^2}B^{M^2} > \frac{\left(4+\left(a^F\frac{B^{F^2}}{2}\right)^2\right)}{\left(4-\left(a^F\frac{B^{F^2}}{2}\right)^2\right)^2}B^{F^2}$$

but we can see that,

$$\frac{\partial\frac{\left(4+(aB)^2\right)B}{\left(4-(aB)^2\right)^2}}{\partial B} = \frac{24a^2B^2+16+a^4B^4}{\left(4-(aB)^2\right)^3} > 0$$

Therefore it holds that

$$\frac{\left(4+\left(a^M\frac{B^{M^2}}{2}\right)^2\right)}{\left(4-\left(a^M\frac{B^{M^2}}{2}\right)^2\right)^2}B^{M^2} > \frac{\left(4+\left(a^F\frac{B^{F^2}}{2}\right)^2\right)}{\left(4-\left(a^F\frac{B^{F^2}}{2}\right)^2\right)^2}B^{F^2} \tag{48}$$

and the equilibrium C^P exists. Observe that for the case where $a^M = a^F$ this holds iff $B^M > B^F$. If condition (48) were violated, we would have the opposite inequality and then an equilibrium would exist with $B^M \leq CB^F$. The equilibrium would then be defined using

$$\overline{b^{M^2}} = \frac{B^{M^2}}{2} \tag{49}$$

and

$$\overline{b^{F^2}} = \frac{1}{6}\frac{4B^{F^3} - C^3B^{M^3}}{2B^F - CB^M} \tag{50}$$

Similarly, we can use (41) and (43) to express C_E and rearranging we get

$$\left(1 - a^{F^2}\frac{B^{F^4}}{4}\right)C_E^2 = 1 - \frac{a^{M^2}}{36}\left(B^{F^2}C_E^2 + 2B^FB^MC_E + 4B^{M^2} - \frac{4B^{M^3}}{2B^M - C_EB^F}\right)^2 \tag{51}$$

The solution of which needs to be compared with the maximum in C of:

$$\overline{b^{M^2}} = B^{F^2}C_E^2 + 2B^F B^M C_E + 4B^{M^2} - \frac{4B^{M^3}}{2B^M - C_E B^F}$$

but notice that the LHS of (51) is increasing in C_E and the RHS is decreasing for $B^M > C_E B^F$ (as per 45) hence in the relevant range equilibrium, when it exists, is unique.

The condition for existence of C_E is that for the maximum possible $C = B^M/B^F$ the LHS (51) is higher than the RHS.

$$\frac{B^{M^2}}{1 - a^{M^2}\frac{B^{M^4}}{4}} > \frac{B^{F^2}}{1 - a^{F^2}\frac{B^{F^4}}{4}} \tag{52}$$

for the case where $a^M = a^F$ this holds iff $B^M > B^F$. If condition (52) were violated, we would have the opposite inequality and then an equilibrium would exist with $B^M \leq CB^F$.

A.3 Proof of Proposition 3

(i) The social planner would choose C to maximize social welfare with socially optimal investments in productive and socialization efforts where social welfare is given by

$$\begin{aligned} w(C) &= \frac{1}{B^F B^M}\left[\int_0^{B^F}\int_0^{Cb_i^F} \frac{b_i^{F^2}}{2}\left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right) db_i^M db_i^F \right. \\ &\quad \left. + \int_0^{B^F}\int_{Cb_i^F}^{B^M} \frac{b_i^{M^2}}{2}\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right) db_i^M db_i^F \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial w(C)}{\partial C} &= \frac{1}{B^F B^M}\left[\int_0^{B^F} \frac{b_i^{F^3}}{2}\left(\left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right) - C^2\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)\right) db_i^F \right. \\ &\quad \left. + \int_0^{B^F}\int_{Cb_i^F}^{B^M} \frac{\partial\left(\frac{b_i^{M^2}}{2}\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)\right)}{\partial C} db_i^M db_i^F \right] \end{aligned} \tag{53}$$

We already established when proving Lemma 5) that for all C

$$\int_0^{B^F}\int_{Cb_i^F}^{B^M}\frac{\partial\left(\frac{b_i^{M2}}{2}\left(\frac{1}{1-a^{M2}\overline{b^{M2}}^2}\right)\right)}{\partial C}db_i^M db_i^F>0 \tag{54}$$

by showing the integrand is positive as $\frac{\partial\overline{b^{M2}}}{\partial C}>0$ (46). Letting

$$H(C)=\frac{1}{B^FB^M}\left[\int_0^{B^F}\frac{b_i^{F3}}{2}\left(\left(\frac{1}{1-a^{F2}\overline{b^{F2}}^2}\right)-C^2\left(\frac{1}{1-a^{M2}\overline{b^{M2}}^2}\right)\right)db_i^F\right]$$

$$\frac{\partial H(C)}{\partial C}=\frac{1}{B^FB^M}\left[\int_0^{B^F}\frac{b_i^{F3}}{2}\left(-2C\left(\frac{1}{1-a^{M2}\overline{b^{M2}}^2}\right)-C^2\frac{\partial\left(\frac{1}{1-a^{M2}\overline{b^{M2}}^2}\right)}{\partial C}\right)db_i^F\right]$$

and again by Lemma 5 we know that

$$\int_0^{B^F}\frac{b_i^{F3}}{2}\left(\frac{\partial\left(\frac{1}{1-a^{M2}\overline{b^{M2}}^2}\right)}{\partial C}\right)db_i^F>0$$

so

$$\frac{\partial H(C)}{\partial C}<0. \tag{55}$$

It is also easy to see that for $C_E=\sqrt{\frac{1-a^{M2}\overline{b^{M2}}^2}{1-a^{F2}\overline{b^{F2}}^2}}$

$$H(C)|_{C=C_E}=0$$

and hence by (55) we have that $H(C)>0$ for $C<C_E$ and the result follows for C_E.

(ii) The social planner would choose C to maximize social welfare taking the optimal socialization and productive effort choices by individuals as given so that social welfare is given by

$$w(C) = \frac{1}{B^F B^M}\left[\int_0^{B^F}\int_0^{Cb_i^F} 2b_i^{F^2}\frac{\left(4+a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4-a^{F^2}\overline{b^{F^2}}^2\right)^2}db_i^M db_i^F + \int_0^{B^F}\int_{Cb_i^F}^{B^M} 2b_i^{M^2}\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}db_i^M db_i^F\right]$$

$$\frac{\partial w(C)}{\partial C} = \frac{1}{B^F B^M}\left[\int_0^{B^F} 2b_i^{F^3}\left(\frac{\left(4+a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4-a^{F^2}\overline{b^{F^2}}^2\right)^2} - C^2\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\right)db_i^F\right] + \int_0^{B^F}\int_{Cb_i^F}^{B^M}\frac{\partial\left(2b_i^{M^2}\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\right)}{\partial C}db_i^M db_i^F$$

We already established when proving Lemma 5) that for all C

$$\int_0^{B^F}\int_{Cb_i^F}^{B^M}\frac{\partial\left(2b_i^{M^2}\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\right)}{\partial C}db_i^M db_i^F > 0 \qquad (56)$$

by showing the integrand is positive as $\frac{\partial\overline{b^{M^2}}}{\partial C} > 0$ (46). Letting

$$H_P(C) = \frac{1}{B^F B^M}\left[\int_0^{B^F} 2b_i^{F^3}\left(\frac{\left(4+a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4-a^{F^2}\overline{b^{F^2}}^2\right)^2} - C^2\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\right)db_i^F\right]$$

$$\frac{\partial H_P(C)}{\partial C} = \frac{1}{B^F B^M}\left[\int_0^{B^F} 2b_i^{F^3}\left(-2C\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2} - C^2\frac{\partial\left(\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\right)}{\partial C}\right)db_i^F\right]$$

and again by Lemma 5 we know that

$$\int_0^{B^F} 2b_i^{F^3}\frac{\partial\left(\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\right)}{\partial C}db_i^F > 0$$

So

$$\frac{\partial H_P(C)}{\partial C} < 0. \tag{57}$$

It is also easy to see that for $C_P = \sqrt{\frac{\left(4+a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}\frac{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}{\left(4-a^{F^2}\overline{b^{F^2}}^2\right)^2}}$

$$H_P\left(C\right)|_{C=C_P} = 0$$

and hence by (57) we have that $H_P\left(C\right) > 0$ for $C < C_P$ and the result follows for C_P.

A.4 Pareto distribution of individual productivities

We assume now that returns b follow a Pareto distribution with shape parameter α

$$f\left(b\right) = \frac{\alpha}{b^{\alpha+1}} \text{ for } 1 \leq b \leq \infty$$

We will derive the results under the assumption that the C that defines the dividing line $b_i^M = C_E b_i^F$ is such that $C \geq 1$.[37]

[37] If $C < 1$, the same results hold with the names of the networks interchanged.

Lemma 6. *If $C \geq 1$, then*

$$\overline{b^{F^2}} = \frac{\alpha}{\alpha-2}\frac{1}{\alpha-1}\frac{(2\alpha-2)\,C^{\alpha}-(\alpha-2)}{2C^{\alpha}-1} \tag{58}$$

and

$$\overline{b^{M^2}} = \frac{\alpha}{\alpha-2}\frac{\alpha}{\alpha-1}C^2 \tag{59}$$

Proof. Let $C \geq 1$. Then

$$\overline{b^{F^2}} = E\left(b_i^{F^2}\left|b_i^M < Cb_i^F\right.\right) = \frac{\int_1^{\infty}\int_1^{Cb_F} b^{F^2}\frac{\alpha}{b^{F\alpha+1}}\frac{\alpha}{b^{M\alpha+1}}db^M db^F}{\int_1^{\infty}\int_1^{Cb_F}\frac{\alpha}{b^{F\alpha+1}}\frac{\alpha}{b^{M\alpha+1}}db^M db^F} = \frac{1}{\alpha}\left(1-\frac{1}{2C^{\alpha}}\right)$$

Similarly,

$$\overline{b^{M^2}} = E\left(b_i^{M^2}\left|b_i^M > Cb_i^F\right.\right) = \frac{\alpha}{\alpha-2}\frac{2\alpha}{2\alpha-2}C^2$$

□

Lemma 7. *The optimal choice C_E defined by (15) exists and is unique.*

Proof. Using the Lemma 6 C_E can be rewritten as:

$$C_E = \sqrt{\frac{1-a^{M^2}\left(\frac{\alpha}{\alpha-2}\frac{\alpha}{\alpha-1}C_E^2\right)^2}{1-a^{F^2}\left(\frac{\alpha}{\alpha-2}\frac{1}{\alpha-1}\left((\alpha-1)+\frac{1}{2C_E^{\alpha}-1}\right)\right)^2}} \tag{60}$$

Note that the LHS of (60) is increasing in C_E and the RHS is decreasing in C_E so that a unique equilibrium exists. □

Moreover,

Lemma 8. $C_E > 1 \Leftrightarrow a^{M^2} < a^{F^2}$

Proof. Note also that if $a^{M^2} = a^{F^2}$ the solution of (60) is at $C^E = 1$. An increase of a^{M^2} with respect to a^{F^2} displaces the RHS to the left so that the new equilibrium entails $C^E < 1$. □

We are now in a position to check how a decentralized network choice deviates from the efficient network choice C^S implemented by a social planner who maximizes social welfare. We study the case where the social planner also implements the socially optimal investments in productive and socialization effort.

Proposition 9. *If $C_E > 1$, there might be too few $\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} > 0$ or too many people $\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} < 0$ in the F network compared to the social optimum. The alternative network will be underpopulated if and only if*

$$\frac{a^{M^2}}{a^{F^2}} > \frac{\left((2\alpha-2)\,C^{\alpha}-(\alpha-2)\right)^2 C^{\alpha}}{\alpha^2\left(2C^{\alpha}-1\right)^3 C^2} \tag{61}$$

Proof. The social planner would choose C to maximize social welfare with socially optimal investments in productive and socialization efforts where social welfare is given by

$$\begin{aligned} w(C) \;=\; & \int_1^{\infty}\int_1^{Cb_F}\frac{b_i^{F^2}}{2}\left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right)\frac{\alpha}{b_i^{F^{\alpha+1}}}\frac{\alpha}{b_i^{M^{\alpha+1}}}db_i^M db_i^F \\ & +\int_1^{\infty}\int_{Cb_F}^{\infty}\frac{b_i^{M^2}}{2}\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)\frac{\alpha}{b_i^{F^{\alpha+1}}}\frac{\alpha}{b_i^{M^{\alpha+1}}}db_i^M db_i^F \end{aligned}$$

$$\begin{aligned} & \frac{\partial w(C)}{\partial C} \\ = & \left[\int_1^{\infty}\frac{b_i^{F^3}}{2}\left(\left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right)-C^2\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)\right)\frac{\alpha}{b_i^{F^{\alpha+1}}}\frac{\alpha}{\left(Cb_i^F\right)^{\alpha+1}}db_i^F\right] \\ & +\int_1^{\infty}\int_1^{Cb_F}\frac{\partial\left(\frac{b_i^{F^2}}{2}\left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right)\right)}{\partial C}\frac{\alpha}{b_i^{F^{\alpha+1}}}\frac{\alpha}{b_i^{M^{\alpha+1}}}db_i^M db_i^F \\ & +\int_1^{\infty}\int_{Cb_F}^{\infty}\frac{\partial\left(\frac{b_i^{M^2}}{2}\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)\right)}{\partial C}\frac{\alpha}{b_i^{F^{\alpha+1}}}\frac{\alpha}{b_i^{M^{\alpha+1}}}db_i^M db_i^F \end{aligned}$$

Now at $C_E = \sqrt{\frac{1-a^{M^2}\overline{b^{M^2}}^2}{1-a^{F^2}\overline{b^{F^2}}^2}}$

$$\begin{aligned}\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} &= a^{F^2}2\overline{b^{F^2}}\frac{\alpha}{\alpha-2}\frac{1}{\alpha-1}\left(\frac{-2\alpha C^{\alpha-1}}{(2C^{\alpha}-1)^2}\right)\left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right)^2\frac{\alpha\left(\left(2\alpha-2\right)C^{\alpha}-(\alpha-2)\right)}{\left(2\alpha-2\right)\left(\alpha-2\right)C^{\alpha}} \\ &+a^{M^2}2\overline{b^{M^2}}\frac{\alpha}{\alpha-2}\frac{\alpha}{\alpha-1}2C\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)^2\frac{\alpha^2}{\alpha-2}\frac{1}{C^{\alpha-2}}\frac{1}{(2\alpha-2)}\end{aligned}$$

$C_E = \sqrt{\frac{1-a^{M^2}\overline{b^{M^2}}^2}{1-a^{F^2}\overline{b^{F^2}}^2}} \rightarrow C^4\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)^2 = \left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right)^2$

Therefore

$$\begin{aligned}\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} &> 0 \Longleftrightarrow -a^{F^2}\frac{\left(\left(2\alpha-2\right)C^{\alpha}-(\alpha-2)\right)^2}{\left(2C^{\alpha}-1\right)^3 C}+a^{M^2}\frac{\alpha^2}{C^{\alpha-1}}>0 \\ &\Longleftrightarrow \frac{a^{M^2}}{a^{F^2}}>\frac{\left(\left(2\alpha-2\right)C^{\alpha}-(\alpha-2)\right)^2 C^{\alpha}}{\alpha^2\left(2C^{\alpha}-1\right)^3 C^2}\end{aligned}$$

and

$$\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E}<0 \Longleftrightarrow \frac{a^{M^2}}{a^{F^2}}<\frac{\left(\left(2\alpha-2\right)C^{\alpha}-(\alpha-2)\right)^2 C^{\alpha}}{\alpha^2\left(2C^{\alpha}-1\right)^3 C^2}$$

By Lemma 8 since $C_E > 1 \Leftrightarrow a^{M^2} < a^{F^2}$, hence $\frac{a^{M^2}}{a^{F^2}} < 1$.

We will now show that

$$1>\frac{\left(\left(2\alpha-2\right)C^{\alpha}-(\alpha-2)\right)^2 C^{\alpha}}{\alpha^2\left(2C^{\alpha}-1\right)^3 C^2}=\frac{\left(\left(\alpha-1\right)\left(2C^{\alpha}-1\right)+1\right)^2 C^{\alpha}}{\alpha^2\left(2C^{\alpha}-1\right)^3 C^2} \tag{62}$$

Note that

$$\left(\left(\alpha-1\right)\left(2C^{\alpha}-1\right)+1\right)^2<\alpha^2\left(2C^{\alpha}-1\right)^2$$

since that expression is equivalent to

$$\begin{aligned}\left(\alpha-1\right)\left(2C^{\alpha}-1\right)+1 &< \alpha\left(2C^{\alpha}-1\right) \\ &\Leftrightarrow 1<2C^{\alpha}-1 \Longleftrightarrow 1<C^{\alpha}\end{aligned}$$

thus

$$\frac{\left(\left(\alpha-1\right)\left(2C^{\alpha}-1\right)+1\right)^2 C^{\alpha}}{\alpha^2\left(2C^{\alpha}-1\right)^3 C^2}<\frac{C^{\alpha}}{\left(2C^{\alpha}-1\right)C^2}<\frac{1}{C}<1 \tag{63}$$

where the last two inequalities hold since $C > 1$, noting that in that case $2C^\alpha - 1 > C^\alpha$. Thus equation (63) establishes (62).

Lemmas 9 shows that parameter values exist so that $\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} < 0$. while Lemma 10 shows the existence of parameter values that $\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} > 0$.

Lemma 9. *Let* $\frac{a^{M^2}}{a^{F^2}} = r < 1$. *For a fixed* α *and* r *there exists an* a^{F^2} *low enough that*

$$r = \frac{a^{M^2}}{a^{F^2}} < \frac{\left((2\alpha - 2)\, C^\alpha - (\alpha - 2)\right)^2 C^\alpha}{\alpha^2 \left(2C^\alpha - 1\right)^3 C^2}$$

Proof. Since

$$C_E = \sqrt{\frac{1 - r a^{F^2}\left(\frac{\alpha}{\alpha-2}\frac{\alpha}{\alpha-1} C_E^2\right)^2}{1 - a^{F^2}\left(\frac{\alpha}{\alpha-2}\frac{1}{\alpha-1}\left((\alpha-1) + \frac{1}{2C_E^\alpha - 1}\right)\right)^2}}$$

we have that

$$\lim_{a^{F^2}\to 0} C_E\left(\alpha, r, a^{F^2}\right) = 1$$

thus

$$\lim_{a^{F^2}\to 0} \frac{\left((2\alpha - 2)\, C^\alpha - (\alpha - 2)\right)^2 C^\alpha}{\alpha^2 \left(2C^\alpha - 1\right)^3 C^2} = \lim_{a^{F^2}\to 0} \frac{\left((2\alpha - 2) - (\alpha - 2)\right)^2}{\alpha^2} = 1 > r.$$

□

Lemma 10. *Let* $\frac{a^{M^2}}{a^{F^2}} = r < 1$. *For a fixed* a^{F^2} *and* r *such that* C_E *exists, there is an* α *high enough that*

$$r = \frac{a^{M^2}}{a^{F^2}} > \frac{\left((2\alpha - 2)\, C_E^\alpha - (\alpha - 2)\right)^2 C_E^\alpha}{\alpha^2 \left(2C_E^\alpha - 1\right)^3 C_E^2}$$

Proof. For a bounded C_E

$$C \equiv \lim_{\alpha\to\infty} C_E^2 = \lim_{a^{F^2}\to 0} \frac{1 - r a^{F^2} C_E^4}{1 - a^{F^2}}$$

Hence

$$ra^{F^2}C^4 + \left(1 - a^{F^2}\right)C^2 - 1 = 0$$

and thus

$$C^2 = \frac{-\left(1 - a^{F^2}\right) \pm \sqrt{\left(1 - a^{F^2}\right)^2 + 4ra^{F^2}}}{2ra^{F^2}}$$

Now since

$$\lim_{\alpha\to\infty} \frac{\left(\left(2\alpha - 2\right)C_E^\alpha - (\alpha - 2)\right)^2 C_E^\alpha}{\alpha^2\left(2C_E^\alpha - 1\right)^3 C_E^2} = \lim_{\alpha\to\infty} \frac{\left(2C^\alpha - 1\right)^2 \alpha^2 C^\alpha}{\alpha^2\left(2C^\alpha - 1\right)^3 C^2} = \frac{1}{2C^2}$$

In other words, we would like to show that for α high enough

$$C^2 > \frac{1}{2r}$$

or

$$\frac{-\left(1 - a^{F^2}\right) + \sqrt{\left(1 - a^{F^2}\right)^2 + 4ra^{F^2}}}{2ra^{F^2}} > \frac{1}{2r} \tag{64}$$

$$\begin{aligned}\sqrt{\left(1 - a^{F^2}\right)^2 + 4ra^{F^2}} &> 1\\ a^{F^2}\left(a^{F^2} + 4r - 2\right) &> 0\end{aligned}$$

which requires $r > \frac{2 - a^{F^2}}{4}$ which is true for example if $r > \frac{1}{2}$. □

Proposition 9 immediately follows from these Lemmas. □

B Appendix B: Congestion

B.1 Proof of lemma 3

Proof. Under congestion, the welfare of the network F remains unchanged while the welfare of network M is given by

$$W_E^M(C) = \frac{1}{B^F B^M} \int_0^{B^F} \int_{Cb_i^F}^{B^M} f\left(C, b_i^M\right) \frac{b_i^{M^2}}{2} \left(\frac{1}{1 - a^{M^2}\overline{b^{M^2}}^2}\right) db_i^M db_i^F \tag{65}$$

when the government induces efficient socialization and productive efforts within a network and by

$$W_P^M(C) = \frac{1}{B^F B^M} \int_0^{B^F} \int_{Cb_i^F}^{B^M} f\left(C, b_i^M\right) 2b_i^{M^2} \frac{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2} db_i^M db_i^F \quad (66)$$

These expressions (65) and (66) can be decomposed in the welfare of those member of M below $b_i^M < CB^F$ for whom congestion does not matter and those above $b_i^M \geq CB^F$ for whom congestion impinges. For the case where the government induces efficient efforts

$$\begin{aligned} W_E^M(C) = & \frac{1}{B^F B^M} \int_0^{B^F} \int_{Cb_i^F}^{CB^F} \frac{b_i^{M^2}}{2} \left(\frac{1}{1 - a^{M^2}\overline{b^{M^2}}^2} \right) db_i^M db_i^F \\ & + \frac{1}{B^F B^M} \int_0^{B^F} \int_{CB^F}^{B^M} \left(\frac{\left(C^* B^F\right)^2}{b_i^{M^2}} + (1 - v(C)) \left(1 - \frac{\left(C^* B^F\right)^2}{b_i^{M^2}} \right) \right) \frac{b_i^{M^2}}{2} \left(\frac{1}{1 - a^{M^2}\overline{b^{M^2}}^2} \right) db_i^M db_i^F \end{aligned}$$

The second line captures welfare of those for whom congestion matters. After some calculations this second line becomes

$$\begin{aligned} & \frac{\left(C^* B^F\right)^2}{2} \left(\frac{1}{1 - a^{M^2}\overline{b^{M^2}}^2} \right) \frac{\left(B^M - CB^F\right) B^F}{B^F B^M} \\ & + \frac{(1 - v(C))}{B^F B^M} \left(\int_0^{B^F} \int_{CB^F}^{B^M} \frac{b_i^{M^2}}{2} \left(\frac{1}{1 - a^{M^2}\overline{b^{M^2}}^2} \right) db_i^M db_i^F \right. \\ & \left. - \frac{\left(B^M - CB^F\right) B^F \left(C^* B^F\right)^2}{2} \left(\frac{1}{1 - a^{M^2}\overline{b^{M^2}}^2} \right) \right) \end{aligned}$$

which decomposes the welfare of people beyond the $b_i^M \geq C_E^* B^F$ boundary in two parts. First the welfare for types exactly at the boundary is $\frac{\left(CB^F\right)^2}{2} \left(\frac{1}{1 - a^{M^2}\overline{b^{M^2}}^2} \right)$ times the fraction of people in that area is $\frac{\left(B^M - CB^F\right) B^F}{B^F B^M}$,which gives the first line. And the second line is the surplus welfare for those types, in addition to what the boundary types get, and on which the congestion impinges. Using (17) we can now write total welfare when the government induces efficient socialization and productive efforts as

$$
\begin{aligned}
w_E(C) \quad = \quad & \frac{CB^{F^3}}{8B^M}\left[\left(\frac{1}{1-a^{F^2}\overline{b^{F^2}}^2}\right)+C^2\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)\right] \\
& +\left(1-C\frac{B^F}{B^M}\right)\frac{\left(CB^F\right)^2}{2}\left(\frac{1}{1-a^{M^2}\overline{b^{M^2}}^2}\right)+\left(1-v\left(C\right)\right)G_E\left(C\right)
\end{aligned}
$$

Welfare in absence of any government intervention $w_P(C)$ is derived in a parallel way. □

B.2 Proof of lemma 4

Proof. We first show that it suffices to have $1 - a^{M^2}\frac{B^{M^4}}{9} \approx 0$ (sufficiently small), to have $C_E = \varepsilon \approx 0$ (very small). Observe that when $C_E \approx 0$

$$\overline{b^{M^2}} = \frac{1}{6}\frac{4B^{M^3} - C^3 B^{F^3}}{2B^M - CB^F} \simeq \frac{B^{M^2}}{3} \tag{67}$$

Hence

$$C_E = \sqrt{\frac{1 - a^{M^2}\overline{b^{M^2}}^2}{1 - a^{F^2}\overline{b^{F^2}}^2}} \tag{68}$$

in which case C_E small by having

$$C_E^2 \simeq \frac{1 - a^{M^2}\frac{B^{M^4}}{9}}{1 - a^{F^2}\frac{B^{F^4}}{4}} \tag{69}$$

and thus it suffices to have

$$1 - a^{M^2}\frac{B^{M^4}}{9} \simeq \varepsilon^2 \tag{70}$$

small to have C_E small.

Next we show that $C_p > 0$ for these parameter values (70). Recall that

$$C_P = \sqrt{\frac{\left(4 + a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)} \frac{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2}{\left(4 - a^{F^2}\overline{b^{F^2}}^2\right)^2}} \tag{71}$$

Assume for contradiction that $C_P = \varepsilon \approx 0$. Since $\overline{b^{M^2}} = \frac{1}{6}\frac{4B^{M^3} - C^3 B^{F^3}}{2B^M - CB^F}$, in this case $\overline{b^{M^2}}^2 = \frac{B^{M^4}}{9}$. But using (70) we get

$$4 - a^{M^2}\frac{B^{M^4}}{9} = 3 + \underbrace{1 - a^{M^2}\frac{B^{M^4}}{9}}_{\approx \varepsilon^2} \simeq 3$$

$$4 + a^{M^2}\overline{b^{M^2}}^2 = 5 - \left(\underbrace{1 - a^{M^2}\frac{B^{M^4}}{9}}_{\approx \varepsilon^2}\right) \simeq 5$$

and so

$$C_P = \sqrt{\frac{\left(4 + a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4 + a^{M^2}\overline{b^{M^2}}^2\right)} \frac{\left(4 - a^{M^2}\overline{b^{M^2}}^2\right)^2}{\left(4 - a^{F^2}\overline{b^{F^2}}^2\right)^2}} \simeq \sqrt{\frac{\left(4 + a^{F^2}\frac{B^{F^4}}{4}\right)}{\left(4 - a^{F^2}\frac{B^{F^4}}{4}\right)^2}\frac{9}{5}} > 0 \tag{72}$$

which contradicts our assumption that $C_P = \varepsilon \approx 0$. Hence $C_P \neq 0$.

Observe that rewriting (15) as

$$\frac{C_E^2}{1 - a^{M^2}\frac{B^{M^4}}{9}} \simeq \frac{1}{1 - a^{F^2}\frac{B^{F^4}}{4}} \tag{73}$$

we can express welfare when optimal socialization and productive efforts are induced in the network (Lemma 3) when (70) holds and hence C_E is very small as

$$w_E(C_E) = \frac{CB^{F^3}}{8B^M}\left[\left(\frac{1}{1-a^{F^2}\frac{B^{F^4}}{4}}\right)+\frac{1}{1-a^{F^2}\frac{B^{F^4}}{4}}\right]+ \left(1-C\frac{B^F}{B^M}\right)\frac{B^{F^2}}{2}\frac{1}{1-a^{F^2}\frac{B^{F^4}}{4}}+(1-v(C))\,G_E(C)$$

which reduces to

$$w_E(C_E \approx 0) \approx \frac{B^{F^2}}{2}\frac{1}{1-a^{F^2}\frac{B^{F^4}}{4}}. \tag{74}$$

To calculate welfare without any government intervention $w_P(C_P)$ recall that $C_P \neq 0$ and hence $v(C_P) = 0$. Hence

$$\begin{aligned} w_P(C_P) &= \frac{CB^{F^3}}{2B^M}\left[\frac{\left(4+a^{F^2}\frac{B^{F^4}}{4}\right)}{\left(4-a^{F^2}\frac{B^{F^4}}{4}\right)^2}+\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}C^2\right] \\ &\quad +2\left(1-C\frac{B^F}{B^M}\right)\frac{\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\left(CB^F\right)^2+G_P(C) \\ &= \frac{1}{B^F B^M}\left[\frac{4\left(4+a^{F^2}\frac{B^{F^4}}{4}\right)}{\left(4-a^{F^2}\frac{B^{F^4}}{4}\right)^2}\frac{C_P B^{F^4}}{8}+\frac{4\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2}\left(\frac{B^{M^3}B^F}{6}-C^3\frac{B^{F^4}}{24}\right)\right] \end{aligned}$$

which coincides with the expression for welfare without congestion. Indeed $v(C_P) = 0$ is equivalent to no congestion in the network.

Note that

$$\frac{4\left(4+a^{M^2}\overline{b^{M^2}}^2\right)}{\left(4-a^{M^2}\overline{b^{M^2}}^2\right)^2} = \frac{4\left(4+a^{M^2}\left(\frac{1}{6}\frac{4B^{M^3}-C^3B^{F^3}}{2B^M-CB^F}\right)^2\right)}{\left(4-a^{M^2}\left(\frac{1}{6}\frac{4B^{M^3}-C^3B^{F^3}}{2B^M-CB^F}\right)^2\right)^2}$$

is increasing in C so we can have a bound on that term by taking $C_P = 0$ and on $C^3\frac{B^{F^4}}{24}$ by taking $C_P = 1$

$$
\begin{aligned}
w(C_P) &= \frac{1}{B^F B^M}\left[\frac{4\left(4+a^{F^2}\overline{b^{F^2}}^2\right)}{\left(4-a^{F^2}\overline{b^{F^2}}^2\right)^2}\frac{C_P B^{F^4}}{8}+\frac{4\left(4+a^{M^2}\left(\frac{1}{6}\frac{4B^{M^3}-C^3B^{F^3}}{2B^M-CB^F}\right)^2\right)}{\left(4-a^{M^2}\left(\frac{1}{6}\frac{4B^{M^3}-C^3B^{F^3}}{2B^M-CB^F}\right)^2\right)^2}\left(\frac{B^{M^3}B^F}{6}-C_P^3\frac{B^{F^4}}{24}\right)\right] \\
&\geq \frac{1}{B^F B^M}\left[\frac{4\left(4+a^{M^2}\left(\frac{B^{M^2}}{3}\right)^2\right)}{\left(4-a^{M^2}\left(\frac{B^{M^2}}{3}\right)^2\right)^2}\left(\frac{B^{M^3}B^F}{6}-\frac{B^{F^4}}{24}\right)\right]
\end{aligned}
$$

We know by (70) that $a^{M^2}\left(\frac{B^{M^2}}{3}\right)^2 \simeq 1$, so then

$$
w(C_P) \geq \frac{1}{B^F B^M}\left[\frac{20}{9}\left(\frac{B^{M^3}B^F}{6}-\frac{B^{F^4}}{24}\right)\right]
$$

Using (74)

$$
\begin{aligned}
w(C_E)-w(C_P) &\leq \frac{1}{2}\times\frac{B^{F^2}}{1-a^{F^2}\frac{B^{F^4}}{4}}-\frac{1}{B^F B^M}\left[\frac{20}{9}\left(\frac{B^{M^3}B^F}{6}-\frac{B^{F^4}}{24}\right)\right] \\
&= \frac{1}{2}\frac{B^{F^2}}{1-a^{F^2}\frac{B^{F^4}}{4}}+\frac{5}{54}\frac{B^{F^4}}{B^M}-\frac{10}{27}B^{M^2}
\end{aligned}
$$

which is negative for B^M big enough. □

C Appendix C: Parental investment

Lemma 11. *Equation (23) can be written as a function with at most two branches:*

$$
E\left[\max\left[2b_i^{F^2}F, 2b_i^{M^2}M\right]\right] = \begin{cases} 2F\frac{e_{p_i}^{F^3}}{6e_{p_i}^M}\sqrt{\frac{F}{M}}+2M\frac{e_{p_i}^{M^2}}{3} & \text{if } e_{p_i}^M > \sqrt{\frac{F}{M}}e_{p_i}^F \\ 2M\frac{e_{p_i}^{M^3}}{6e_{p_i}^F}\sqrt{\frac{M}{F}}+2F\frac{e_{p_i}^{F^2}}{3} & \text{if } e_{p_i}^M < \sqrt{\frac{F}{M}}e_{p_i}^F \end{cases} \tag{75}
$$

where the two branches exist if the inequalities in (75) are non-empty.

Proof. Under the assumption that

$$e^M_{p_i} > \sqrt{\frac{F}{M}} e^F_{p_i}$$

$E\left[\max\left[2b_i^{F^2}F, 2b_i^{M^2}M\right]\right]$ becomes

$$\begin{aligned} & E\left[\max\left[2b_i^{F^2}F, 2b_i^{M^2}M\right]\right] \\ = & \frac{1}{e^F_{p_i}e^M_{p_i}}\left(2F\int_0^{e^F_{p_i}}\int_0^{b_i^F\sqrt{\frac{F}{M}}} b_i^{F^2}db_i^M db_i^F + 2M\int_0^{e^F_{p_i}}\int_{b_i^F\sqrt{\frac{F}{M}}}^{e^M_{p_i}} b_i^{M^2}db_i^M db_i^F\right) \\ = & \frac{1}{e^F_{p_i}e^M_{p_i}}\left(2F\frac{e^{F^4}_{p_i}}{6}\sqrt{\frac{F}{M}} + 2M\frac{e^{M^3}_{p_i}e^F_{p_i}}{3}\right) \end{aligned}$$

so that

$$E\left[\max\left[2b_i^{F^2}F, 2b_i^{M^2}M\right]\right] = 2F\frac{e^{F^3}_{p_i}}{6e^M_{p_i}}\sqrt{\frac{F}{M}} + 2M\frac{e^{M^2}_{p_i}}{3} \tag{76}$$

Suppose instead that

$$e^M_{p_i} < \sqrt{\frac{F}{M}} e^F_{p_i}$$

then

$$E\left[\max\left[2b_i^{F^2}F, 2b_i^{M^2}M\right]\right] = 2M\frac{e^{M^3}_{p_i}}{6e^F_{p_i}}\sqrt{\frac{M}{F}} + 2F\frac{e^{F^2}_{p_i}}{3} \tag{77}$$

□

Thus, for a parent who belongs to network F, $x^F_{p_i} + x^M_{p_i} = K$, $e^F_{p_i} = \overline{A} + x^F_{p_i}$ and $e^M_{p_i} = \underline{A} + K - x^F_{p_i}$. It immediately follows:

Lemma 12.

$$x^F_{p_i} < \frac{K + \left(\underline{A} - \overline{A}\sqrt{\frac{F}{M}}\right)}{1 + \sqrt{\frac{F}{M}}} \quad \textit{if } e^M_{p_i} > \sqrt{\tfrac{F}{M}} e^F_{p_i} \tag{78}$$

$$x^F_{p_i} > \frac{K + \left(\underline{A} - \overline{A}\sqrt{\frac{F}{M}}\right)}{1 + \sqrt{\frac{F}{M}}} \quad \textit{if } e^M_{p_i} < \sqrt{\tfrac{F}{M}} e^F_{p_i} \tag{79}$$

Analogously, for a parent who belongs to network M, $x^F_{p_i} + x^M_{p_i} = K$, $e^F_{p_i} = \underline{A} + x^F_{p_i}$ and $e^M_{p_i} = \overline{A} + K - x^F_{p_i}$. It immediately follows:

Lemma 13.

$$x^F_{p_i} < \frac{K + \left(\overline{A} - \underline{A}\sqrt{\frac{F}{M}}\right)}{1 + \sqrt{\frac{F}{M}}} \quad \textit{if } e^M_{p_i} > \sqrt{\frac{F}{M}} e^F_{p_i} \tag{80}$$

$$x^F_{p_i} > \frac{K + \left(\overline{A} - \underline{A}\sqrt{\frac{F}{M}}\right)}{1 + \sqrt{\frac{F}{M}}} \quad \textit{if } e^M_{p_i} < \sqrt{\frac{F}{M}} e^F_{p_i} \tag{81}$$

C.1 Convexity of $g_1\left(x^F_{p_i}\right)$

Lemma 14. *The function $g_1\left(x^F_{p_i}\right)$ is convex.*

Proof. For $x^F_{p_i} < \frac{K + \left(\underline{A} - \overline{A}\sqrt{\frac{F}{M}}\right)}{1 + \sqrt{\frac{F}{M}}}$

$$\frac{\partial g_1\left(x^F_{p_i}\right)}{\partial x^F_{p_i}} = -\frac{4}{3} M\left(\underline{A} + K - x^F_{p_i}\right) + 6F\left(\overline{A} + x^F_{p_i}\right)^2 \frac{\sqrt{\frac{F}{M}}}{6\left(\underline{A} + K - x^F_{p_i}\right)}$$
$$+12F\left(\overline{A} + x^F_{p_i}\right)^3 \frac{\sqrt{\frac{F}{M}}}{6\left(\underline{A} + K - x^F_{p_i}\right)^2}$$

Since

$$6F\left(\overline{A} + x^F_{p_i}\right)^2 \frac{\sqrt{\frac{F}{M}}}{6\left(\underline{A} + K - x^F_{p_i}\right)} + 12F\left(\overline{A} + x^F_{p_i}\right)^3 \frac{\sqrt{\frac{F}{M}}}{6\left(\underline{A} + K - x^F_{p_i}\right)^2}$$

is increasing in $x^F_{p_i}$ this implies that

$$\frac{\partial^2 g_1\left(x^F_{p_i}\right)}{\partial x^{F^2}_{p_i}} > 0$$

so $g_1(.)$ is convex.

For $x_{p_i}^F > \frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}$

$$\frac{\partial g_1\left(x_{p_i}^F\right)}{\partial x_{p_i}^F} == 4F\frac{\left(\overline{A}+x_{p_i}^F\right)}{3}-\sqrt{\frac{M}{F}}\left(M\frac{\left(\underline{A}+K-x_{p_i}^F\right)^2}{\left(\overline{A}+x_{p_i}^F\right)}+2M\frac{\left(\underline{A}+K-x_{p_i}^F\right)^3}{\left(\overline{A}+x_{p_i}^F\right)^2}\right)$$

and since $\left(M\frac{\left(\underline{A}+K-x_{p_i}^F\right)^2}{\left(\overline{A}+x_{p_i}^F\right)}+2M\frac{\left(\underline{A}+K-x_{p_i}^F\right)^3}{\left(\overline{A}+x_{p_i}^F\right)^2}\right)$ is decreasing in $x_{p_i}^F$ it is easy to see that

$$\frac{\partial^2 g_1\left(x_{p_i}^F\right)}{\partial x_{p_i}^{F2}} > 0$$

□

C.2 Nonexistence of an interior solution

We need to compare the value of the parental objective functions at three possible points. For F-parents we will need to establish whether either $x_{p_i}^F = 0$ or $x_{p_i}^F = \frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}$ are optimal in the range $x_{p_i}^F < \frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}$ and whether either $x_{p_i}^F = K$ or $x_{p_i}^F = \frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}$ are optimal in the range $x_{p_i}^F > \frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}$. Therefore the existence of an internal optimally global solution implying that F-parents invest in both networks requires first that these extreme points are defined, namely

$$\frac{K+\underline{A}}{\overline{A}} > \sqrt{\frac{F}{M}} > \frac{\underline{A}}{K+\overline{A}} \tag{82}$$

and that

$$g\left(\frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}\right) > g(0) \quad \text{if } e_{p_i}^M > \sqrt{\frac{F}{M}}e_{p_i}^F$$

$$g\left(\frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}\right) > g(K) \quad \text{if } e_{p_i}^M < \sqrt{\frac{F}{M}}e_{p_i}^F$$

Since,

$$g(0) = 2F\frac{\overline{A}^3}{6\left(\underline{A}+K\right)}\sqrt{\frac{F}{M}}+2M\frac{\left(\underline{A}+K\right)^2}{3}$$

$$g\left(\frac{K+\left(\underline{A}-\overline{A}\sqrt{\frac{F}{M}}\right)}{1+\sqrt{\frac{F}{M}}}\right) = F\left(\frac{\overline{A}+\underline{A}+K}{1+\sqrt{\frac{F}{M}}}\right)^2$$

$$g(K) = 2M\frac{\underline{A}^3}{6\left(\overline{A}+K\right)}\sqrt{\frac{M}{F}}+2F\frac{\left(\overline{A}+K\right)^2}{3}$$

Thus, the conditions for optimality of an internal global solution when it exists are:

$$F\left(\frac{\overline{A}+\underline{A}+K}{1+\sqrt{\frac{F}{M}}}\right)^2 > 2F\frac{\overline{A}^3}{6\left(\underline{A}+K\right)}\sqrt{\frac{F}{M}}+2M\frac{\left(\underline{A}+K\right)^2}{3} \tag{83}$$

$$F\left(\frac{\overline{A}+\underline{A}+K}{1+\sqrt{\frac{F}{M}}}\right)^2 > 2M\frac{\underline{A}^3}{6\left(\overline{A}+K\right)}\sqrt{\frac{M}{F}}+2F\frac{\left(\overline{A}+K\right)^2}{3} \tag{84}$$

Or equivalently

$$3\frac{F}{M}\left(\frac{\overline{A}+\underline{A}+K}{1+\sqrt{\frac{F}{M}}}\right)^2 > \frac{F}{M}\frac{\overline{A}^3}{(\underline{A}+K)}\sqrt{\frac{F}{M}}+2\left(\underline{A}+K\right)^2 \tag{85}$$

$$3\frac{F}{M}\left(\frac{\overline{A}+\underline{A}+K}{1+\sqrt{\frac{F}{M}}}\right)^2 > \frac{\underline{A}^3}{(\overline{A}+K)}\sqrt{\frac{M}{F}}+2\frac{F}{M}\left(\overline{A}+K\right)^2 \tag{86}$$

If either of these conditions does not hold, then parents will put all their available effort in developing the abilities of one network only.

We now look when an interior solution for M-parents exists. To be defined it requires that

$$\frac{K+\underline{A}}{\overline{A}} > \frac{1}{\sqrt{\frac{F}{M}}} > \frac{\underline{A}}{K+\overline{A}} \tag{87}$$

Moreover, the interior solution should maximize the parental objective function which requires

$$3\frac{F}{M}\left(\frac{\overline{A}+\underline{A}+K}{1+\sqrt{\frac{F}{M}}}\right)^2 > \frac{F}{M}\frac{\underline{A}^3}{(\overline{A}+K)}\sqrt{\frac{F}{M}}+2\left(\overline{A}+K\right)^2 \tag{88}$$

$$3\frac{F}{M}\left(\frac{\overline{A}+\underline{A}+K}{1+\sqrt{\frac{F}{M}}}\right)^2 > \frac{\overline{A}^3}{(\underline{A}+K)}\sqrt{\frac{M}{F}}+2\frac{F}{M}\left(\underline{A}+K\right)^2 \tag{89}$$

Proposition 10. *An interior solution does not exist for* $\underline{A}=0$

Proof. We will prove this for F-parents here (The proof for M-parents is analogous). Condition (85) tells us when the interior solution for F-parents is better than no investment in F (full investment in M). Condition (86) tells us when the interior solution for F-parents is better than full investment in F. For the interior and the M corner solution to exist condition (82) is required. The proof consists in showing that these conditions are incompatible. We first rewrite condition (85) as

$$3\left(\frac{1+\frac{\underline{A}+K}{\overline{A}}}{1+\sqrt{\frac{F}{M}}}\right)^2 > \frac{\overline{A}}{(\underline{A}+K)}\sqrt{\frac{F}{M}}+\frac{2\left(\frac{\underline{A}+K}{\overline{A}}\right)^2}{\frac{F}{M}} \tag{90}$$

Let $y=\frac{\underline{A}+K}{\overline{A}}$ and $x=\frac{F}{M}$. Then conditions (90) and 82) become

$$3\left(\frac{1+y}{1+\sqrt{x}}\right)^2 > \frac{\sqrt{x}}{y}+\frac{2\,(y)^2}{x} \tag{91}$$

$$y > \sqrt{x}>0\text{for}\underline{A}=0 \tag{92}$$

Condition (91) can also be written as,

$$\frac{y^2x+6yx+3x-2y^2-4y^2\sqrt{x}}{x\,(1+2\sqrt{x}+x)} > \frac{\sqrt{x}}{y} \tag{93}$$

Let $y=a\sqrt{x}$. Then (93) becomes

$$\left(a^3-1\right)x+\left(6a^2-4a^3-2\right)\sqrt{x}+3a-2a^3-1>0$$

Since $3a-2a^3-1<0$ for $a>1$ and $a^3-1>0$ the inequality can only be true for sufficiently high x. The remainder of the proof consists in showing that these high values of x are inconsistent with condition (86). We prove this for the special case where $\underline{A}=0$ where condition (86) reduces to

$$\frac{3}{2}>\left(1+\sqrt{x}\right)^2 \tag{94}$$

and therefore $x^{\max}=\left(\sqrt{\frac{3}{2}}-1\right)^2$. Hence

$$\begin{aligned} f_a\left(x^{\max}\right) &= \left(a^3-1\right)\left(\sqrt{\frac{3}{2}}-1\right)^2+\left(6a^2-4a^3-2\right)\left(\sqrt{\frac{3}{2}}-1\right)+3a-2a^3-1 \\ &= \left(\frac{9}{2}-6\sqrt{\frac{3}{2}}\right)a^3+6\left(\sqrt{\frac{3}{2}}-1\right)a^2+3a-2\left(\sqrt{\frac{3}{2}}-1\right)-\left(\sqrt{\frac{3}{2}}-1\right)^2-1 \end{aligned}$$

$$f_{a=1}\left(x^{\max}\right)=\frac{9}{2}-6+5-1-\frac{5}{2}=2-6+5-1=0$$

$$\frac{\partial f_a\left(x^{\max}\right)}{\partial a}=3\left(\frac{9}{2}-6\sqrt{\frac{3}{2}}\right)a^2+12\left(\sqrt{\frac{3}{2}}-1\right)a+3$$

$$\left.\frac{\partial f_a\left(x^{\max}\right)}{\partial a}\right|_{a=1}=-6\sqrt{\frac{3}{2}}+\frac{9}{2}<0$$

Since $\partial f_a\left(x^{\max}\right)/\partial a$ is a parabola $\left.\partial f_a\left(x^{\max}\right)/\partial a\right|_{a=1}<0$ implies that it is decreasing for all $a>1$. Thus, $\partial f_a\left(x^{\max}\right)/\partial a<0$ for all $a<1$. This implies that $f_{a=1}\left(x^{\max}\right)<0$ for all $a>1$. But this is a contradiction with the condition (94) and the result follows. □

Proposition 10 shows that when $\underline{A}^l=0$ parents never invest in enhancing the abilities of their children in both networks. This result holds in general in the current setup. Notice that for $\underline{A}^l>0$ we can express $\overline{A^l}=t^l\underline{A}^l$ where $t^l>0$. The required inequalities for an interior solution for F-parents in this case are

$$\begin{aligned}
3\frac{F}{M}\left(\frac{1+t+\frac{K}{\underline{A}}}{1+\sqrt{\frac{F}{M}}}\right)^2 &> \frac{F}{M}\frac{t^3}{\left(1+\frac{K}{\underline{A}}\right)}\sqrt{\frac{F}{M}}+2\left(1+\frac{K}{\underline{A}}\right)^2 && (95)\\
3\frac{F}{M}\left(\frac{1+t+\frac{K}{\underline{A}}}{1+\sqrt{\frac{F}{M}}}\right)^2 &> \frac{1}{\left(t+\frac{K}{\underline{A}}\right)\sqrt{\frac{F}{M}}}+2\frac{F}{M}\left(t+\frac{K}{\underline{A}}\right)^2 && (96)\\
\frac{K}{\underline{A}}+1 &> t\sqrt{\frac{F}{M}} && (97)\\
\frac{K}{\underline{A}}+t &> \frac{1}{\sqrt{\frac{F}{M}}} && (98)
\end{aligned}$$

We did not prove the non-existence of an interior solution analytically, but rather by plotting these inequalities in a three-dimensional plot with axis $F/M, K/\underline{A}$ and t in Mathematica which gives an empty intersection. Hence, in the current setup no interior solution seems possible.

C.3 Proof of Proposition 7

Let the proportion of parents in network M be m and of parents in network F be $(1-m)$. We prove the proposition for $B^M \geq CB^F$.

(i) When everybody invests in the same network, let's say network F,[38] then $e^F_{p_iF} = \overline{A} + K$ while $e^F_{p_iM} = \underline{A} + K$ and $e^M_{p_iM} = \overline{A}$ and $e^M_{p_iF} = \underline{A}$. This gives different distributions from which children's talents are drawn for the different parental traits unless $\underline{A} = \overline{A} = A$. We need to calculate $\overline{b^{M^2}}$ and $\overline{b^{F^2}}$ given that the children coming from different networks face different uniform distributions recalling that

$$\overline{b^{M^2}} = E\left(b_i^{M^2} \left| b_i^M > Cb_i^F \right.\right), \overline{b^{F^2}} = E\left(b_i^{M^2} \left| b_i^M < Cb_i^F \right.\right)$$

For all F children in the entire society we have

$$\overline{b^{F^2}} = m \left.\overline{b^{F^2}}\right|_M + (1-m) \left.\overline{b^{F^2}}\right|_F$$

where $\left.\overline{b^{F^2}}\right|_M$ refers to the F children coming from M parents and $\left.\overline{b^{F^2}}\right|_F$ refers to the F children coming from F parents. Using (41) we can calculate

$$\left.\overline{b^{F^2}}\right|_M = \frac{B^{F^2}}{2} = \frac{\left(\underline{A}+K\right)^2}{2} \text{ and } \left.\overline{b^{F^2}}\right|_F = \frac{B^{F^2}}{2} = \frac{\left(\overline{A}+K\right)^2}{2}$$

Therefore for all F children in the whole society we get

$$\overline{b^{F^2}} = m\frac{\left(\underline{A}+K\right)^2}{2} + (1-m)\frac{\left(\overline{A}+K\right)^2}{2} \tag{99}$$

Similarly, for all the M children in society

$$\overline{b^{F^2}} = m \left.\overline{b^{M^2}}\right|_M + (1-m) \left.\overline{b^{M^2}}\right|_F$$

which after using (43) becomes

[38] The proof for everybody investing in network M is analogous.

$$\overline{b^{M^2}} = \frac{1}{6}\left(m\frac{4\overline{A}^3 - C^3\left(\underline{A}+K\right)^3}{2\overline{A} - C\left(\underline{A}+K\right)} + (1-m)\frac{4\underline{A}^3 - C^3\left(\overline{A}+K\right)^3}{2\underline{A} - C\left(\overline{A}+K\right)}\right) \quad (100)$$

Introducing these expression into the defining equation of C_P given by (14) and rearranging, we get

$$G(C) = \frac{\left(4 - a^{F^2}\left(m\frac{(\underline{A}+K)^2}{2} + (1-m)\frac{\left(\overline{A}+K\right)^2}{2}\right)^2\right)^2}{\left(4 + a^{F^2}\left(m\frac{(\underline{A}+K)^2}{2} + (1-m)\frac{\left(\overline{A}+K\right)^2}{2}\right)^2\right)} C_P^2$$
$$- \frac{\left(4 - a^{M^2}\left(\frac{1}{6}\left(m\frac{4\overline{A}^3 - C^3(\underline{A}+K)^3}{2\overline{A} - C(\underline{A}+K)} + (1-m)\frac{4\underline{A}^3 - C^3\left(\overline{A}+K\right)^3}{2\underline{A} - C\left(\overline{A}+K\right)}\right)\right)^2\right)^2}{\left(4 + a^{M^2}\left(\frac{1}{6}\left(m\frac{4\overline{A}^3 - C^3(\underline{A}+K)^3}{2\overline{A} - C(\underline{A}+K)} + (1-m)\frac{4\underline{A}^3 - C^3\left(\overline{A}+K\right)^3}{2\underline{A} - C\left(\overline{A}+K\right)}\right)\right)^2\right)}$$

Since $G(0) < 0$ and $\lim_{C\to\infty} G(C) \to \infty$ then $G(.)$ has a fixed point.

(ii) When everybody invests in their own network, then for F parents $e^F_{p_iF} = \overline{A} + K$ and $e^M_{p_iF} = \underline{A}$ while for M parents $e^F_{p_iM} = \underline{A}$ and $e^M_{p_iM} = \overline{A} + K$. Therefore

$$\overline{b^{F^2}} = m\frac{\underline{A}^2}{2} + (1-m)\frac{\left(\overline{A}+K\right)^2}{2}$$

while

$$\overline{b^{M^2}} = \frac{1}{6}\left(m\frac{4\left(\overline{A}+K\right)^3 - C^3\left(\underline{A}\right)^3}{2\left(\overline{A}+K\right) - C\underline{A}} + (1-m)\frac{4\underline{A}^3 - C^3\left(\overline{A}+K\right)^3}{2\underline{A} - C\left(\overline{A}+K\right)}\right)$$

Introducing these expression into the defining equation of C_P given by (14) we get

$$G(C) = \frac{\left(4 - a^{F^2}\left(m\frac{\underline{A}^2}{2} + (1-m)\frac{\left(\overline{A}+K\right)^2}{2}\right)^2\right)^2}{\left(4 + a^{F^2}\left(m\frac{\underline{A}^2}{2} + (1-m)\frac{\left(\overline{A}+K\right)^2}{2}\right)^2\right)} C_P^2$$

$$- \frac{\left(4 - a^{M^2}\left(\frac{1}{6}\left(m\frac{4\left(\overline{A}+K\right)^3 - C^3(\underline{A})^3}{2\left(\overline{A}+K\right) - C\underline{A}} + (1-m)\frac{4\underline{A}^3 - C^3\left(\overline{A}+K\right)^3}{2\underline{A} - C\left(\overline{A}+K\right)}\right)\right)^2\right)^2}{\left(4 + a^{M^2}\left(\frac{1}{6}\left(m\frac{4\left(\overline{A}+K\right)^3 - C^3(\underline{A})^3}{2\left(\overline{A}+K\right) - C\underline{A}} + (1-m)\frac{4\underline{A}^3 - C^3\left(\overline{A}+K\right)^3}{2\underline{A} - C\left(\overline{A}+K\right)}\right)\right)^2\right)}$$

Since $G(0) < 0$ and $\lim_{C\to\infty} G(C) \to \infty$ then $G(.)$ has a fixed point.

C.4 Proof of Proposition 8

Since an interior solution is impossible we have to check which of the corner solution gives a higher utility which is done by comparing $g(0)$ with $g(K)$. It is better to invest in the F network only when $g(0) < g(K)$. Defining

$$x = \frac{F}{M} \text{ and } y = \frac{K}{A}$$

for both types of parents this is equivalent to

$$(1-x)\left[2\sqrt{x}(1+y)^2 + \frac{1+x}{1+y}\right] > 0$$

Since the expression in the square bracket is always positive, parents want to invest in F whenever $1 < x = \frac{F}{M}$ hence when $M < F$ and in M otherwise.