



Which Club Should I Attend, Dad?: Targeted Socialization and Production

**Facundo Albornoz
Antonio Cabrales
Esther Hauk**

November 2014

Barcelona GSE Working Paper Series

Working Paper n° 797

Which club should I attend, Dad?: Targeted socialization and production*

Facundo Albornoz[†] Antonio Cabrales[‡] Esther Hauk[§]

November 4, 2014

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. We first show that individuals underinvest in productive and social effort, but that solving only the investment problem can exacerbate the misallocations due to network choice, to the point that it may generate an even lower social welfare if one of the networks is sufficiently disadvantaged. We also study the interaction of parental investment with network choice. We relate these equilibrium results with characteristics that we find in the data on economic co-authorship and field transmission between advisors and advisees.

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

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

1 Introduction

Many productive processes are mediated by social interaction. The accumulation of human capital (Moretti, 2004), innovation activities (Cassiman and

*Hauk acknowledges financial support from the Spanish Ministry of Economy and Competitiveness through the Severa Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075) and through CICYT project number ECO2012-37065 and from the government of Catalonia. Pablo Warnes and Ezequiel García Lembergman provided outstanding assistance.

[†]Universidad de San Andrés and CONICET

[‡]Department of Economics, University College London; email: a.cabrales@ucl.ac.uk

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

Veugelers, 2002), and crime (Glaeser, Sacerdote, and Scheinkman, 2003), are all affected by the actions and abilities of others around us. As a result, economic agents devote considerable effort to developing social interactions with productive consequences, and to interacting with the “right” individuals. Until now, the literature has explored these two efforts separately. Benabou (1993)’s initial study, which has been expanded upon by recent economic literature on social networks,¹ sought to understand the process of selecting the best neighborhood to profit from spillovers. Cabrales, Calvó-Armengol, and Zenou (2011) examined the interaction between (undirected) socialization and production efforts, but did so within the confines of a single network.

This paper examines the impact of both the productive consequences of social interactions and the interaction with “right” individuals simultaneously. It analyzes how investment in (and choice of) network/neighborhood interact with decisions regarding socialization and production. Our approach allows us to generate novel results with implications for policy interventions. We model the interaction between directed network selection, indirect socialization within the network, and the choice of productive effort in a tractable framework where agents make decisions on all of those aspects optimally. This allows for a complete equilibrium and welfare analysis of individual decisions from which we derive comparative statics results. There are a variety of potential applications of our model, however, we focus on the economic production and human capital acquisition in environments where individuals have diverse backgrounds and abilities.

This paper begins by providing some empirical observations concerning a particular domain of production. Academic life provides a good example of an environment in which participants make interrelated productive and socialization decisions. Scientists develop social interactions with productive impact within different academic fields. At some point (often during their PhD studies), academics choose their networks or fields of interaction. This decision is based on idiosyncratic characteristics but it may also be influenced by the relative opportunities available in alternative networks. As these opportunities change over time (some networks become more productive or more rewarding than others) advisees may end up working in different fields than their advisors. Using data from RePEc we establish the following observations. First, more productive fields are characterized by higher levels of coauthorship. Second, intergenerational field mobility is less likely to occur in larger fields. Third, there is a negative relationship between intergenerational field mobility and the demand for new PhD Assistant Professors by field. And last, those fields with

¹See e.g. books by Goyal (2012), Jackson (2010) or Vega-Redondo (2007).

higher demand for Assistant Professors are more likely to host those advisees who have left the fields of their advisors. These observations have value in their own right, but our model also looks to shed light on the reasons for these outcomes.

Our model has three principal components. First, it recognizes that there are complementarities in productive investments (direct human capital acquisition) within networks. We follow Cabrales, Calvó-Armengol, and Zenou (2011) in assuming multiplicative spillovers between an agent's effort and those of other members of her/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 with other network members' knowledge stock. An individual's return on productive effort is idiosyncratic and can vary across networks. Second, the amount of spillover also depends positively on the costly socialization effort of the individual and the other network members. That is, taking advantage of others' information requires socialization and relies on other network members' socializing as well as one's own. One's incentives to socialize increases with the information everyone has, so socialization and productive investments are complementary. Finally, the agents also decide in which network they would like to interact, which depends on the relative abilities to gain surplus in each of them. 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 are 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, meetings are random, but the club or bar determines the people into whom he/she 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, so that we can use standard Nash equilibrium analysis. However, the fact that agents also choose their networks enhances our ability to evaluate realistic implications and connect these to our empirical observations.

Let us now turn to our results. We first fully characterize the unique

symmetric (in the sense that agents of same type choose the same options) equilibrium, both in terms of socialization and production effort as well as for network choice. From the equilibrium characterization, we observe a corollary result: the average socialization is increasing in the average type of the group. This accounts for our first empirical observation.

We then compare the equilibrium outcomes with those that would be chosen by a utilitarian social planner. As would be 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. There are more people than the socially desirable number in the network whose distribution of types has a larger mean if individual productivities are uniformly distributed (Proposition 3). The reason for this result is that the average welfare in the network with higher average productivity is unchanged if an additional person moves in, whereas the average welfare in the alternative network increases due to this change, since the average type improves. This is noteworthy because it is the more productive network that is overpopulated with decentralized sorting and a uniform distribution of individual productivities.² This can be interpreted, contrary to popular wisdom, as implying that there was too much integration into the mainstream labor market.

Moreover, there are parameter values such that a government that operates on one margin only, inducing efficient socialization and production effort within a network, may harm global efficiency by exacerbating misallocation due to network choice. This occurs to the extent that policy intervention may even reduce social welfare if the alternative network is sufficiently disadvantaged. This is an important novel point of our paper. The fact that individuals choose not just their efforts within networks, but also the networks to which they belong (in a sense intensive and extensive margins of socialization interact) make policy design more challenging, as there needs to be a coordination between the local within-network choice and the overall process of network selection.

The paper then explores an expanded model where parents invest in the training of children before they go into the labor market. The first result is again a full characterization of the equilibrium of the parental investment game. We establish existence and uniqueness and show that in the absence of asymmetries in the initial distribution of abilities across networks, the parents will put all of their effort in the a priori most productive network (Proposition

²With alternative distributions of talents this outcome holds only for some parameter values (see Appendix A.3) but not over the full range of parameters.

6). A tendency to choose “conformism” with the more productive network can also arise under asymmetries in the initial distribution of abilities across networks. If the initial endowment of ability is higher in the family network than in the alternative one, parents will only invest in the network that is more productive a priori (i.e. without parental investment) if they have enough resources to invest, or if differences in average productivities are very high. Our empirical observations are consistent with conformism with the most productive network. We find that intergenerational field mobility is less likely to occur in larger fields, that there is a negative relationship between intergenerational field mobility and the demand for new Assistant Professors by field, and that switchers are more likely to end up in fields with higher demand. According to our model, these observations arise in situations where advising skills are not field specific, advisors have access to sufficient investment resources, and/or the outcomes from taking the wrong field are sufficiently dire.

Finally, we 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.³

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.⁴ 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

³We 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.

⁴Bisin, Patacchini, Verdier, and Zenou (2010) evidences how cultural identity investments of minority groups increases under stronger group pressure.

sense, our paper is related to a growing literature on friendship formation. Currarini, Jackson, and Pin (2009) study individual preferences in friend choice. Their model can explain various empirical patterns of homophily, some of which are reminiscent of our empirical observations. For example, while they document that larger groups tend to exhibit more homophily, we find that advisees continue working in the same fields as their advisors when those fields are larger and more profitable. They also find that individuals in larger groups tend to socialize more (they have more ties). In our case, we document that the number of coauthors is larger in more productive fields. 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 with 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.

Our empirical observations about the intergenerational mobility between fields of academic economists is largely consistent with the empirical findings in the literature on intergenerational occupational mobility.⁵ 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

⁵See 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.

the same occupational category as her/his parent is increasing in the size of this occupational category; this is equivalent to our empirical observation 2 which states that intergenerational mobility is less likely to occur in larger fields.⁶

Our model finds that strong intergenerational job persistence is more likely when the initial endowment of ability is higher in the family network than in the alternative network. Moreover, intergenerational job persistence should be higher in the more profitable jobs (network) even if initial endowments of abilities are similar. There is a literature on the intergenerational transmission of employers (see e.g. Stinson and Wignall (2014)), which finds that sharing probabilities between father and sons in the U.S. in 2010 are much higher than the baseline probability that a father shares a job with an unrelated male similar to her/his son.⁷ 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.

This paper is organized as follows. Section 2 describes our illustrative empirical exercise. Section 3 describes the model. Section 4 contains the equilibrium and welfare analysis. Section 5 describes the effects of parental effort. Section 5 concludes. Most proofs are gathered in the Appendix.

2 Networking and Productivity and intergenerational transmission of research topics

Academic life is an example of a situation in which individual productive outcomes are affected by the abilities and activities of other people involved in the same production process. In this context, socialization decisions become key productive choices. Individuals do not only decide their socialization effort, but they also have to select with whom they will interact. To illustrate this phenomenon, we present a descriptive overview of trends of socialization and intergenerational mobility between research fields in Economics. Beyond the intrinsic value of learning about socialization patterns in the academic field of

⁶See Appendix C.

⁷The 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).

Economics, we believe this exercise will impose discipline on our analysis since we will require that these empirical observations emerge as equilibrium results within the confines of our theoretical framework.

2.1 Empirical observations

We extracted data from four main sources. First, we used the RePEc Genealogy project to construct a data set of advisors and advisees for all cohorts from 1980 to 2014. Second, we scrap from IDEAS-RePEc website every research paper by the authors listed in the RePEc Genealogy project. Third, we used JEL identifiers to associate an author with a field.⁸ Last, we construct measures of coauthorship scraping data from CollEc.⁹ Our final data set consists of 9063 researchers, 6421 advisor-advisee pairs and include all their papers, advisors, students and coauthors.

Networking and Productivity

If socialization carries productive implications, we expect stronger socialization in more productive networks. To establish this relationship empirically, we make a series of arbitrary definitions that facilitates the task. First, we associate networks to fields in economics. Second, we capture socialization by a measure of coauthorship. For each economist in our data set, we counted the number of her/his coauthors and then estimated the median number of coauthors by JEL fields,¹⁰ in order to obtain a measure of coauthorship by field.¹¹ Table 1 displays in column (1) the median number of coauthors for each field. This measure ranges from 3.5 in Economic Thought and Methodology to 9.5 in

⁸More specifically, we conducted the analysis thus: we summed up for each author all the JEL identifiers at the uppermost level (a single letter without numbers). For every individual author, we constructed a vector with the sum of all of the JEL information contained in her papers, divided by field. For example, if the author has three papers registered in IDEAS classified as A1, B2 and B31 according to the JEL, a second paper classified as B4 and B21, and the third getting C1 and A as classification, then we obtained the following vector of JEL fields: (2, 2, 1, 0, ..., 0), because she has 2 papers corresponding to A category, two papers in field B and another paper classified as C. After creating this vector for each author we define the author's field as the one in which she has the maximum value in this vector. In the example above, that author would have A and B as her main fields. We have dropped fields A, P and Z which represented less than 0.2 % of the total sample.

⁹A RePEc service of rankings by co-authorship centrality for authors registered in the RePEc Author Service.

¹⁰That is, the median number of coauthors amongst all the authors in each field.

¹¹The median instead of the mean was used in order to mitigate the concern that subfields are highly influenced by some authors who are clearly outliers in the less popular fields.

Table 1: Median number of coauthors by JEL field and Top 10% authors by field.

JEL field	Median Coauthors	“Top Author” Share
Economic Thought and Methodology (B)	3.5	0.14
Mathematical and Quantitative Methods (C)	5	0.25
Microeconomics (D)	5	0.27
Macroeconomics And Monetary Economics (E)	6	0.36
International Economics (F)	7	0.34
Finance (G)	6	0.42
Public Economics (H)	7	0.35
Health, Education, And Welfare (I)	5	0.21
Labor And Demographic (J)	7	0.36
Law And Economics (K)	4	0.31
Industrial Organization (L)	6	0.27
Business Administration and Economics (M)	6	0.18
Economic History (N)	6	0.32
Economic Development, Technological Change, and Growth (O)	8	0.36
Agricultural Economics (Q)	9.5	0.35
Urban, Rural, Regional (R)	8.5	0.39

Agricultural Economics. Last, we required a measure of productivity. IDEAS-RePEc generates a series of rankings by author, from which we selected the Average Rank Score.¹² For each author, we defined whether she or he is a “top author” according to whether she or he is included in the Top 10% of the IDEAS-RePEc of authors.¹³ Finally, we calculated the share of “top authors” for each JEL field as a measure of field productivity (Table 1, column 2).¹⁴

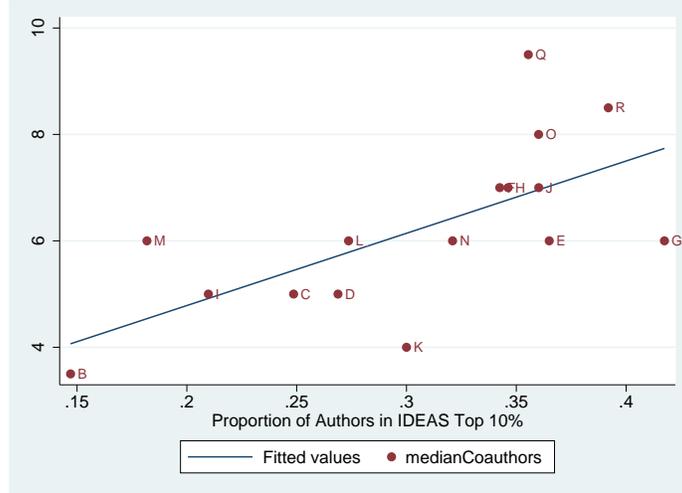
Based on Figure 1, we examine whether there is a relationship between the median number of coauthors and the share of “top authors”. Clearly, there is

¹²This score is determined by taking a harmonic mean of the ranks in each method, except the first one (number of works), the best, and the worst rank.

¹³We also used a more strict definition of “top author” using the Top 5% threshold.

¹⁴Note that our data is biased towards “top researchers”. In almost all the fields, more than 10% of researchers are considered a “top researcher”.

Figure 1: Association between the median number of coauthors and the proportion of authors in the Top 10% in IDEAS ranking.



a positive association that we summarize as:

Empirical Observation 1 *More productive fields are characterized by higher levels of coauthorship.*

Note that in interpreting coauthorship as the observable consequence of socialization, this empirical observation is related to the result in Currarini, Jackson, and Pin (2009) (section 3.2.) that shows that individuals belonging to larger groups have more friendship connections *per capita*.

Intergenerational transmission of research topics

To explore the patters of “intergenerational” field mobility, let s_i be the number of student-advisors pairs working in the same field and b_i be the number of student-advisor pairs in different fields, where i refers to the field of the advisor. Thus, the probability that a student chooses the field i conditional on her/his advisor working on i is equal to:

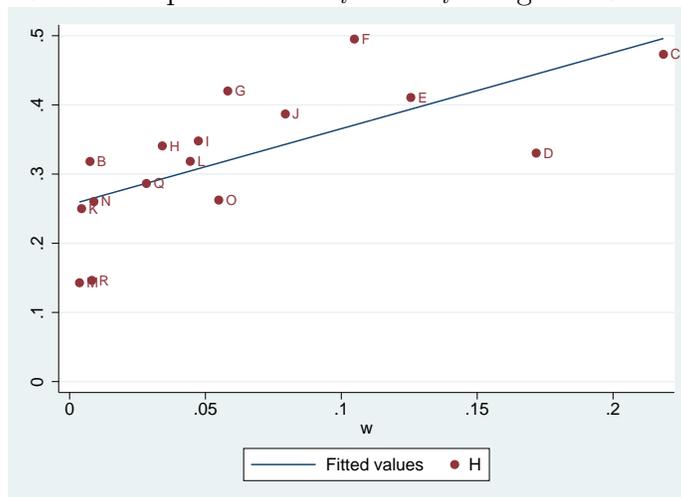
$$H_i = \frac{s_i}{s_i + b_i}$$

This measure is reminiscent of what Currarini, Jackson, and Pin (2009) refer to as relative homophily. Table 2 reports (H_i) together with the unconditional probability of a student of working on field i (w_i).

Table 2: Conditional probability (H) and unconditional probability (w) using new data on papers by JEL classification.

JEL Field Name	H	w
Economic Thought and Methodology (B)	0.32	0.01
Mathematical and Quantitative Methods (C)	0.47	0.22
Microeconomics (D)	0.33	0.17
Macroeconomics and Monetary (E)	0.41	0.13
International (F)	0.49	0.10
Finance (G)	0.42	0.06
Public Economics (H)	0.34	0.03
Health, Education and Welfare (I)	0.35	0.05
Labor and Demographic (J)	0.39	0.08
Law and Economics (K)	0.25	0.01
Industrial Organization (L)	0.32	0.01
Business Administration and Economics Marketing Accounting (M)	0.14	0.01
Economic History (N)	0.26	0.01
Economic Development, Technological Change, and Growth (O)	0.26	0.05
Agricultural Economics (Q)	0.29	0.03
Urban, Rural, Regional, Real Estate and Transportation (R)	0.15	0.01

Figure 2: Relationship between H_i and w_i using new JEL fields data.



Notice that $H_i > w_i$ for all fields, which is driven by the fact that students often select advisors working in the fields that they like. Even more interesting, the difference between w_i and H_i varies across fields. In fact, Figure 2 plots both measures and shows that H_i increases in w_i . This observation leads to:

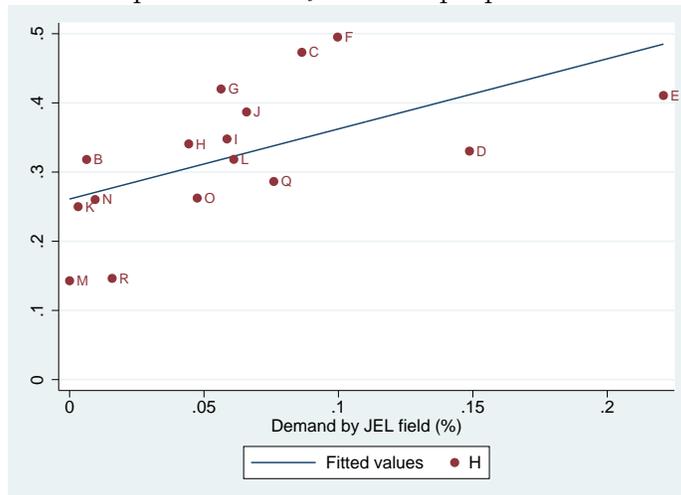
Empirical Observation 2 *Intergenerational field mobility is less likely to occur in larger fields.*

To infer whether intergenerational field mobility is related to the appeal of the field, we need a measure of profitability. As a proxy, we use a measure of success in the PhD labor market for each field based on the *Survey of the Labor Market for New PhD Hires in Economics* published by the Center for Business and Economic Research at the University of Arkansas. This survey collects information on over 200 organizations, including the number of new hires for each JEL field. Using this information, we construct an aggregate relative demand by field for the 2009-2012 period and regress this variable against the H_i index. Based on Figure 3, we obtain:

Empirical Observation 3 *There is a negative relationship between intergenerational field mobility and the demand for new PhD Assistant Professors by field.*

We examine whether those advisees switching fields end up joining fields with higher appeal (defined as having higher demand for new PhD Assistant

Figure 3: Relationship between H_i and the proportion of demand by field.



Professors). A way to uncover this feature is to build an index of “receptiveness” by field. That is, the share of switchers received by field. In Figure 4 we plot the relationship between “receptiveness” and the relative demand by field. This relationship is clearly positive implying:

Empirical Observation 4 *The fields with higher demand in the market for new PhD Assistant Professors are those that attract a higher proportion of researchers working in different fields than their advisors.*

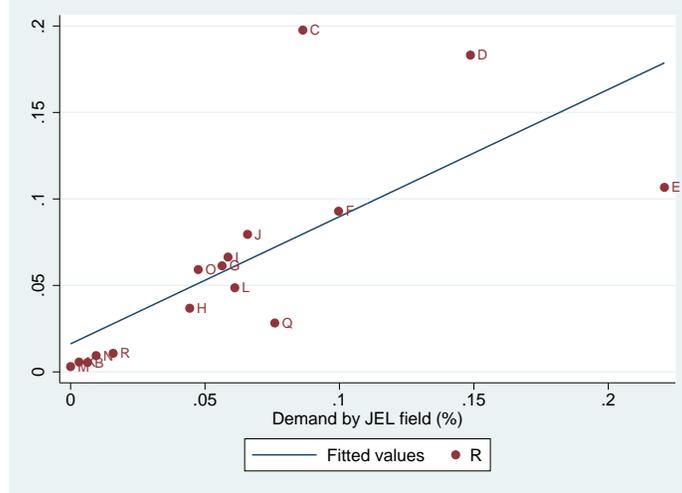
While observations 1, 2, 3 and 4 convey interesting facts about economics research, we take these results with caution as simply suggesting some characteristics of a multi-network framework, addressed in the next section.¹⁵

3 The model - payoffs

Our agents live in a world where they can receive utility from two different networks, M and F , in which they can participate. Each of those networks is

¹⁵Of course, these results simply suggest some trends in economics research. Although establishing these observations as robust empirical facts goes beyond the scope of this paper, we have undertaken some robustness checks that include different field classification as the NEP classification used by IDEAS-RePEc and measure of field productivity based on Card and Della Vigna (2013), which indicates the share of published papers in the top five Journals (American Economic Review (AER), Econometrica (EMA), the Journal of Political Economy (JPE), the Quarterly Journal of Economics (QJE), and the Review of Economic Studies (RES)) for each field.

Figure 4: Relationship between Receptiveness and the proportion of demand by field.



thus 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 . Agents' payoff arises from the activities of individuals' belonging to the networks. One such action is the direct productive effort (k_i^n), but there are also socialization activities (s_i^n), which serve the purpose of taking advantage of the productive efforts of other members of the network. Consequently, the payoff within a particular network n is the sum of two components, a private component, and one derived from the interactions.

The private component has a linear-quadratic cost-benefit structure. The benefit results from productive investment k_i^n as well as from a measure of individual productivity b_i^n summarizing any relevant idiosyncratic traits. As a result, we can write the private component P_i as:

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

The other component of payoffs is derived from interactions. Each player's payoff within a network depends on her/his own actions and the actions undertaken by all her/his network partners through a network externality. As in Cabrales, Calvó-Armengol, and Zenou (2011), we assume complementarity in productive efforts, in particular productive investment is multiplicative in the effort of pairs of individuals in the same network. Together with an assumption of symmetry, and overall constant returns, this implies that we multiply the square roots of productive efforts and add them up. These synergistic returns (payoffs generated by the network) also depend on a parameter a which captures the overall strength of synergies, as well as the individual

productivities of the members, b_i^n . These abilities are distributed uniformly and independently in $[0, B^n]$.

However, taking advantage of this network externality necessitates, and thus is mediated by, the socialization efforts of members of the network, s_i whose generation involves quadratic costs. We summarize the returns generated by the individual interactions with the social network by S_i^n , which letting \mathcal{N}_i be the network(s) to which individual i belongs, can be written as:

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

We must define the term, "incorporating socialization," $g_{ij}(\mathbf{s})$. One preliminary observation is that the socialization within each network is undirected. That is, individuals choose the group of people with whom to socialize (the network), but within the network they only choose the amount of interaction s_i , not the identity of the individuals with whom they interact. We believe this is the way in which socialization 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.

More formally, each player i selects a socialization effort, $s_i \geq 0$. Let s be a profile of socialization efforts. Then, i and j interact with a link intensity given by:

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

The functional form in (1) can be related to simple properties of the link intensity $g_{ij}^n(s_i^n, s_j^n)$, as per the following

Lemma 1 *Suppose that, for all $\mathbf{s} \neq \mathbf{0}$, the link intensity satisfies:*

(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) *overall constant returns to scale and symmetry: $\int_{j \in \mathcal{N}_i} g_{ij}^n(s_i^n, s_j^n) dj = \frac{1}{N^n} \int_{j \in \mathcal{N}_i} (s_i^n)^{1/2} (s_j^n)^{1/2} dj$;*

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

then, the link intensity is given by (1).

Proof of Lemma 1: Fix \mathbf{s} . Combining (A1) and (A3) gives $(s_k^n)^{1/2} g_{ij}^n(s_i^n, s_j^n) = s_j^{1/2} g_{ij}^n(s_i^n, s_k^n)$. Integrating across all j 's and using (A2) gives $g_{ij}^n(s_i^n, s_k^n) = \frac{1}{N^n} (s_i^n)^{1/2} (s_k^n)^{1/2}$. ■

Notice that given our definition of constant returns to scale, and given a level of socialization effort for all members of the network, total socialization $\int_{j \in \mathcal{N}_i} g_{ij}^n(s_i^n, s_j^n) dj$ 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 $1/(N^n)^\beta$ for some β different from 1.

Combining the private returns and the network externality yields individual payoffs described by:

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

Each individual i first chooses in which network to participate, possibly in both of them, a decision that carries no direct cost, and then takes the decisions over k_i and s_i simultaneously.

4 The equilibrium

We solve the game by backward induction both for the individual and the social planner who maximizes the sum of individual utilities and compare the solutions.

4.1 Choice of production and socialization efforts

For every network in which an individual participates we have to find her/his optimal productive and socialization effort (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

$$\int_{i \in \mathcal{N}^F \cup \mathcal{N}^F} u_i(b_i) di = \int_{i \in \mathcal{N}^F \cup \mathcal{N}^F} \left(b_i k_i + a b_i \sqrt{k_i s_i} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj - \frac{1}{2} k_i^2 - \frac{1}{2} s_i^2 \right) di \quad (3)$$

The FOC for the decentralized problem are

$$\begin{aligned} k_i &= b_i + \frac{a}{2} b_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} \\ s_i &= \frac{a}{2} b_i \sqrt{\frac{k_i}{s_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} \end{aligned}$$

while the FOC for the social planner simplify to

$$\begin{aligned} k_i^s &= b_i + a b_i \sqrt{\frac{s_i^s}{k_i^s}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j^s s_j^s}}{N^i} dj \\ s_i^s &= a b_i \sqrt{\frac{k_i^s}{s_i^s}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j^s s_j^s}}{N^i} dj \end{aligned}$$

We conjecture that both the individual optimum as well as the social optimum depend on its own productivity and is multiplicative in a parameter common to all individuals in the network that somehow captures the network effects and then shows that the conjecture is indeed possible. This implies for the individual problem $k_i = b_i k$ and $s_i = b_i s$ while for the centralized solution $k_i^s = b_i k^s$ and $s_i^s = b_i s^s$. Observe that the conjecture implies that $\int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} di = \int_{j \in \mathcal{N}_i} \frac{b_j^2 \sqrt{k s}}{N^i} di = \bar{b}^2 \sqrt{k s}$ where

$$\bar{b}^2 = \int_{j \in \mathcal{N}_i} \frac{b_j^2}{N^i} di.$$

With these conjectures we get two simultaneous equations with two unknowns, namely

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

for the decentralized problem and

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

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

Proposition 1 *The optimal common network parameter for productive and socialization effort are given by*

$$k = \frac{1}{1 - \frac{a^2\bar{b}^2}{4}} = \frac{4}{4 - a^2\bar{b}^2} \quad (4)$$

$$s = \frac{\frac{a}{2}\bar{b}^2}{1 - \frac{a^2\bar{b}^2}{4}} = \frac{2a^2\bar{b}^2}{4 - a^2\bar{b}^2} \quad (5)$$

for the individual choice problem and by

$$k^s = \frac{1}{1 - a^2\bar{b}^2} \quad (6)$$

$$s^s = \frac{ab^2}{1 - a^2\bar{b}^2} \quad (7)$$

for the social planner.

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

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

Corollary 1 is reflected in our empirical observation 1, which shows that scholars in more productive fields have more coauthors, if we identify coauthorship as the observable correlate of socialization. Thus, individuals within more productive networks socialize more on average.

In addition it is easy to see that

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

This happens because 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, namely,

$$u_i(b_i) = b_i^2 k + ab_i^2 k s \bar{b}^2 - \frac{b_i^2}{2} k^2 - \frac{b_i^2}{2} s^2$$

$$u_i(b_i) = 2b_i^2 \frac{(4 + a^2 \bar{b}^2)}{(4 - a^2 \bar{b}^2)^2} \quad (8)$$

is lower than the individual utility resulting from the social planner solution given by

$$u_i^s(b_i) = b_i^2 k^s + ab_i^2 k^s s^s \bar{b}^2 - \frac{b_i^2}{2} (k^s)^2 - \frac{b_i^2}{2} (s^s)^2$$

$$u_i^s(b_i) = \frac{b_i^2}{2} \left(\frac{1}{1 - a^2 \bar{b}^2} \right). \quad (9)$$

4.2 Choice of network

Assume now that there are two different networks M and F and individuals or the social planner must decide individual investment in both networks. Individual utility is given by

$$u_i(b_i^M, b_i^F) = b_i^M k_i^M + b_i^F k_i^F + a^M b_i^M \sqrt{k_i^M s_i^M} \int_{j \in \mathcal{N}_i} \frac{b_j^M \sqrt{k_j^M s_j^M}}{N^i} dj$$

$$+ a^F b_i^F \sqrt{k_i^F s_i^F} \int_{j \in \mathcal{N}_i} \frac{b_j^F \sqrt{k_j^F s_j^F}}{N^i} dj - \frac{1}{2} (k_i^F + k_i^M)^2 - \frac{1}{2} (s_i^F + s_i^M)^2$$

It is immediate that individuals will only choose to invest in one of the networks. Also the social planner will want individuals to invest in one of the networks only.

We will first analyze the individual choice of the network. 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 $b_i^M = C b_i^F$ such that individuals who fall below the line will choose network F while individuals above the line will choose network M . This implies that

$$\bar{b}^{M^2} = E \left(b_i^{M^2} \mid b_i^M > C b_i^F \right)$$

$$\bar{b}^{F^2} = E \left(b_i^{F^2} \mid b_i^M < C b_i^F \right)$$

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

$$\max \left[2b_i^{M^2} \frac{(4 + a^{M^2} \overline{b^{M^2}})}{(4 - a^{M^2} \overline{b^{M^2}})^2}, 2b_i^{F^2} \frac{(4 + a^{F^2} \overline{b^{F^2}})}{(4 - a^{F^2} \overline{b^{F^2}})^2} \right] \quad (10)$$

where the dividing line - if it exists - is defined when both terms of (10) are equal or equivalently when

$$b_i^M = b_i^F \sqrt{\frac{(4 + a^{F^2} \overline{b^{F^2}})}{(4 + a^{M^2} \overline{b^{M^2}})} \frac{(4 - a^{M^2} \overline{b^{M^2}})^2}{(4 - a^{F^2} \overline{b^{F^2}})^2}} = b_i^F C$$

Now the personally (equilibrium) optimal C_P is unique and given as the fixed point of

$$C_P = \sqrt{\frac{(4 + a^{F^2} \overline{b^{F^2}})}{(4 + a^{M^2} \overline{b^{M^2}})} \frac{(4 - a^{M^2} \overline{b^{M^2}})^2}{(4 - a^{F^2} \overline{b^{F^2}})^2}} \quad (11)$$

If we have induced the s^s and k^s (say via subsidies) that are appropriate at the social optimum, people would adjust themselves to networks using

$$\frac{b_i^{F^2}}{2(1 - a^{A^2} \overline{b^{F^2}})} = \frac{b_i^{M^2}}{2(1 - a^{M^2} \overline{b^{M^2}})}$$

In that case the (again, unique) dividing line is defined as $b_i^M = C_E b_i^F$ would solve

$$C_E = \sqrt{\frac{1 - a^{M^2} \overline{b^{M^2}}}{1 - a^{F^2} \overline{b^{F^2}}}} \quad (12)$$

In order to proceed we must make assumption on how individual productivities are distributed. For most part of this analysis, we will concentrate on the uniform distribution. Hence we assume that $b_i^l \sim U[0, B^l]$ in network l . Under this assumption we show in Appendix A.1 that

$$\overline{b^{F^2}} = \frac{B^{F^2}}{2} \quad (13)$$

and

$$\overline{b^{M^2}} = \frac{1}{6} \frac{4B^{M^3} - C^3 B^{F^3}}{2B^M - CB^F} \quad (14)$$

Moreover, we can prove:

Lemma 2 *Both C_P defined by (11) and C_E defined by (12) exist and are unique.*

Proof. See Appendix A.1. ■

We can check now whether the decentralized and centralized networks reach a social optimum. The following results shows that this is not the case.

Proposition 3 *If $B^M > C_E B^F$, social welfare is increasing in C at $C = C_E$ and at $C = C_P$*

Proof. See Appendix A.2. ■

Recall that the social planner in a centralized network chooses the optimal productive and socialization investments but has no influence on the process through which individuals self-select into one of the two networks. Proposition 3 implies that with a uniform distribution of individual talent, too few people join the F network, independent of whether there is a social planner or not.¹⁶ One can explain this result 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 social welfare. Also, average welfare in the F network remains almost unchanged if an additional person joins in because the average type $\overline{b^{F^2}} = B^{F^2}/2$ does not depend on C . Moreover, the average welfare in the M network does increase due to this change, since the average type improves ($\overline{b^{M^2}}$ does depend on C). 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.3 illustrates that social welfare might be increasing or decreasing in C at $C = C_E$ when individual productivities follow a Pareto distribution (with a Pareto distribution $\overline{b^{F^2}}$ depends on C).¹⁷

¹⁶The 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.

¹⁷We also give a sufficient condition for the same result to obtain with a Pareto distribution, when the distribution's tail is sufficiently thin.

Interestingly, when people freely sort themselves into networks, it is the more efficient network that becomes overpopulated. Consider e.g. 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 teaching and research activities. In the case of uniformly distributed individual productivities and assuming that productivity in research has a higher upper bound than teaching, the research community will be overpopulated and too few people will voluntarily classify themselves as teachers.

Inducing the optimal socialization and productive efforts within a network does not necessarily lead to a better overall outcome when individuals freely sort themselves into networks.

Claim 1 *There are parameter values for which $C_P > C_E$,*

In other words, there are cases for which local efficiency for a given network reduces global efficiency. The claim is proved by simulations in Appendix 1. From the simulations it is safe to conclude that C_P always outperforms C_E when the disadvantage of the F network is sufficiently large (low p and low l). Under C_E the government is operating on one margin inducing efficient productive and socialization effort within the network. However, when the F network is sufficiently disadvantaged the relative gains for the M network from local efficiency are larger than for the F network which increases the overall appeal of the M network when individuals make their own sorting decisions. In 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. Returning to our immigrant example, if the immigrant network has clearly a productivity disadvantage and a synergy disadvantage, no government intervention is socially better than local intervention via transfers and taxes that induce the efficient effort levels within the networks.¹⁸

¹⁸A 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.

5 Parental influence on the child’s private return to the network

In this section, we explore the role of parents in inducing their children to choose a particular network. By doing this, our model can generate a variety of interesting social cultural tendencies in a two-network society.

We let parents involve in improving/expanding the abilities of their children in one or both of the networks.¹⁹ Specifically, we assume that parental inculturation effort can change the support of the distribution of b^l s from which the child’s b_i^l is drawn. The realistic consequence of this assumption is that parents can influence the process of network selection without fully determining which network their children eventually select.

For simplicity, we further impose the distribution of b^l s to be uniform. More formally, let

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

where

$$e_{p_i}^l = \begin{cases} \overline{A^l} + x_{p_i}^l, & \text{if } i \text{ belongs to network } l \\ \underline{A^l} + x_{p_i}^l, & \text{if } i \text{ does not belong to network } l \end{cases}$$

and $x_{p_i}^l$ summarizes the effort devoted by parent p_i of agent (child) i .²⁰ Parameters $(\underline{A^l}, \overline{A^l})$ capture the initial endowment of ability in the parental and the non-parental networks, respectively. For example, consider a parent i belonging to network F. In this case, $\overline{A^F}$ summarizes how network F (e.g. a neighborhood, a religious community, or any particular cultural identity) influences the distribution of abilities from which i will draw a b_i . However, the other network, in this case M, may also favor the acquisition of its specific abilities through, for example, mainstream education. This is captured by $\underline{A^F}$. Both parameters $(\underline{A^l}, \overline{A^l})$ may differ in each network and their different combinations can describe different cultural contexts. For example, a society characterized by a dominant culture may be described as a case where $\underline{A^F} > \overline{A^F}$. Similarly, an alternative network may also be more pressing for a child born in a F-family than the mainstream if $\underline{A^F} < \overline{A^F}$.

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:

¹⁹With this assumption we depart from direct socialization efforts aiming at affecting children preferences, a possibility explored in Bisin and Verdier (2001).

²⁰This is the same as to assume that $b_i^l \sim U [0, 1]$ where parental effort simply enlarges the support of the distribution, increasing both mean and variance.

$$E \left[\max \left[2b_i^{F^2} \frac{(4 + a^{F^2} \overline{b^{F^2}})}{(4 - a^{F^2} \overline{b^{F^2}})^2}, 2b_i^{M^2} \frac{(4 + a^{M^2} \overline{b^{M^2}})}{(4 - a^{M^2} \overline{b^{M^2}})^2} \right] \right] \quad (15)$$

subject to the constraint that total effort cannot exceed an exogenous time endowment (K) such as $x_{pi}^F + x_{pi}^M = K$.²¹

Our parents are perfectly altruistic and do not want to impose any particular network *per se*. They only care about their children's welfare, irrespectively of whether they join their same own network or not. This is different from 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). As discussed below, altruistic parental involvement may generate intergenerational cultural mobility within the same family.

We proceed by finding the equilibrium levels of x_{pi}^F and x_{pi}^M . As the time allocated to enhance the abilities the mainstream requires (x_{pi}^M) may be expressed as $K - x_{pi}^F$, the exercise boils down to find the optimal x_{pi}^F . Notice that x_{pi}^F and x_{pi}^M are individual decisions and therefore parents take $\overline{b^{F^2}}$ and $\overline{b^{M^2}}$ as given. This allows us to ease the analysis by defining:

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

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

$$E \left[\max \left[2b_i^{F^2} F, 2b_i^{M^2} M \right] \right] \quad (16)$$

subject to

$$x_{pi}^F + x_{pi}^M = K \text{ and } 0 \leq x_{pi}^F \leq K \quad (17)$$

In Appendix B we show that equation (16) for F -parents can be expressed as:²²

²¹In principle, inasmuch K captures the time parents allocate to influence their children's decisions, K could vary across networks and parents. However, endogenizing K would complicate the analysis with no additional insights.

²²For M -parents the equation becomes

$$g(x_{p_i}^F) = \begin{cases} 2F \frac{(\bar{A} + x_{p_i}^F)^3}{6(\underline{A} + K - x_{p_i}^F)} \sqrt{\frac{F}{M}} + 2M \frac{(\underline{A} + K - x_{p_i}^F)^2}{3} & \text{if } x_{p_i}^F < \frac{K + (\underline{A} - \bar{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}} \\ 2M \frac{(\underline{A} + K - x_{p_i}^F)^3}{6(\bar{A} + x_{p_i}^F)} \sqrt{\frac{M}{F}} + 2F \frac{(\bar{A} + x_{p_i}^F)^2}{3} & \text{if } x_{p_i}^F > \frac{K + (\underline{A} - \bar{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}} \end{cases} \quad (19)$$

In the Appendix B.1 we prove that $g(x_{p_i}^F)$ is convex in both branches (Lemma 12). 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. In the appendix, we show that the existence of the two extreme points requires the following conditions to hold:

$$\frac{K + \underline{A}}{\bar{A}} > \sqrt{\frac{F}{M}} > \frac{\underline{A}}{K + \bar{A}} \quad (20)$$

and

$$\frac{K + \underline{A}}{\bar{A}} > \frac{1}{\sqrt{\frac{F}{M}}} > \frac{\underline{A}}{K + \bar{A}}, \quad (21)$$

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 B.2 states that this is not an equilibrium result.²³ Proposition 4 clarifies the parental investment subgame.

Proposition 4 *The decisions on cultural parental involvement are as follows:*

$$g(x_{p_i}^F) = \begin{cases} 2F \frac{(\underline{A} + x_{p_i}^F)^3}{6(\bar{A} + K - x_{p_i}^F)} \sqrt{\frac{F}{M}} + 2M \frac{(\bar{A} + K - x_{p_i}^F)^2}{3} & \text{if } x_{p_i}^F < \frac{K + (\bar{A} - \underline{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}} \\ 2M \frac{(\bar{A} + K - x_{p_i}^F)^3}{6(\underline{A} + x_{p_i}^F)} \sqrt{\frac{M}{F}} + 2F \frac{(\underline{A} + x_{p_i}^F)^2}{3} & \text{if } x_{p_i}^F > \frac{K + (\bar{A} - \underline{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}} \end{cases} \quad (18)$$

²³The 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.

1. *F*-parents will fully invest in their own network *F* if condition (20) is not satisfied and

$$\frac{\underline{A}^3}{(\bar{A} + K)} \sqrt{\frac{M}{F}} + 2 \frac{F}{M} (\bar{A} + K)^2 > \frac{(\underline{A} + K)^3}{\bar{A}} \sqrt{\frac{M}{F}} + 2 \frac{F}{M} \bar{A}^2 \quad (22)$$

or if condition(20) is satisfied and

$$\frac{\underline{A}^3}{(\bar{A} + K)} \sqrt{\frac{M}{F}} + 2 \frac{F}{M} (\bar{A} + K)^2 > \frac{F}{M} \frac{\bar{A}^3}{(\underline{A} + K)} \sqrt{\frac{F}{M}} + 2 (\underline{A} + K)^2 \quad (23)$$

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 (21) is not satisfied and

$$\frac{F}{M} \frac{\underline{A}^3}{(\bar{A} + K)} \sqrt{\frac{F}{M}} + 2 (\bar{A} + K)^2 > \frac{F}{M} \frac{(\underline{A} + K)^3}{\bar{A}} \sqrt{\frac{F}{M}} + 2 \bar{A}^2 \quad (24)$$

or if condition(21) is satisfied and

$$\frac{F}{M} \frac{\underline{A}^3}{(\bar{A} + K)} \sqrt{\frac{F}{M}} + 2 (\bar{A} + K)^2 > \frac{\bar{A}^3}{(\underline{A} + K)} \sqrt{\frac{M}{F}} + 2 \frac{F}{M} (\underline{A} + K)^2 \quad (25)$$

while they will fully invest in the other network *F* whenever condition (25) is violated.

Proof. Condition (23) says that investing fully in *F* is better than investing fully in *M* for *F*-parents, while condition (25) says that investing fully in *M* is better than investing fully in *F* for *M*-parents when both branches of (43) 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 (22) and (24). ■

Next, the existence of an equilibrium with endogenous parental influence must be formally proven. In the proof we must distinguish between several cases depending on which corner is chosen, (i) everybody invests in the same network, (ii) parents invest in their own network only.²⁴

To prove existence we need to show that in both cases in the current context the defining equation of C_P given by (11) has a fixed point. The

²⁴It is not possible to have both types of parents investing in the opposite network.

only difficulty consists in calculating $\overline{b^{M^2}}$ and $\overline{b^{F^2}}$ since we need to take into account that the children coming from different networks may face different uniform distributions recalling that

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

In Appendix B.3 we calculate the corresponding expressions taking into account the proportions of F/M children coming from F/M parents and establish

Proposition 5 *An equilibrium with endogenous parental influence exists and is unique.*

Proof. See Appendix B.3. ■

Discussion

In this section, we elaborate on the implications of parental cultural investment for social tendencies in a (2-network) society. We clarify how different cultural dynamics depend on the ratio of overall productivity, an exogenous time endowment, and the magnitude of the relative initial ability endowment in the parental and non-parental network.

Let's begin by considering the case where both networks are symmetric in terms of influence they exert in the inculturation process. To establish symmetry, we assume $\overline{A^l} = \underline{A^l} \forall l$. In this case, the only driving parameter for parental cultural investment is overall network efficiency (x) as we show in the next result:

Proposition 6 *If $\overline{A^l} = \underline{A^l}$ then*

$$\begin{aligned} x_{p_i}^F &= 0 \text{ and } x_{p_i}^M = K && \text{if } M > F \\ x_{p_i}^F &= K \text{ and } x_{p_i}^M = 0 && \text{if } M < F \end{aligned}$$

Proof. See Appendix B.4. ■

This result carries an interesting implication: in absence of asymmetries in network influence, there exists a general tendency toward conformism where parents spend all their inculturation effort on generating abilities for the most profitable network. Therefore, the reasons for a child joining the “alternative” network are purely idiosyncratic.²⁵

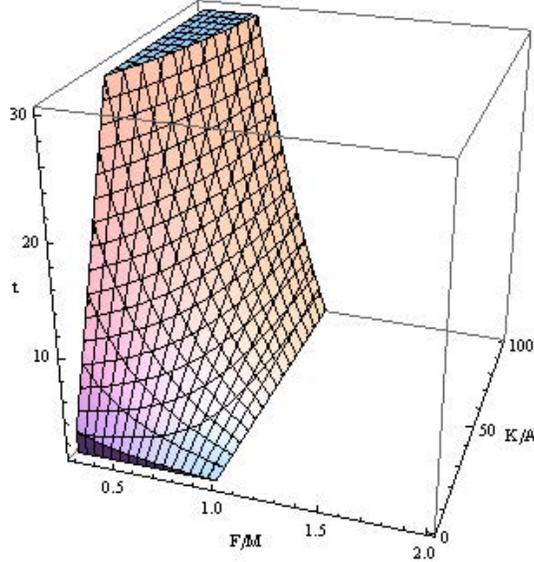
²⁵Notice that this situation where all parents invest in the mainstream network does not depend on K .

Proposition 6 can be related to our empirical observations 2 (intergenerational field mobility is less likely to occur in larger fields), 3 (there is a negative relationship between intergenerational field mobility and the demand for new Assistant Professors by field) and 4 (switchers are more likely to be attracted by fields with a higher demand). In other words, we find that even though PhD students select field by affinity, when selecting research topics, they are influenced by the current professional impact of each field. This is consistent with advising skills being sufficiently transferrable between fields. However, as will be shown, it is also consistent with $\bar{A}^l/\underline{A}^l > 1$ when K is sufficiently large and/or one network has a sufficient productivity advantage. In other words, conformity can arise from absence of parental asymmetry of ability, or because of the abundance of investment resources.

To see this, we allow for differences in network influence. Although we cannot obtain analytical solutions, the model is easy to solve numerically under the relevant parameters. 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, $t^l = \bar{A}^l/\underline{A}^l \in (0, \infty)$ 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 she/he receives more influence from the other network. To capture the case of parents' network dominance, it suffices to assume $\bar{A}^l/\underline{A}^l > 1 \forall l$. 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 presence of network influence asymmetry. Thus, if 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 5 where we depict F-parents' decision of investing on 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 (relatively low value of F/M). This is, as M-network overall productivity relative to F 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 5: Cultural conformism under asymmetric network influence: F-parents investing in the M-network



cultural conformism emerges in equilibrium even under asymmetric network influence.²⁶

However, if parents' time endowment (K) is not enough to revert the initial influence of the child's environment (t), then the society will be characterized by high intergenerational transmission within family, low mobility between networks and a tendency to fragmentation into two well defined networks. Although F-parents would like to educate their child to be competent for M-Network they don't have enough resources/time to do so. In these societies, intergenerational differences are purely associated with idiosyncratic characteristics of children. Thus, children switching networks have to be natural born talented in the specific activities of the high-productivity network.

We now consider the existence of an influential established network (not necessary advantageous in terms of profits). Consider, for example, a society where the institutions are designed to promote a dominant network (M with no loss of generality). In many modern societies schools enhance the skills and abilities required by mainstream activities (the M-network in our model). To capture this, assume $\underline{A}^F > \bar{A}^F$ while $\underline{A}^M < \bar{A}^M$. If network M is not only

²⁶Note 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.

dominant but also more efficient, clearly, parents from both networks are willing to spend all their education effort to improve their child’s skills to succeed in M network. The case when the M-network is not the most profitable one (i.e.: $F > M$) is particularly interesting. In this case, parents in both networks might be trapped into investing in the preponderant network M if the time endowment is not large enough. While the former case is similar (but stronger) to the cases associated with conformism, the latter case explains societies where rational parents tend to invest in the less profitable but culturally dominant networks.

Figure 6 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 7 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 in network M-abilities for $F/M < 1$. Surprisingly this might happen even if $F/M \geq 1$. As Figure 6 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.: $t^F \rightarrow 0$). Similar results hold for M-parents (Figure 7) but in this case, the strong influence of the predominant network is captured by $t^M \rightarrow \infty$. Notice that if the M-network is deeply rooted in the society’s culture (this is, if $t^M \rightarrow \infty, t^F \rightarrow 0$), 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 low productivity cultural dynamics.

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. We have shown that, as expected in a model with complementarities, individuals underinvest in productive and social effort, but that solving only the investment problem can exacerbate the misallocations due to network choice, to the point that it may generate an even lower social welfare if the alternative network is sufficiently disadvantaged. This is an important novel conclusion of this paper. Individuals 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. We relate these equilibrium results with characteristics that we find

Figure 6: Cultural conformism under an hegemonic network: F-parents investing in the M-network

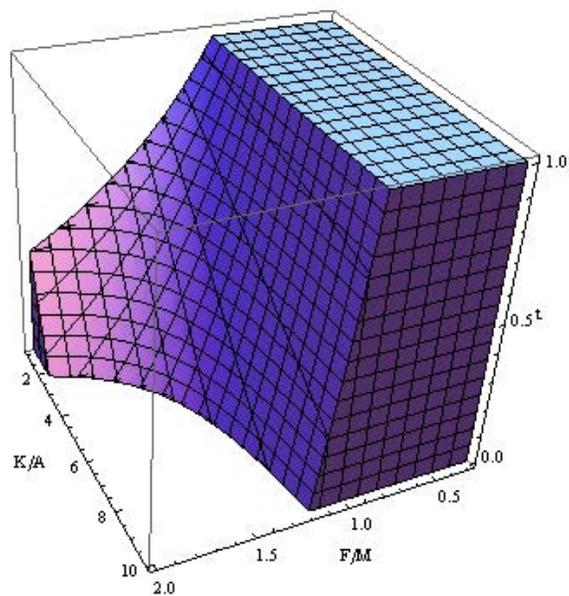
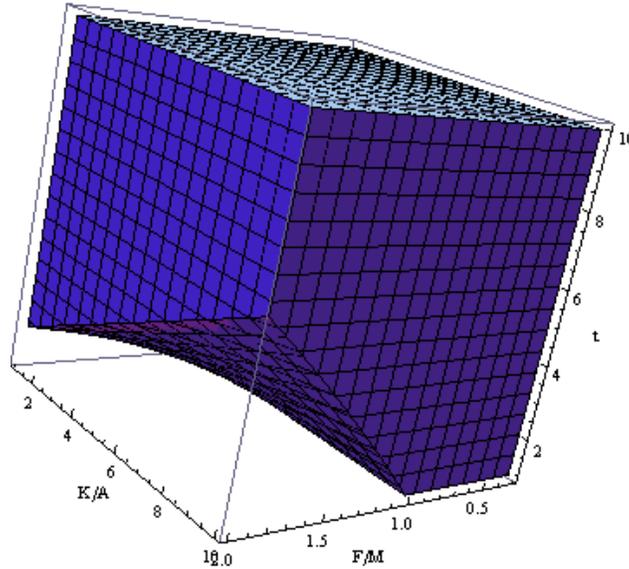


Figure 7: Cultural trap under with an hegemonic network: M-parents investing in the M-network



in the data of economic coauthorship and field transmission between advisors and advisees.

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.²⁷ 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

²⁷Another 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).

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.

References

- ALBORNOZ, F., S. BERLINSKI, AND A. CABRALES (2014): “Motivation, resources and the organization of the school system,” mimeo.
- 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,” .
- BISIN, A., E. PATACCINI, 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,” Discussion paper, National Bureau of Economic Research.
- 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.
- CABRALES, A., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2011): “Social interactions and spillovers,” *Games and Economic Behavior*, 72(2), 339–360.
- CARD, D., AND S. DELLA VIGNA (2013): “Nine Facts about Top Journals in Economics,” *Journal of Economic Literature*, 51(1), 144–161.

- 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 Equilibrium in the network game

A.1 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 \bar{b}^{F^2} and \bar{b}^{M^2} .

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

So

$$\bar{b}^{F^2} = \frac{B^{F^2}}{2} \tag{26}$$

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

So

$$\bar{b}^{M^2} = \frac{1}{6} \left(B^{F^2} C^2 + 2B^F B^M C + 4B^{M^2} - \frac{4B^{M^3}}{2B^M - CB^F} \right) \tag{27}$$

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

Lemma 3 $\overline{b^{M^2}}$ is maximized at $C = \frac{B^M}{B^F}$ and obtains the value

$$\overline{b_{\max}^{M^2}} = \frac{B^{M^2}}{2} \quad (28)$$

Proof. Observe that

$$\frac{\partial \overline{b^{M^2}}}{\partial C} = \frac{1}{6} \left(2CB^{F^2} + 2B^F B^M - \frac{4B^{M^3} B^F}{(2B^M - CB^F)^2} \right) \quad (29)$$

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 (28). ■

Using the expressions derived for (26) and (27) we can calculate C_P and C_E . In the case of C_P

the expression (11) becomes

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

Rearranging we get

$$\frac{\left(4 - \left(a^F \frac{B^{F^2}}{2}\right)^2\right)^2}{\left(4 + \left(a^F \frac{B^{F^2}}{2}\right)^2\right)^2} C_P^2 = \frac{\left(4 - \left(a^M \frac{1}{6} \frac{4B^{M^3} - C_P^3 B^{F^3}}{2B^M - C_P B^F}\right)^2\right)^2}{\left(4 + a^{M^2} \left(\frac{1}{6} \frac{4B^{M^3} - C_P^3 B^{F^3}}{2B^M - C_P B^F}\right)^2\right)^2} \quad (30)$$

We define

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

and check how it changes we 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 $4 - a^{M^2} \overline{b^{M^2}} > 0$ and $\frac{\partial \overline{b^{M^2}}}{\partial C} > 0$ in the relevant range. Hence the LHS of (30) 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 (30) is smaller than the LHS. Since the value of $\overline{b^{M^2}}(C = \frac{B^M}{B^F}) = \overline{b_{\max}^{M^2}} = \frac{B^{M^2}}{2}$ by Lemma 3 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,

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

Therefore it holds that

$$\frac{\left(4 + \left(a^M \frac{B^{M^2}}{2}\right)^2\right)}{\left(4 - \left(a^M \frac{B^{M^2}}{2}\right)^2\right)^2} B^{M^2} > \frac{\left(4 + \left(a^F \frac{B^{F^2}}{2}\right)^2\right)}{\left(4 - \left(a^F \frac{B^{F^2}}{2}\right)^2\right)^2} B^{F^2} \quad (31)$$

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 (31) 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

$$\overline{b^{M^2}} = \frac{B^{M^2}}{2} \quad (32)$$

and

$$\overline{b^{F^2}} = \frac{1}{6} \frac{4B^{F^3} - C^3B^{M^3}}{2B^F - CB^M} \quad (33)$$

Similarly, we can use (26) and (27) to express C_E and rearranging we get

$$\left(1 - a^{F^2} \frac{B^{F^4}}{4}\right) C_E^2 = 1 - \frac{a^{M^2}}{36} \left(B^{F^2} C_E^2 + 2B^F B^M C_E + 4B^{M^2} - \frac{4B^{M^3}}{2B^M - C_E B^F}\right)^2 \quad (34)$$

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 (34) is increasing in C_E and the RHS is decreasing for $B^M > C_E B^F$ (as per 29) 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 (34) is higher than the RHS.

$$\frac{B^{M^2}}{1 - a^{M^2} \frac{B^{M^4}}{4}} > \frac{B^{F^2}}{1 - a^{F^2} \frac{B^{F^4}}{4}} \quad (35)$$

for the case where $a^M = a^F$ this holds iff $B^M > B^F$. If condition (35) were violated, we would have the opposite inequality and then an equilibrium would exist with $B^M \leq C B^F$.

A.2 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

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

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

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

$$\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} = \int_0^{B^F} \int_{C b_i^F}^{B^M} \frac{\partial \left(\frac{b_i^{M^2}}{2} \left(\frac{1}{1 - a^{M^2} \overline{b^{M^2}{}^2}} \right) \right)}{\partial C} db_i^M db_i^F > 0$$

and the last inequality is true by Lemma 3.

(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

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

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

$$\text{Now at } C_P = \sqrt{\frac{(4 + a^{F^2} \overline{b^{F^2}})}{(4 + a^{M^2} \overline{b^{M^2}})} \frac{(4 - a^{M^2} \overline{b^{M^2}})^2}{(4 - a^{F^2} \overline{b^{F^2}})^2}}$$

$$\frac{\partial w(C)}{\partial C} \Big|_{C=C_P} = \int_0^{B^F} \int_{Cb_i^F}^{B^M} \frac{\partial \left(2b_i^{M^2} \frac{(4 + a^{M^2} \overline{b^{M^2}})}{(4 - a^{M^2} \overline{b^{M^2}})^2} \right)}{\partial C} db_i^M db_i^F > 0$$

and the last inequality is true by Lemma 3

A.3 Pareto distribution of individual productivities

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

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

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 $C \geq 1$.²⁸

²⁸If $C < 1$, the same results hold with the names of the networks interchanged.

Lemma 4 *If $C \geq 1$, then*

$$\overline{b^{F^2}} = \frac{\alpha}{\alpha - 2} \frac{1}{\alpha - 1} \frac{(2\alpha - 2) C^\alpha - (\alpha - 2)}{2C^\alpha - 1} \quad (36)$$

and

$$\overline{b^{M^2}} = \frac{\alpha}{\alpha - 2} \frac{\alpha}{\alpha - 1} C^2 \quad (37)$$

Proof. Let $C \geq 1$. Then

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

Similarly,

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

■

Lemma 5 *The optimal choice C_E defined by (12) exists and is unique.*

Proof. Using the Lemma 4 C_E can be rewritten as:

$$C_E = \sqrt{\frac{1 - a^{M^2} \left(\frac{\alpha}{\alpha-2} \frac{\alpha}{\alpha-1} C_E^2\right)^2}{1 - a^{F^2} \left(\frac{\alpha}{\alpha-2} \frac{1}{\alpha-1} \left((\alpha-1) + \frac{1}{2C_E^{\alpha-1}}\right)\right)^2}} \quad (38)$$

Note that the LHS of (38) is increasing in C_E and the RHS is decreasing in C_E so that a unique equilibrium exists. ■

Moreover,

Lemma 6 $C_E > 1 \Leftrightarrow a^{M^2} < a^{F^2}$

Proof. Note also that if $a^{M^2} = a^{F^2}$ the solution of (38) 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 7 *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*

$$\frac{a^{M^2}}{a^{F^2}} > \frac{((2\alpha - 2)C^\alpha - (\alpha - 2))^2 C^\alpha}{\alpha^2 (2C^\alpha - 1)^3 C^2} \quad (39)$$

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

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

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

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

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

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

Therefore

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

and

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

By Lemma 6 since $C_E > 1 \iff a^{M^2} < a^{F^2}$, hence $\frac{a^{M^2}}{a^{F^2}} < 1$.

We will now show that

$$1 > \frac{((2\alpha-2)C^\alpha - (\alpha-2))^2 C^\alpha}{\alpha^2 (2C^\alpha-1)^3 C^2} = \frac{((\alpha-1)(2C^\alpha-1)+1)^2 C^\alpha}{\alpha^2 (2C^\alpha-1)^3 C^2} \quad (40)$$

Note that

$$((\alpha-1)(2C^\alpha-1)+1)^2 < \alpha^2 (2C^\alpha-1)^2$$

since that expression is equivalent to

$$\begin{aligned} (\alpha-1)(2C^\alpha-1)+1 &< \alpha(2C^\alpha-1) \\ &\iff 1 < 2C^\alpha-1 \iff 1 < C^\alpha \end{aligned}$$

thus

$$\frac{((\alpha-1)(2C^\alpha-1)+1)^2 C^\alpha}{\alpha^2 (2C^\alpha-1)^3 C^2} < \frac{C^\alpha}{(2C^\alpha-1)C^2} < \frac{1}{C} < 1 \quad (41)$$

where the last two inequalities hold since $C > 1$, noting that in that case $2C^\alpha-1 > C^\alpha$. Thus equation (41) establishes (40).

Lemmas 7 shows that parameter values exist so that $\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} < 0$. while Lemma 8 shows the existence of parameter values that $\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} > 0$.

Lemma 7 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

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

Proof. Since

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

we have that

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

thus

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

■

Lemma 8 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

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

Proof. For a bounded C_E

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

Hence

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

and thus

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

Now since

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

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

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

or

$$\frac{-\left(1 - a^{F^2}\right) + \sqrt{\left(1 - a^{F^2}\right)^2 + 4ra^{F^2}}}{2ra^{F^2}} > \frac{1}{2r} \quad (42)$$

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

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

Proposition 7 immediately follows from these Lemmas. ■

A.4 Proof of Claim 1

We will establish this claim with numerical simulations. For this purpose let $p = B^F/B^M$ and $l = a^{F^2}/a^{M^2}$ and assume that network F has a productivity disadvantages so that $p < 1$. If $l < 1$ network A also provides worse synergies than network M while for $l > 1$ synergies in network F are stronger. We will plot $C_P - C_E$ as a function of p and/or l when p takes values between 0.1 and 0.9 and l takes values between 0.1 and 2. In the figures we use $B^M = 5$ and $a^{M^2} = 0.0015$. The general shape of these figures does not depend on this parameter choice. Figure A.1 clearly illustrates that there are parameter values for which $C_P - C_E$ are positive and others for which this difference is negative. From A.1 we learn that $C_P - C_E$ is monotonically decreasing in l for all values of $p \in (0, 1)$. In other words C_P outperforms C_E for both low l and low p but C_E starts picking up when synergies favor network F over network M : sufficiently high synergies even offset the productivity advantage of network M .

B Parental investment

Lemma 9 *Equation (16) can be written as a function with at most two branches:*

$$E \left[\max \left[2b_i^{F^2} F, 2b_i^{M^2} M \right] \right] = \begin{cases} 2F \frac{e_{p_i}^{F^3}}{6e_{p_i}^M} \sqrt{\frac{F}{M}} + 2M \frac{e_{p_i}^{M^2}}{3} & \text{if } e_{p_i}^M > \sqrt{\frac{F}{M}} e_{p_i}^F \\ 2M \frac{e_{p_i}^{M^3}}{6e_{p_i}^F} \sqrt{\frac{M}{F}} + 2F \frac{e_{p_i}^{F^2}}{3} & \text{if } e_{p_i}^M < \sqrt{\frac{F}{M}} e_{p_i}^F \end{cases} \quad (43)$$

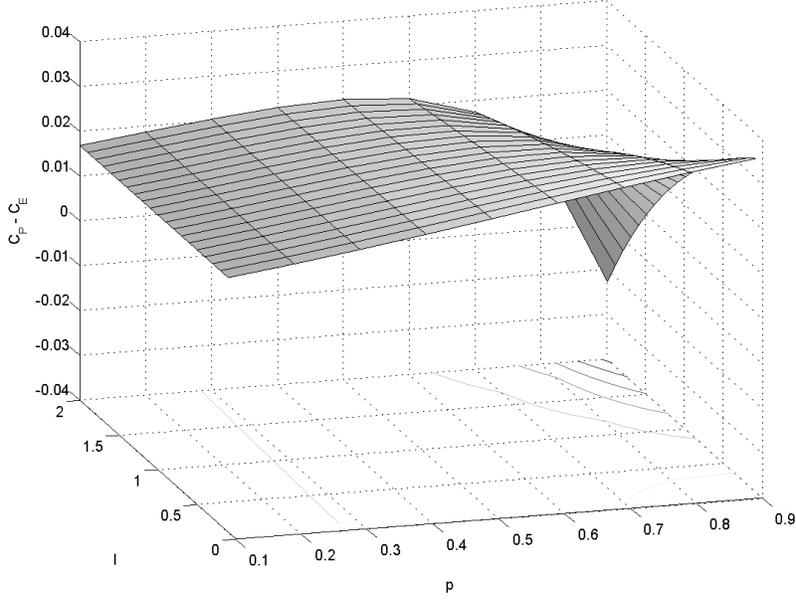


Figure A.1: $C_P - C_E$ as a function of l and p , for $B^M = 5$ and $a^{M^2} = 0.0015$.

where the two branches exist if the inequalities in (43) are non-empty.

Proof. Under the assumption that

$$e_{p_i}^M > \sqrt{\frac{F}{M}} e_{p_i}^F$$

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

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

so that

$$E \left[\max \left[2b_i^{F^2} F, 2b_i^{M^2} M \right] \right] = 2F \frac{e_{p_i}^{F^3}}{6e_{p_i}^M} \sqrt{\frac{F}{M}} + 2M \frac{e_{p_i}^{M^2}}{3} \quad (44)$$

Suppose instead that

$$e_{p_i}^M < \sqrt{\frac{F}{M}} e_{p_i}^F$$

then

$$E \left[\max \left[2b_i^{F^2} F, 2b_i^{M^2} M \right] \right] = 2M \frac{e_{p_i}^{M^3}}{6e_{p_i}^F} \sqrt{\frac{M}{F}} + 2F \frac{e_{p_i}^{F^2}}{3} \quad (45)$$

■

Thus, for a parent who belongs to network F , $x_{p_i}^F + x_{p_i}^M = K$, $e_{p_i}^F = \bar{A} + x_{p_i}^F$ and $e_{p_i}^M = \underline{A} + K - x_{p_i}^F$. It immediately follows:

Lemma 10

$$x_{p_i}^F < \frac{K + \left(\underline{A} - \bar{A} \sqrt{\frac{F}{M}} \right)}{1 + \sqrt{\frac{F}{M}}} \quad \text{if } e_{p_i}^M > \sqrt{\frac{F}{M}} e_{p_i}^F \quad (46)$$

$$x_{p_i}^F > \frac{K + \left(\underline{A} - \bar{A} \sqrt{\frac{F}{M}} \right)}{1 + \sqrt{\frac{F}{M}}} \quad \text{if } e_{p_i}^M < \sqrt{\frac{F}{M}} e_{p_i}^F \quad (47)$$

Analogously, for a parent who belongs to network M , $x_{p_i}^F + x_{p_i}^M = K$, $e_{p_i}^F = \underline{A} + x_{p_i}^F$ and $e_{p_i}^M = \bar{A} + K - x_{p_i}^F$. It immediately follows:

Lemma 11

$$x_{p_i}^F < \frac{K + \left(\bar{A} - \underline{A} \sqrt{\frac{F}{M}} \right)}{1 + \sqrt{\frac{F}{M}}} \quad \text{if } e_{p_i}^M > \sqrt{\frac{F}{M}} e_{p_i}^F \quad (48)$$

$$x_{p_i}^F > \frac{K + \left(\bar{A} - \underline{A} \sqrt{\frac{F}{M}} \right)}{1 + \sqrt{\frac{F}{M}}} \quad \text{if } e_{p_i}^M < \sqrt{\frac{F}{M}} e_{p_i}^F \quad (49)$$

B.1 Convexity of $g_1(x_{p_i}^F)$

Lemma 12 *The function $g_1(x_{p_i}^F)$ is convex.*

Proof. For $x_{p_i}^F < \frac{K + (\underline{A} - \bar{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}}$

$$\begin{aligned} \frac{\partial g_1(x_{p_i}^F)}{\partial x_{p_i}^F} &= -\frac{4}{3} M (\underline{A} + K - x_{p_i}^F) + 6F (\bar{A} + x_{p_i}^F)^2 \frac{\sqrt{\frac{F}{M}}}{6(\underline{A} + K - x_{p_i}^F)} \\ &\quad + 12F (\bar{A} + x_{p_i}^F)^3 \frac{\sqrt{\frac{F}{M}}}{6(\underline{A} + K - x_{p_i}^F)^2} \end{aligned}$$

Since

$$6F (\bar{A} + x_{p_i}^F)^2 \frac{\sqrt{\frac{F}{M}}}{6 (\underline{A} + K - x_{p_i}^F)} + 12F (\bar{A} + x_{p_i}^F)^3 \frac{\sqrt{\frac{F}{M}}}{6 (\underline{A} + K - x_{p_i}^F)^2}$$

is increasing in $x_{p_i}^F$ this implies that

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

so $g_1 (\cdot)$ is convex.

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

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

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

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

■

B.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 $x_{p_i}^F = \frac{K + (\underline{A} - \bar{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}}$ are optimal in the range $x_{p_i}^F < \frac{K + (\underline{A} - \bar{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}}$ and whether either $x_{p_i}^F = K$ or $x_{p_i}^F = \frac{K + (\underline{A} - \bar{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}}$ are optimal in the range $x_{p_i}^F > \frac{K + (\underline{A} - \bar{A} \sqrt{\frac{F}{M}})}{1 + \sqrt{\frac{F}{M}}}$. 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

$$\frac{K + \underline{A}}{\bar{A}} > \sqrt{\frac{F}{M}} > \frac{\underline{A}}{K + \bar{A}} \quad (50)$$

and that

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

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

Since,

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

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

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

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

$$F \left(\frac{\bar{A} + \underline{A} + K}{1 + \sqrt{\frac{F}{M}}}\right)^2 > 2F \frac{\bar{A}^3}{6(\underline{A} + K)} \sqrt{\frac{F}{M}} + 2M \frac{(\underline{A} + K)^2}{3} \quad (51)$$

$$F \left(\frac{\bar{A} + \underline{A} + K}{1 + \sqrt{\frac{F}{M}}}\right)^2 > 2M \frac{\underline{A}^3}{6(\bar{A} + K)} \sqrt{\frac{M}{F}} + 2F \frac{(\bar{A} + K)^2}{3} \quad (52)$$

Or equivalently

$$3 \frac{F}{M} \left(\frac{\bar{A} + \underline{A} + K}{1 + \sqrt{\frac{F}{M}}}\right)^2 > \frac{F}{M} \frac{\bar{A}^3}{(\underline{A} + K)} \sqrt{\frac{F}{M}} + 2(\underline{A} + K)^2 \quad (53)$$

$$3 \frac{F}{M} \left(\frac{\bar{A} + \underline{A} + K}{1 + \sqrt{\frac{F}{M}}}\right)^2 > \frac{\underline{A}^3}{(\bar{A} + K)} \sqrt{\frac{M}{F}} + 2 \frac{F}{M} (\bar{A} + K)^2 \quad (54)$$

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

$$\frac{K + \underline{A}}{\underline{A}} > \frac{1}{\sqrt{\frac{F}{M}}} > \frac{\underline{A}}{K + \underline{A}} \quad (55)$$

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

$$3 \frac{F}{M} \left(\frac{\bar{A} + \underline{A} + K}{1 + \sqrt{\frac{F}{M}}} \right)^2 > \frac{F}{M} \frac{\underline{A}^3}{(\bar{A} + K)} \sqrt{\frac{F}{M}} + 2(\bar{A} + K)^2 \quad (56)$$

$$3 \frac{F}{M} \left(\frac{\bar{A} + \underline{A} + K}{1 + \sqrt{\frac{F}{M}}} \right)^2 > \frac{\bar{A}^3}{(\underline{A} + K)} \sqrt{\frac{M}{F}} + 2 \frac{F}{M} (\underline{A} + K)^2 \quad (57)$$

Proposition 8 *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 (53) tells us when the interior solution for F-parents is better than no investment in F (full investment in M). Condition (54) 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 (50) is required. The proof consists in showing that these conditions are incompatible. We first rewrite condition (53) as

$$3 \left(\frac{1 + \frac{\underline{A} + K}{\underline{A}}}{1 + \sqrt{\frac{F}{M}}} \right)^2 > \frac{\bar{A}}{(\underline{A} + K)} \sqrt{\frac{F}{M}} + \frac{2 \left(\frac{\underline{A} + K}{\underline{A}} \right)^2}{\frac{F}{M}} \quad (58)$$

Let $y = \frac{\underline{A} + K}{\underline{A}}$ and $x = \frac{F}{M}$. Then conditions (58) and 50) become

$$3 \left(\frac{1 + y}{1 + \sqrt{x}} \right)^2 > \frac{\sqrt{x}}{y} + \frac{2(y)^2}{x} \quad (59)$$

$$y > \sqrt{x} > 0 \text{ for } \underline{A} = 0 \quad (60)$$

Condition (59) can also be written as,

$$\frac{y^2 x + 6yx + 3x - 2y^2 - 4y^2 \sqrt{x}}{x(1 + 2\sqrt{x} + x)} > \frac{\sqrt{x}}{y} \quad (61)$$

Let $y = a\sqrt{x}$. Then (61) becomes

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

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 (54). We prove this for the special case where $\underline{A} = 0$ where condition (54) reduces to

$$\frac{3}{2} > (1 + \sqrt{x})^2 \tag{62}$$

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

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

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

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

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

Since $\partial f_a(x^{\max})/\partial a$ is a parabola $\partial f_a(x^{\max})/\partial a|_{a=1} < 0$ implies that it is decreasing for all $a > 1$. Thus, $\partial f_a(x^{\max})/\partial a < 0$ for all $a < 1$. This implies that $f_{a=1}(x^{\max}) < 0$ for all $a > 1$. But this is a contradiction with the condition (62) and the result follows. ■

Proposition 8 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

$$3\frac{F}{M} \left(\frac{1+t+\frac{K}{\underline{A}}}{1+\sqrt{\frac{F}{M}}} \right)^2 > \frac{F}{M} \frac{t^3}{\left(1+\frac{K}{\underline{A}}\right)} \sqrt{\frac{F}{M}} + 2 \left(1+\frac{K}{\underline{A}}\right)^2 \quad (63)$$

$$3\frac{F}{M} \left(\frac{1+t+\frac{K}{\underline{A}}}{1+\sqrt{\frac{F}{M}}} \right)^2 > \frac{1}{\left(t+\frac{K}{\underline{A}}\right) \sqrt{\frac{F}{M}}} + 2\frac{F}{M} \left(t+\frac{K}{\underline{A}}\right)^2 \quad (64)$$

$$\frac{K}{\underline{A}} + 1 > t\sqrt{\frac{F}{M}} \quad (65)$$

$$\frac{K}{\underline{A}} + t > \frac{1}{\sqrt{\frac{F}{M}}} \quad (66)$$

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.

B.3 Proof of Proposition 5

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 ,²⁹ then $e_{p_i F}^F = \bar{A} + K$ while $e_{p_i M}^F = \underline{A} + K$ and $e_{p_i M}^M = \bar{A}$ and $e_{p_i F}^M = \underline{A}$. This gives different distributions from which children's talents are drawn for the different parental traits unless $\underline{A} = \bar{A} = A$. 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

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

For all F children in the entire society we have

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

where $\overline{b^{F^2}} \Big|_M$ refers to the F children coming from M parents and $\overline{b^{F^2}} \Big|_F$ refers to the F children coming from F parents. Using (13) we can calculate

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

²⁹The proof for everybody investing in network M is analogous.

Therefore for all F children in the whole society we get

$$\overline{b^{F^2}} = m \frac{(\underline{A} + K)^2}{2} + (1 - m) \frac{(\overline{A} + K)^2}{2} \quad (67)$$

Similarly, for all the M children in society

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

which after using (14) becomes

$$\overline{b^{M^2}} = \frac{1}{6} \left(m \frac{4\overline{A}^3 - C^3 (\underline{A} + K)^3}{2\overline{A} - C (\underline{A} + K)} + (1 - m) \frac{4\underline{A}^3 - C^3 (\overline{A} + K)^3}{2\underline{A} - C (\overline{A} + K)} \right) \quad (68)$$

Introducing these expression into the defining equation of C_P given by (11) and rearranging, we get

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

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

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

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

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

while

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

Introducing these expression into the defining equation of C_P given by (11) we get

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

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

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

B.4 Proof of Proposition 6

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.

C Occupational mobility

This appendix is based on Long and Ferrie (2013) who study occupational mobility in Britain and the U.S., Azam (2013) who looks at India and Binzel and Carvalho (2013) who examines Egypt. We restrict ourselves to these papers on intergenerational occupational mobility because they either calculate

or 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 . This allows us to relate their findings to our empirical observations on intergenerational field mobility of academic economists. 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.

Below we provide a summary table for the unconditional w_i and conditional probability H_i to belong to the different occupational categories. The results for Britain and the U.S. are calculations based 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 11357 father-son pairs.³⁰ 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.

As in the case of academic fields, the conditional probability to work in a certain occupational category is always higher than the unconditional probability, i.e. $H_i > w_i$. Also, the difference between w_i and H_i varies across occupational categories. Figure C.1 clearly shows that H_i increases in w_i and that this holds even when mixing different countries. Hence our empirical observation 2 also applies to intergenerational occupational mobility.

Finally, notice that intergenerational persistence among farmers varies across countries. In U.K and the U.S., where agriculture is relatively productive with respect to the cases of India and Egypt, the effect of H is stronger, which is consistence with Proposition 6 and Observations 2 and 3.

³⁰Azam (2013) does not find any differences in mobility in successive ten year birth cohorts.

Table C.1: Occupational mobility across job categories and countries

Country	Category	w_i	H_i
Britain	high white collar (HWC)	0,259	0,628
	low white collar (LWC)	0,123	0,209
	farmer	0,013	0,209
	skilled/semi skilled	0,542	0,622
	unskilled	0,062	0,09
U.S.	high white collar (HWC)	0,372	0,617
	low white collar (LWC)	0,111	0,161
	farmer	0,025	0,135
	skilled/semi skilled	0,398	0,466
	unskilled	0,093	0,133
India	white collar (WC)	0,121	0,444
	skilled/semi skilled	0,407	0,716
	unskilled	0,325	0,552
	farmer	0,147	0,275
Egypt	farmer	0,174	0,403
	semi skilled/unskilled manual	0,155	0,294
	skilled manual	0,235	0,43
	white collar (WC)	0,16	0,254
	professional	0,276	0,556

Figure C.1: Intergenerational occupational mobility: relationship between H_i and w_i

