

Occupational choice with endogenous spillovers*

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

June 8, 2017

Abstract

We study a model that integrates productive and socializing efforts with occupational choice in the presence of endogenous spillovers. Among other results, we show that more talented individuals work harder and contribute more to the emergence of externalities, but also have incentives to segregate. Average socializing increases in the average productivity of the occupation. Also, the size of an occupation grows in its network synergies. Turning to efficiency, we show that individuals underinvest in productive and socializing effort, and sort themselves inefficiently into occupations. We derive the optimal subsidy to achieve efficient effort within occupations and show that efficient sorting into occupations can always be achieved by a linear tax. We illustrate the importance for the government to intervene on both margins, as solving only the within occupation investment problem can exacerbate misallocations due to network choice and may even reduce welfare in presence of congestion costs.

JEL-Classification: D85, H21, H23, J24

Key-words: occupational choice, social interactions, endogenous spillovers, optimal taxation

*This paper is part of a substantial revision of the paper circulated under the title "Targeted socialization and production". It is based on the first part of the older paper only and introduces mainly new results. Hauk acknowledges financial support from the Spanish Ministry of Economy and Competitiveness through the Severo Ochoa Program for Centers of Excellence in R&D (SEV-2015-0563) and through CICYT project number ECO2015-67171-P (MINECO/FEDER) and from the government of Catalonia (2014 SGR 1064). We thank participants at numerous conferences and seminars for helpful comments.

[†]University of Nottingham and IIEP-CONICET; facundo.albornoz@nottingham.ac.uk

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

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

1 Introduction

Many productive processes are mediated by social interactions. The accumulation of human capital (Moretti, 2004), innovation (Cassiman and Veugelers, 2002), and crime (Glaeser, Sacerdote, and Scheinkman, 2003), are examples of activities carried out by individuals whose actions are affected by the activities and abilities of others they establish connections with. Since social interactions have productive consequences, economic agents naturally devote a considerable effort to developing them. From the perspective of an individual, socializing in production activities involves two different but interconnected decisions: first selecting who to interact with, and then choosing the strength of these interactions, together with the productive effort. However, the literature has explored these two dimensions of socializing separately.¹ This paper brings together both dimensions of socializing by studying a model of occupational choice with endogenous spillovers emerging from individual productive efforts and socializing decisions.

More specifically, we study how the choice between employment and entrepreneurship is affected by those who participate in each occupation, their productive effort, and the interactions they establish within the same occupational group. In our model, individuals are endowed with different (occupation-specific) abilities as, say, entrepreneurs and employees, and socializing is multidimensional. First, each individual make a decision about which occupation to join. Once this decision is taken, they choose a level of productive effort, together with the intensity of their social interactions. Employees or entrepreneurs socialize to take advantage, in an endogenous way, of spillovers emerging from the productive efforts of those in the same occupation. With these features, we provide a tractable model that allows for a complete equilibrium and welfare analysis, and generates novel results with implications for policy interventions.²

Embedding endogenous spillovers in a model of occupational choice is important for several reasons. First, because these spillovers exist: empirically

¹The study of how peers are selected has been conducted from various angles; among them, neighborhoods (Benabou, 1993), schools (Epple and Romano, 1998; Ferreyra and Kosenok, 2015), social networks (Goyal, 2012; Jackson, 2010; Vega-Redondo, 2007) and even specialties within occupations (Arcidiacono and Nicholson, 2005). The issue of within group socializing has been studied by Cabrales, Calvó-Armengol, and Zenou (2011) and Canen and Trebbi (2016).

²Although specifically designed to study occupational decisions, the model admits other interpretations, which will be occasionally discussed.

the importance of social connections for entrepreneurs (see for example Guiso and Schivardi, 2011; Guiso, Pistaferri, and Schivardi, 2015; Hoanga and Antoncic, 2003), for professionals (see for example West, Barron, Dowsett, and Newton (1999) for the medical and Ogas (2002) for the legal profession) and even for the unemployed Korpi (2001) has been widely established. Second, because they matter: Guiso and Schivardi (2011) find that spillovers rather than heterogeneous entry costs are the explanation for differences in entrepreneurial activities across Italian regions.³ Third, spillovers are likely to be endogenous: if spillovers are beneficial (damaging), rational individuals will look for ways to enhance (reduce) them. The scarce existing literature introducing spillovers into occupational choice takes them as exogenous (e.g. Guiso and Schivardi, 2011; Cicala, Fryer Jr, and Spenkuch, 2016; Chandra and Staiger, 2007). Last, as we spell out in detail below, analyzing endogenous spillovers leads to important insights. For example, we derive optimal policies to achieve efficiency in occupational choice, to correct both the within sector inefficiency caused by the externality of spillovers and the misallocations across occupations. These results allow us to rationalize why many measures implemented by governments to boost entrepreneurship have failed (see e.g. Henrekson and Stenkula, 2010; Acs, Åstebro, Audretsch, and Robinson, 2016).

In our model, productive efforts within an occupation are assumed to be complementary (i.e. spillovers are multiplicative) and the size of the resulting spillovers does not only depend on individual productive effort but also on the degree of social interactions. To fix ideas, one can view spillovers as the consequence of information sharing within an occupation, which implies that the individual marginal productivity with respect to one's own stock of human capital increases linearly in the knowledge stock of others in the same occupation. Absorbing others' information requires social interactions. Individual incentives to socialize increase with the information everyone has, which implies a complementarity between socializing and productive investments. As a result of the socializing process, spillovers within an occupation are determined by the collective output of individual productive and socializing decisions.

Our first set of results concern individual decisions of productive and socializing effort for a given occupation. All these results have empirical implications and we contrast them with some existing evidence. We show that more

³The approach on heterogeneous entry costs is implicitly followed by a large literature that focuses on (particularly financial) factors that keep the would-be entrepreneurs from actually creating a new firm (e.g. Quadrini, 2009).

talented individuals not only work harder but also generate more spillovers.⁴ Furthermore, our model predicts that on average individuals in more productive occupations work harder and socialize more.⁵ But average socializing and, hence, learning spillovers are also increasing in network synergies. As a consequence, occupations with weaker synergies should experience lower interactions and fewer spillovers.⁶ Insofar as synergies capture institutional and technological aspects of socializing, we can provide an explanation for the intensity of spillovers varying across geographical regions (e.g. Bottazzi and Peri, 2003) and over time (e.g. Jaffe, 1986). We complete our characterization of individual decisions by showing how the benefits of interaction are greater for highly productive workers; this provides a rationalization for the existence of fraternities and elite societies (e.g. Popov and Bernhardt, 2012).

Turning to the policy implications of endogenous spillovers and multidimensional socializing, we identify and distinguish two different types of inefficiencies. For a given distribution of individuals between employees and entrepreneurs, (socializing and productive) effort decisions are inefficient as a consequence of the spillovers. We show, however, that an optimal subsidy to fix effort inefficiencies for a given composition of the occupation exists and provide its explicit form.

But, since individuals sort themselves into each of the occupations, we also show that allocative inefficiencies matter. We first prove that, independently of productive and socializing efforts, there is an equilibrium division of individuals between occupations. We show that linear taxes/subsidies can be used to alter the composition of each occupation, which implies that socially optimal sorting between entrepreneurship and employment can be achieved in equilibrium.

The presence of two different sorts of inefficiencies in occupational choices, requiring specific policy instruments, implies that optimal policy responses are quite complex and single instruments may fail or even worsen the problem to be addressed. For example, the usual policy response to boost entrepreneurship involves differential taxation. But this policy response does not take into

⁴This result is consistent, for example, with Azoulay, Zivin, and Wang (2010), who show that researchers collaborating with a superstar scientist experience a significant decline in their productivity (quality adjusted publication rate) after the unexpected death of their superstar collaborators. Similarly, Waldinger (2010) find that the expulsion of high quality Jewish scientists from Nazi Germany harmed, in a significant way, their students left behind.

⁵The connection between occupation productivity and individual socializing effort is in line with Currarini, Jackson, and Pin (2009) and consistent with observations provided by Albornoz, Cabrales, Hauk, and Warnes (2017).

⁶This result is observed by Nix (2015) for the case of Sweden.

account that within sector externalities typically require a specific level of taxation to correct the “internal” effort inefficiency. It is therefore possible that a single tax to correct an allocative inefficiency might distort the internal allocation of effort. Similarly, correcting only for within sector inefficiency might worsen the across sector allocation. To examine this question more rigorously and to explore further the implications of endogenous spillovers and multidimensional socializing, we need to make specific assumptions about the distributions of occupation-specific abilities. We analyze two cases: the uniform and the Pareto distribution. In both cases, we show that the equilibrium is unique, which is a useful feature of our framework.

In the case of the uniform distribution function, we find that, independently of the effort choices, there are always fewer entrepreneurs than optimally required. Crucially, optimal subsidies designed to reach internal efficiency exacerbate allocative problems by inducing congestion, which -if costly- could imply a reduction in global efficiency. This rationalizes the observed failures in the use of subsidies to spur entrepreneurship (Henrekson and Stenkula, 2010; Acs, Åstebro, Audretsch, and Robinson, 2016). We show by example that this reduction in global efficiency can occur in our model.

Allocative inefficiency also emerges when productivities are assumed to be given by a Pareto distribution, although whether it results in overpopulation or underpopulation of the entrepreneurial sector depends on the Pareto shape parameter and the strength of the synergies. In spite of this general ambiguity, we are able to provide some sufficient conditions for under and overpopulation of a sector. For example, we show that the sector with a higher intensity of synergies (which we think likely to be the entrepreneurial sector) could be underpopulated for distributions of abilities exhibiting a relatively low level of dispersion (a shape parameter sufficiently high). Notice that a low dispersion in a Pareto distribution means that the number of superstars in that sector will be very small. Thus, if both sectors have a low number of superstars, the sector with a larger impact of synergies is likely to be underpopulated. Although this is a matter that obviously requires more research, we think it is an interesting rule of thumb to detect underpopulation in a productive sector. On the other hand, we also show that the size of the sector with higher synergies could be sub-optimally large if the synergies are sufficiently small in both sectors.

Since wages and income, at least at the top of the distribution, are well described by a Pareto (Guvenen, Karahan, Ozkan, and Song, 2015, for a recent reference), our model can associate inequality in talent with the possibil-

ity of allocative inefficiency. We show that changes that make distributions more disperse increase the average level of socializing and productive effort. This effect is not confined to the occupation where inequality increased, but also takes place in the other occupation. Distributional spillovers across occupations imply that greater inequality may lead to higher productive and socializing intensity, connecting two phenomena that are generally considered as independent from each other.⁷

Our model contributes to several aspects of the literature of occupational choice. This literature generally builds upon the seminal contribution by Lucas (1978). In Lucas (1978)'s model as well as in several follow-up papers, ability has a single dimension which implies the counterfactual prediction that all entrepreneurs should earn more income than every employee. The literature has accounted for low and high income in both sectors by adding a second dimension of ability à la Roy (1951).⁸ We follow this approach and allow for occupation-specific abilities. As a consequence, occupational choices are determined by comparative rather than absolute advantage. In this context, Rothschild and Scheuer (2012) and Scheuer (2014) study the optimal design of redistributive income taxes. We also study optimal policy instruments but our concern is efficiency not redistribution. A fundamental contribution of our approach is introducing endogenous spillovers. The few papers studying the effect of spillovers in occupational choice take them as exogenously given. In Guiso and Schivardi (2011), exogenous spillovers affect occupational choices by shifting productivity. In Cicala, Fryer Jr, and Spenkuch (2016); Chandra and Staiger (2007), exogenous spillovers change relative benefits from different activities. We complement this literature by providing a framework where individual efforts affect the level of spillovers they enjoy and derive its policy implications.

There is plenty of evidence of excessive or insufficient number of participants in specific occupations. Many countries make it a priority to spur entrepreneurship. Shakhnov (2014) finds that financial markets are overcrowded with respect to entrepreneurship and that the model matches well US data. Khabibulina and Hefti (2015) find a negative correlation of relative wages in the financial sector with respect to the manufacturing sector in the

⁷Inequality spilling over across occupations is a relatively unstudied possibility. In a recent paper, Clemens, Gottlieb, Hémous, and Olsen (2016) show that higher inequality in one occupation spills over into other occupations through consumption demand across occupations, yielding further increases in inequality.

⁸Early examples are Heckman and Sedlacek (1985, 1990) and Jovanovic (1994).

U.S. states from 1977 to 2011. Lopez-Martin (2015) obtain similar results for the allocation of workers between the formal and informal sectors. Our paper provides an explanation for these phenomena and shows that overpopulation/underpopulation can emerge in a model without much structure. More generally, our results have concrete implications for economic growth, as misallocation of talent and resources is viewed as a major force of cross country GDP and productivity differences (e.g. Murphy, Shleifer, and Vishny, 1991; Restuccia and Rogerson, 2013; Hsieh and Klenow, 2009).

There is a huge research effort to understand the effect of social relations and occupational decisions and outcomes (e.g. Granovetter, 1995; Calvo-Armengol and Jackson, 2004; Bentolila, Michelacci, and Suarez, 2010, to mention some of many contributions). The main goal of this literature is to clarify how previous social connections affect future employment decisions. In our analysis, occupational choice is driven by future socializing, not past connections. In this sense, our paper offers a new direction to explore the relationship between socializing and productive decisions.

This paper is organized as follows. Section 2 describes the model. Section 3 contains the equilibrium analysis and the general results valid for any occupation specific ability distribution. Section 4 contains the results when making specific distributional assumptions. Section 5 concludes. Most proofs are gathered in the Appendix.

2 The model - payoffs

We consider an economy with a continuum of heterogeneous individuals that choose their occupation. They can be either employees (employed in occupation M) or entrepreneurs (occupation F). Each individual i has an occupation-specific individual productivity parameter b_i^n for $n \in \{M, F\}$, which is randomly and independently drawn for each occupation. For the time being, we make no specific assumptions on how these abilities are distributed.⁹

After choosing their activity, all agents within the same occupation simultaneously decide their direct productive effort k_i^n and their socializing effort s_i^n . Socializing activities allow them to take advantage of productive efforts made by the other members of the same occupation. Consequently, the payoff within a particular occupation n is the sum of two components, a private com-

⁹In section 4, we study the cases where abilities are distributed uniformly or according to a Pareto distribution.

ponent P_i^n , and a synergistic component S_i^n derived from social interactions. The private component P_i^n has a linear-quadratic cost-benefit structure and is given by

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

where d^n is an occupation-specific parameter and is multiplicative in individual ability in occupation n . It is a useful normalization about the distribution of “final” abilities $d^n b_i^n$ that allows us to discuss the comparative statics of a change in the mean of the ability distribution while fixing the distribution of b_i^n . Obviously a shift that increases d^n involves a specific way to introduce a first order stochastically dominating shift in “final” abilities $d^n b_i^n$.

The synergistic component, S_i^n , captures that socializing is required to take advantage of the externalities generated within each occupation, which are due to the complementarity in productive efforts. In addition, socializing within each occupation is undirected.¹⁰ Specifically, this means that within occupational groups the agents only choose the amount of interaction s_i , but not the identity of the individuals with whom they interact. However, we allow individuals to choose the occupational group where they socialize. This is the way in which socializing often occurs in reality: entrepreneurs and employees go to conferences or business fairs, they join professional associations and go to their meetings, or simply share social activities or events. Synergistic effort is mostly generic within the conference, fair or social gathering; but clearly individuals carefully choose the socializing spaces they attend and the associated socializing intensity.

Denoting by \mathcal{N}_i the occupational group to which individual i belongs, the synergistic returns are given by

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

where the parameter a captures the overall strength of synergies, \mathbf{s} is the profile of all socializing efforts and $g_{ij}^n(\mathbf{s})$ is the link intensity of individual i and j , which we define below. Each occupational group is composed by a continuum

¹⁰Undirected socializing and the requirement of socializing to enjoy externalities are features shared with Cabrales, Calvó-Armengol, and Zenou (2011). However, we propose a different functional form for the benefits from synergistic returns. We will show that using our synergistic component S_i^n leads to a game with a *unique* symmetric equilibrium within a network, while the game in Cabrales, Calvó-Armengol, and Zenou (2011) has multiple equilibria. Equilibrium uniqueness in socializing and productive efforts facilitate our analysis of directed occupational choice.

of individuals $\mathcal{N}^n \subset \mathbb{R}$ for $n \in \{M, F\}$, where the measure of the set \mathcal{N}^n is N^n .

Observe that synergistic returns are multiplicative in individual productivity parameters and in the square root of productive efforts additively separable by pairs, hence productive efforts are complementary.¹¹ Adopting this specific functional form implies that synergistic returns are symmetric in pairwise productive efforts and that the synergistic returns exhibit constant returns to scale to overall productive efforts. Similar assumptions are imposed on the link intensity, which captures the extent to which individuals take advantage of the endogenously generated productive occupational externalities. Formally, these assumptions are:¹²

- (A1) Symmetry: $g_{ij}^n(s_i^n, s_j^n) = g_{ji}^n(s_j^n, s_i^n)$, for all i, j, n ;
- (A2) The total interaction intensity of individual i in group n exhibits constant returns to scale to overall inputs in socializing efforts and symmetry:

$$\int_{j \in \mathcal{N}_i} g_{ij}^n(s_i^n, s_j^n) dj = \frac{1}{N^n} \int_{j \in \mathcal{N}_i} (s_i^n)^{1/2} (s_j^n)^{1/2} dj$$
;
- (A3) Anonymous socializing: $g_{ij}^n(s_i^n, s_j^n) / (s_j^n)^{1/2} = g_{ki}^n(s_k^n, s_i^n) / (s_k^n)^{1/2}$, for all i, j, k ;

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

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

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

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

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

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 (A2) and a level of socializing effort for all members of the group, total socializing of an individual in a group $\int_{j \in \mathcal{N}_i} g_{ij}^n(s_i^n, s_j^n) dj$ is

¹¹ Complementarity in productive returns in Cabrales, Calvó-Armengol, and Zenou (2011) is generated by synergistic returns being multiplicative in productive efforts and additively separable by pairs.

¹² While Cabrales, Calvó-Armengol, and Zenou (2011) also model symmetric and anonymous socializing, which is the key for generic socializing, they assume that link intensity satisfies aggregate constant returns to scale.

independent of the size of the group. In other words, individuals will not have more contacts in larger occupational groups if everyone in the same occupation chooses the same s_i^n independent of size. One could easily accommodate other assumptions, where socializing is either easier or more difficult in larger groups by using $1/(N^n)^\beta$ for some β different from 1.

Combining the private returns and the synergistic component yields individual payoffs of an individual i in an occupational group n as:

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

We assume that individuals can only belong to one single group. This assumption is consistent with a number of potential applications: most people are either entrepreneurs or employees. They tend to have only one profession to which they dedicate themselves; academics generally do not work simultaneously in very distinct fields; top athletes generally only excel in one sport; and in spite of “Ingres’ violin” the same thing generally holds for artists.¹³ It can also be justified formally within the model in a variety of ways. For example, by adding a sufficiently large fixed cost to join a group which could arise from training costs. We also assume no specific capital requirements to become an entrepreneur. This could be due to the absence of capital market imperfections or justified by simply assuming that entry costs are similar across occupations. This way, occupational choices are not associated with initial wealth and we can focus on social interactions and productive decisions.¹⁴

Finally, the timing of events is as follows: each individual i first chooses whether to be a employee or an entrepreneur, and then takes the decisions over k_i and s_i simultaneously.

¹³The term “Ingres’ violin” comes from the French neoclassical artist Jean Auguste Dominique Ingres, who while famous for his paintings was also incredibly talented though less well known for his skill on the violin.

¹⁴See Evans and Jovanovic (1989) for the seminal contribution on the analysis of the effect of liquidity constraints on entrepreneurial choice.

3 The equilibrium and general results

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

3.1 Choice of production and socializing efforts

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

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

Denote by $\bar{b}^2 = \int_{j \in \mathcal{N}_i} b_j^2 dj$. We assume that

Assumption 1. $\sup_C \left(a d^2 \bar{b}^2 \right)^2 < 1$.

We can now derive the equilibrium decisions in terms of productive and socializing efforts, which we state as follows:

Proposition 1. *Under assumption 1, both the individual choice problem and the social planner choice problem have a unique (interior) solution which for each individual is equal to her own productivity multiplied by a function that is identical for all individuals in the group.¹⁵ That is*

$$k_i = b_i k \text{ and } s_i = b_i s \text{ for all } i \quad (4)$$

$$k_i^s = b_i k^s \text{ and } s_i^s = b_i s^s \text{ for all } i \quad (5)$$

¹⁵The individual choice problem also has a trivial partial corner solution where $s_i = 0$. If nobody socializes, socializing is not profitable. However, this equilibrium is not stable, since the marginal utility of s_i is positive for any (even infinitesimally small) average level of socializing in the occupational group. We therefore ignore this solution in our analysis.

where the optimal common group functions for productive and socializing effort are given by

$$k = \frac{d}{1 - \left(\frac{a}{2}d^2\bar{b}^2\right)^2}, \quad (6)$$

$$s = \frac{\frac{a}{2}d^3\bar{b}^2}{1 - \left(\frac{a}{2}d^2\bar{b}^2\right)^2}, \quad (7)$$

for the individual choice problem, and by

$$k^s = \frac{d}{1 - \left(ad^2\bar{b}^2\right)^2}, \quad (8)$$

$$s^s = \frac{ad^3\bar{b}^2}{1 - \left(ad^2\bar{b}^2\right)^2}, \quad (9)$$

for the social planner.

Proof. See Appendix A.1. □

There are a couple of aspects worth of notice in these solutions. An immediate observation is that since productivity b_i is complementary to effort, more talented individuals will also tend to work harder.¹⁶ On the other hand, the common group functions are increasing in the average group squared productivity \bar{b}^2 and in average group productivity $\bar{b} = \int_{j \in \mathcal{N}_i} b_j dj$. Since individual socializing is $s_i = b_i s$, average socializing is $\bar{b} s$. Thus, an interesting corollary of Proposition 1 follows:

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

¹⁶The correlation between talent and effort has been observed in education; a sector for which we have good data on both ability and effort (see e.g. Yeo and Neal (2004) and Babcock and Betts (2009)). But these individual features also translate to the group level, something that allows to make intergroup comparisons as well. On the one hand, high talented individuals generate greater externalities on their fellows. Evidence consistent with this result is observed in the academic world. For example, the sudden absence of extremely highly productive researchers provides a natural test for our prediction. Azoulay, Zivin, and Wang (2010) find that researchers collaborating with a superstar scientist experience a lasting and significant decline in their quality adjusted publication rate after the unexpected death of their superstar collaborator. A result similar in spirit is provided by Waldinger (2010) when showing that the expulsion of high quality Jewish scientists from Nazi Germany has a negative effect on the productivity of the PhD students left behind.

The empirical implication of Corollary 1 is that individuals within more productive occupational groups socialize more on average.¹⁷

Average socializing and hence learning spillovers are also increasing in network synergies a . Thus, occupation with fewer synergies should experience lower interactions and fewer spillovers.¹⁸ Insofar as synergies capture institutional and technological aspects of socializing, we can also provide an explanation for the intensity of spillovers varying across geographical regions (e.g. Bottazzi and Peri, 2003) and over time (e.g. Jaffe, 1986).

Since we have derived the optimal efforts, we can obtain the associated individual utilities, which are given by

Proposition 2. *The resulting individual utilities are*

$$u_i(b_i) = \frac{b_i^2 d^2}{2} \left(\frac{\left(1 + \left(\frac{a}{2} d^2 \bar{b}^2\right)^2\right)}{\left(1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)^2\right)^2} \right), \quad (10)$$

in the individual choice problem and

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

for the social planner solution.

Proof. See Appendix A.1. □

An implication of Proposition 2 is that while all individuals benefit from being in a more productive occupational group (since $\partial u_i^s(b_i)/\partial \bar{b}^2 > 0$ so util-

¹⁷This empirical implication of our model is consistent with evidence presented in Currarini, Jackson, and Pin (2009) showing that the number of interactions within friendship groups are increasing in size. Albornoz, Cabrales, Hauk, and Warnes (2017) provide further empirical evidence for this prediction based on the analysis of co-authorships within economics fields. Furthermore, academic life is clearly an example of a situation in which an individual's productive outcomes are affected by the abilities and activities of other researchers involved in the same production process. Hence socializing decisions become key productive choices. Moreover academics choose their field of research: their group. Using data scrapped from the IDEAS-RePEc website Albornoz, Cabrales, Hauk, and Warnes (2017) establish that economic researchers who work in more productive fields tend to have more co-authors.

¹⁸This is indeed found by Nix (2015) for the case of Sweden. After constructing a ranking of interactions with peers using Swedish data on workers, their peers, and their firms from 1985-2012, Nix (2015) compares it to estimated learning spillovers per-occupations and finds a strong correlation between those two measures.

ity increases in \bar{b}^2 for everyone), higher types benefit even more from a given level of within-occupation externality (that is, $\partial^2 u_i^s(b_i)/\partial b_i \partial \bar{b}^2 > 0$, so that individual type and group type are complementary). This creates an incentive for high types to segregate from low types if possible, as productivity is independent of occupational group size for a given average spillover, as shown above by Lemma 1. We certainly observe a tendency for high-skilled employees or entrepreneurs to create elite societies. Good examples are the Freemasons or the Rotary club (Yanagida (1992), Burt (2003)) where access is restrictive and whose objective seems to be mainly to socialize among like-minded high-skilled individuals. These examples are particularly interesting because they are often secretive, i.e., they are not created for the purpose of signaling such quality to the external world.¹⁹

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

Proposition 3. *Individuals underinvest in both productive and socializing effort ($k^s > k$ and $s^s > s$).*

Underinvestment clearly creates a rationale for subsidizing some professional activities where learning spillovers are important for productivity.²⁰ As discussed in the introduction, Guiso and Schivardi (2011) find that spillovers rather than heterogeneous entry costs are the explanation for differences in entrepreneurial activities across Italian regions.²¹ In the last few decades, en-

¹⁹There are other examples where elite groups use restricted settings to socialize, like London clubs in the late 1800s and early 1900s (Brayshay, Cleary, and Selwood (2006), Brayshay, Cleary, and Selwood (2007)). Also, fraternities in college serve the purpose of segregation, are mainly for networking and have a positive effect on future income. Marmaros and Sacerdote (2002) report that fraternity membership is positively associated with networking and with finding a high paying job directly out of college. Routon and Walker (2014) confirm that fraternity membership increases the probability of a recent graduate obtaining a job. Mara, Davis, and Schmidt (2016) find that fraternity membership increases expected future income by roughly 30%.

²⁰This is clear in the high-tech industry. To cite one example, Pirolo and Presutti (2007) analyze the metropolitan high-tech cluster in Rome and show that social interactions are the most significant determinant of the innovation process and relationships based on knowledge sharing are the most important ones.

²¹The approach on heterogeneous entry costs is implicitly followed by a large literature that focuses on (particularly financial) factors that keep the would-be entrepreneurs from actually creating a new firm, as described in Quadrini (2009)

trepreneurship has emerged as a key issue in the policy arena.²² For instance the European Commission launched the “Small Business Act for Europe” in June 2008, which explicitly recognizes the central role of innovative small and medium-size enterprises (SMEs) in the EU economy and sets out a comprehensive policy framework for the EU and its member states. In this document, the Commission proposes that member states should create an environment that rewards entrepreneurship, specifically mentioning taxation in this context. Since entrepreneurial effort in particular, and effort within an occupation in general, is suboptimal in the presence of spillovers, we now turn to the determination of an optimal subsidy within each occupation.

Denote individual output in a given occupation y_i by

$$y_i \equiv db_i k_i + adb_i (k_i)^{1/2} \int_{j \in \mathcal{N}_i} \left(db_j (k_j)^{1/2} g_{ij}(\mathbf{s}) \right) dj.$$

Proposition 4. *A subsidy that achieves efficient effort within an occupation (taking as given the selection into occupations) is given by:*

$$y_i - d \frac{(k_i)^2}{k}. \quad (12)$$

Proof. See Appendix A.2 □

This subsidy, which is based on observable individual output and productive effort, alters the original utility in a way that induces socially optimal levels of effort. However, it takes as given the selection into occupations. For this reason, it is only part of an optimal policy. Individuals choose their occupation, and these individual choices might not be efficient. We now analyze the optimal individual occupational choice and then return to the issue of taxation to induce efficient occupational choices.

3.2 Choice of occupation

Having found the second-stage utilities, we can now solve the first-stage in which individuals sort themselves into either employees (group M) or entrepreneurs (group F). We show now that independently of whether productive or socializing efforts within the occupation are individually chosen (decentralized solution) or by the social planner, the solution is characterized

²²The Economist on 14th March 2009 published a special report on entrepreneurship with the title “Global Heroes”.

by a dividing line $b_i^M = Cb_i^F$ such that individuals who fall below the line choose group F , while individuals above the line choose group M . Such a dividing line implies that

$$\overline{b^{M^2}} = E\left(b_i^{M^2} | b_i^M > Cb_i^F\right), \quad (13)$$

$$\overline{b^{F^2}} = E\left(b_i^{F^2} | b_i^M < Cb_i^F\right). \quad (14)$$

In other words, comparative advantage determines the choice of occupation in a particularly simple way. For individuals whose ratio of types b_i^M/b_i^F is bigger than C , they choose to become employees (group M). On the other hand, individuals whose ratio of types b_i^M/b_i^F is lower than C choose to become entrepreneurs (group F). Naturally, C is an endogenous function of all the parameters in the model, and in general, it need not be unique.

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

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

Under the decentralized solution, individuals choose to become an employee (group M) if and only if $u_i(b_i^M) \geq u_i(b_i^F)$. Hence, whenever

$$\frac{b_i^{M^2} d^{M^2}}{2} \left(\frac{\left(1 + \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)}{\left(1 - \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2} \right) > \frac{b_i^{F^2} d^{F^2}}{2} \left(\frac{\left(1 + \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)}{\left(1 - \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2} \right). \quad (15)$$

If the dividing line exists, its slope is defined when the expressions on either side of the inequality in (15) are equal. In other words, the dividing line is defined by the following expression:

$$b_i^M = b_i^F \frac{d^F}{d^M} \sqrt{\frac{\left(1 + \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right) \left(1 - \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(1 - \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2 \left(1 + \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}} = b_i^F C_P.$$

Hence C_P is the fixed point of the mapping

$$C_P = \frac{d^F}{d^M} \sqrt{\frac{\left(1 + \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right) \left(1 - \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(1 - \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2 \left(1 + \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}}, \quad (16)$$

where the right hand of (16) depends on C_P through $\overline{b^{M^2}}$ and $\overline{b^{F^2}}$, which are defined by equations (13) and (14) respectively. Put differently, C_P is implicitly defined by a zero of the mapping

$$g(C, \cdot) \equiv \frac{d^{F^2}}{d^{M^2}} \frac{\left(1 + \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right) \left(1 - \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(1 - \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2 \left(1 + \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2} - C^2. \quad (17)$$

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

$$C_E = \frac{d^F}{d^M} \sqrt{\frac{\left(1 - \left(a d^{M^2} \overline{b^{M^2}}\right)^2\right)}{\left(1 - \left(a d^{F^2} \overline{b^{F^2}}\right)^2\right)}}, \quad (18)$$

and is implicitly defined by a zero of the mapping

$$f(C, \cdot) \equiv \frac{d^{F^2}}{d^{M^2}} \frac{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2} - C^2. \quad (19)$$

We can use the implicit definitions of C_P and C_E to derive comparative static results based on the implicit function theorem. First observe that

Proposition 5. *For any underlying distribution of abilities, if assumption 1 is satisfied, a zero of the mappings $f(C)$ and $g(C)$, and thus an equilibrium, always exists. Furthermore, in any stable equilibrium, $\frac{\partial f(C)}{\partial C} < 0$ and $\frac{\partial g(C)}{\partial C} < 0$.*

Proof. See Appendix A.3. □

This is a helpful technical result. Basically, $\partial f(C)/\partial C < 0$ and $\partial g(C)/\partial C < 0$ mean that establishing comparative static results only requires checking the sign of the derivatives of the functions defining C_P and C_E with respect to the underlying parameters a^n and d^n for $n \in \{M, F\}$.

Proposition 6. *For any underlying distribution of abilities, if assumption 1 is satisfied, both C_P and C_E are decreasing in a^M and d^M and they are also both increasing in a^F and d^F .*

Proof. See Appendix A.4. □

When C decreases more people become employees (join the M -group). Similarly, an increase in C implies that more people become entrepreneurs (join the F -group). Thus, according to Proposition 6, an increase in the power of synergies, or a specific first order stochastic dominance shift in the distribution of final abilities, will lead to more people joining the affected occupation.

Higher within occupation synergies a might be caused by the introduction of new or improved communication technology facilities. The effect of these technologies on productivities has been widely acknowledged.²³ To our knowledge, there is no study linking the relative sizes of economic sectors with their differential adoption of communication technologies. This paper provides clear predictions linking relative sector sizes with other observable characteristics. These predictions can be tested in future research and exhibit the nice feature of being independent of the underlying distribution of abilities.

Similarly, our model delivers clear and testable predictions for a shift in d . Such shifts could be technological changes that affect the productivity of every individual in a given occupation. Or they could be due to institutional features. For example, in some institutional settings very large (or very small)

²³Some examples are: Lio and Liu (2006) who find positive and significant relationship between the adoption of information and communication technology and agricultural productivity based on data collected in 81 countries for the period 1995–2000. Bayes (2001) find that village phones in Bangladesh allow farmers to obtain better prices through information diffusion. Genius, Koundouri, Nauges, and Tzouvelekas (2013) provides this type of evidence for olive farmers in Greece while Sene (2015) studies peanut farmers in Senegal and obtains a positive impact of connectivity on output.

firms are extremely regulated, while in others there are too many loopholes for politically connected firms. If we interpret our model as choosing to work in the formal or informal sector, for example, a looser degree of control would induce a first-order stochastic dominant shift in the profitability of the informal sector and it will increase the attractiveness of that sector. A high taxation level in the formal sector will have a similar effect.²⁴

The above results only indicate how the relative occupational sector sizes change with the underlying parameters, but they do not inform us about the efficiency or inefficiency of the equilibrium outcomes. Is there an occupational sector which is too big or too small? We will turn to this question in Subsection 4.1 where we make specific assumptions for the underlying talent distributions. But before assuming specific distributions, we can show in all generality that a social optimal sorting into occupations can always be achieved since:

Proposition 7. *Any $C \in [0, \infty)$ can be obtained in equilibrium using a linear tax/subsidy on output.*

Proof. See Appendix A.5. □

From proposition 7 it is immediate that

Corollary 2. *The socially optimal C can be achieved in equilibrium using a linear tax/subsidy on output.*

4 Additional results for specific distributions of talent

We now wish to derive additional insights from our model. At this stage, we assume specific ability distributions. We study the cases of the Uniform and Pareto distributions of talent. The following results state that both distributions deliver a unique equilibrium.

Proposition 8. *If abilities are uniformly and independently distributed in $[0, B^n]$ for $n \in \{M, F\}$, both C_P defined by (16) and C_E defined by (18) exist and are unique.*

Proof. See Appendix A.6. □

²⁴Lopez-Martin (2015) find plentiful evidence for these effects.

Proposition 9. *If abilities are distributed independently and follow a Pareto law in $[1, \infty)$ with shape parameter α_j for $j \in \{M, F\}$, both C_P , defined by (16) and C_E , defined by (18) exist and are unique.*

Proof. See Appendix A.7. □

Uniqueness is a useful feature of our model to study the emergence of efficient allocations of individuals across occupations. Global allocative efficiency is discussed in Section 4.1 for both distributions. The uniform distribution allows to show in a simple way the policy implications of allocative inefficiency (Section 4.2). The Pareto distribution provides a framework to study the effect of inequality on socializing and productive efforts, and the emergence of inequality spillovers across sectors (Section 4.3).

4.1 Global allocative inefficiency in the absence of intervention on location

We focus here on the relationship between decentralized and centralized effort choices within occupations and social optimality in occupational choice. We begin with the case of abilities being uniformly and independently distributed. Without loss of generality, we let $B^M > CB^F$. This assumption simply says that the general employment sector is larger than that of the entrepreneurs. To see whether occupations are over or under populated, we simply need to compare the equilibrium cut-off dividing lines. Notice, for example, that occupation F is overpopulated if the social planner would choose a cutoff that lies to the right of those cutoffs chosen by the individuals in both cases studied above. Formally, overpopulation of the occupation F is implied by $C_E^* > C_E$ and $C_P^* > C_P$, where C_E^* is the cutoff a social planner would choose when effort choices in the groups are centralized while C_P^* is the cutoff the social planner would choose when effort choices in the groups are decentralized. The next result shows that this is indeed the case:

Proposition 10. *If abilities are uniformly and independently distributed in $[0, B]^n$ for $n \in \{M, F\}$, social welfare is increasing in C for all $C \leq C_E$ and $C \leq C_P$.*

Proof. See Appendix A.8. □

Proposition 10 implies that with a uniform distribution of individual talent, too few people join the F group, independently of whether the choice of

productive and socializing efforts are efficient or not. The intuition is simple. At either C_E or C_P , individuals at the margin, for whom b_i^M is very close to Cb_i^F , are almost indifferent about which occupation to choose. For this reason, moving them from one occupational group to the other does not affect much their utilities. However, this relocation affects welfare in both occupational groups by affecting their average type. For entrepreneurs (occupation F), this effect is almost non-existing because its average type does not depend on C under the uniform distribution (see equation (40) in Appendix A.6). For the general occupation (M group), the average type improves with C (see equation (42) in Appendix A.6) and hence average welfare increases when the indifferent, and close to indifferent, M -types in the group are induced to join the entrepreneurial sector. Society would be better off had they joined the F group. This occurs independently of whether productive and socializing efforts are generated in a socially optimal way, or in a decentralized way. This is important because, as we mentioned previously, the fact that some activities are not sufficiently populated from a social point of view motivates interventions, such as incentives to become entrepreneurs (Haufler, Norbäck, and Persson (2014)).

However, the stark result according to which it is always the same occupation that is overpopulated hinges on the assumption that individual productivities are uniformly distributed. When talent follows a Pareto distribution, occupations are still inefficiently populated but in a way that depends on the underlying parameters. To save space, we just compare C_E^* to C_E (i.e. the case where investment in productive and socializing effort are optimally determined by a social planner). The next result characterizes allocative inefficiency for the case of the Pareto distribution:

Proposition 11. *Let abilities be independently distributed, and assume they follow a Pareto law in $[1, \infty)$ for $n \in \{M, F\}$ with a common shape parameter α . Then there might be too few (i.e. $\frac{\partial w(C)}{\partial C} \Big|_{C=C_E} > 0$) or too many people (i.e. $\frac{\partial w(C)}{\partial C} \Big|_{C=C_E} < 0$) in occupation F compared to the social optimum.*

In particular:

- For fixed values of a^F , a^M , d^F and d^M , satisfying $a^{M^2} d^{M^2} < a^{F^2} d^{F^2}$, there is a value of α high enough such that $\frac{\partial w(C)}{\partial C} \Big|_{C=C_E} > 0$.
- Also, for fixed values of d^F , d^M and for $\left(a^{M^2} d^{M^2}\right) \left(a^{F^2} d^{F^2}\right) < 1$, there is an a^F low enough such that $\frac{\partial w(C)}{\partial C} \Big|_{C=C_E} < 0$.

Proof. See Appendix A.9. □

Notice that $a^{M^2}d^{M^2} < a^{F^2}d^{F^2}$ implies that the overall strength of synergies for entrepreneurs is higher than for employees. In Lemma 4 (Appendix A.9), we show that $a^{M^2}d^{M^2} < a^{F^2}d^{F^2}$ is equivalent to $C_E > 1$. An important implication of $C_E > 1$ is that the average type in occupation F decreases with C , while the average type in occupation M increases with C .²⁵ Thus, under $C_E > 1$, reallocating M -types that are close to indifferent to occupation F lead to lower welfare in occupation F , since the average type in occupation F decreases. At the same time, welfare in occupation M increases because the average type in occupation M increases. The overall effect on social welfare is therefore ambiguous.

In spite of the general ambiguity, we are able to provide sufficient conditions for under- and overpopulation of the entrepreneurial sector when $C_E > 1$. Notice that Proposition 11 establishes that occupation F is underpopulated for distributions with relatively low dispersion (high values of α). Notice as well that a low dispersion in a Pareto distribution means that the tails of the distribution are thin, which implies that the number of superstars is very small. Thus, if both sectors have a low number of very able individuals, the entrepreneurial sector, which under $C_E > 1$ has a larger impact of synergies, will be underpopulated. We are not aware of any empirical research documenting whether these conditions ($a^{M^2}d^{M^2} < a^{F^2}d^{F^2}$ and α small) are satisfied. This is a matter that obviously requires further research. However, this result provides a potentially useful rule of thumb to detect underpopulation/overpopulation of different occupational sectors.

Proposition 11 also establishes that the size of the occupation with higher synergies is sub-optimally large when synergies are sufficiently small in both occupations. To see this, notice that occupation F is overpopulated when a_F (the strengths of synergies in F , and also in M , since $a^{M^2}d^{M^2} < a^{F^2}d^{F^2}$ for $C_E > 1$) is very low.

Although the direction of overpopulation is conditional on the distribution of abilities, the general message of our analysis is relevant in its own right: decentralized selection involves sub-optimal composition of occupations. As a consequence, the social optimum is achieved when the central planner intervenes at both margins, i.e. by inducing optimal efforts within the occupation and chooses the optimal ratio C_E^* .

²⁵If $C_E < 1$, we just have to interchange the labels of the group.

There is plenty of evidence of excessive or insufficient size of specific occupations. Shakhnov (2014) shows that financial markets are overcrowded with respect to entrepreneurship and that the model matches well US data. Khabibulina and Hefti (2015) find a negative correlation of relative wages in the financial sector with respect to manufacturing sector in case of the U.S. states from 1977 to 2011. A similar, while somewhat less robust, result applies to the case of relative sector sizes as measured by the labor force. Our paper provides an explanation for these phenomena and shows that productive and informational spillovers are prime candidate mechanisms for overpopulation/underpopulation to emerge in economic sectors.

4.2 Welfare-reducing policy interventions with congestion

In this section, we assume that abilities are uniformly and independently distributed in $[0, B^n]$ for $n \in \{M, F\}$. Proposition 10 established that $C_E^* > C_E$, however, this result is silent towards the position of C_E^* with respect to C_P . This poses an interesting question, since C_E requires intervention by the social planner when choosing productive and socializing efforts, while C_P is the cutoff chosen by individuals in the absence of any intervention. Can no intervention be better than intervening at one margin only? We turn to this question in a setup where we allow for congestion. We first show that inducing the optimal socializing and production effort can induce over-congestion. Then we show that if congestion is costly, inducing within-group efficiency without correcting for misallocations in occupational choice can reduce global efficiency.

Optimal socializing and production effort can induce excessive congestion

We want to show the existence of parameter values for which a decentralized choice of occupation, together with an optimal choice of socializing and production efforts can lead to over-congestion in the occupation of general employees.

Proposition 12. *Suppose $a^M = a^F$, $d^M = d^F$ and $B^M = B^F + \varepsilon$. For ε small enough we have that $C_E < C_P < 1 < C_E^*$.*

Proof. Suppose first that $a^M = a^F$, $d^M = d^F$ and $B^M = B^F$. It is easy to see that in the absence of any asymmetries $C_E^* = C_E = C_P = 1$. It is then optimal,

both socially and individually, for individuals to sort into the occupation where they have the higher productivity draw b_i^j . Now, suppose we give a small advantage to occupation M , by increasing B^M . Since the F parameters are left unchanged also the utility for an individual from choosing occupation F is unaffected (since average type in occupation F is independent of C). From the definitions of C_E and C_P it is easy to see that increasing B^M will reduce C_E and C_P : for a fixed C , the right hand side of both defining equations (16) and (18) decrease if B^M increases. Hence to preserve the equality, the variable C has to fall. An increase in B^M increases average type in occupation M and therefore it draws more people into this occupation. However, individuals with lower b_i^M than before the increase in B^M are now drawn into occupation M , and those individuals lower the average type in occupation M , which eventually stops the inflow. This effect is stronger when efforts are induced optimally since the optimal effort choices allow individuals to take more advantage of the improved parameters in occupation M , hence $C_E < C_P < 1$ when $a^M = a^F$, $d^M = d^F$ and $B^M = B^F + \varepsilon$, for ε small enough. From the point of view of the social planner, when $a^M = a^F$, $d^M = d^F$ and $B^M > B^F$ restricting occupation M (i.e. inducing more people to become entrepreneurs), does not affect the average type in occupation F because it is the marginal type that joins occupation F , while it improves the average type in occupation F . From the above discussion, it is immediate that $C_E < C_P < 1 < C_E^*$ and the result follows. \square

$C_E < C_P < 1 < C_E^*$ corresponds to situations in which achieving within occupation efficiency induces a lower number of individuals to become entrepreneurs. Under these circumstances, the regulating government operating only on one margin is “wasting” part of the effort because it generates a counter reaction on the other margin it does not control. For this reason, inducing efficient efforts for a given composition of occupations induces an even more severe underpopulation of the entrepreneurial occupation than in the absence of any intervention. We next show that this can lead to a reduction in global efficiency if congestion is costly.

Congestion can reduce global efficiency

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

$$f(C, b_i^M) = \begin{cases} 1 & \text{if } b_i^M < C^* B^F \\ \frac{(C^* B^F)^2}{b_i^{M^2}} + (1 - v(C)) \left(1 - \frac{(C^* B^F)^2}{b_i^{M^2}}\right) & \text{if } b_i^M \geq C^* B^F \end{cases} , \quad (20)$$

where for the time being $v(C)$ is any function that guarantees $f'_C(C, b_i^M) \geq 0$.²⁶ This captures the fact that congestion is more harmful the larger the population in M , since there are fewer people in M the larger is C . Observe that (20) takes away part of the welfare of b_i^M types above $C^* B^F$ and make it closer to the welfare of type $b_i^M = C^* B^F$ when C is progressively smaller.²⁷ We could also allow for congestion in the other group, but this would not change the qualitative results. So, for notational simplicity, we apply congestion only in group M . This way of describing congestion has the advantage that it does not alter our equilibrium location analysis. The reason is that it takes welfare away only from agents that are “supramarginal”, i.e., they will choose to go to occupation M anyway. This is admittedly artificial, but the point of this exercise is only to highlight a theoretical possibility, and this particular modeling device is the simplest one that delivers the conclusion in a transparent way.

Proposition 13. *Suppose congestion costs are given by $f(C, b_i^M)$ defined in equation (20), and also that there is no intervention in occupational choice. An intervention designed to optimize the s_i, k_i choice within occupations, taking as given the equilibrium occupational choice might lead to a lower welfare than no intervention. That is, there are parameter values for which welfare under socially optimal socializing and productive efforts within occupations is lower than the welfare with individually optimal choice of both socializing and productive effort.*

To prove Proposition 13, we first derive an expression for welfare.

²⁶We choose a specific $v(C)$ in Lemma 3 since we illustrate with a specific example that an intervention to correct within-occupational inefficiency without accounting for across occupation misallocations can be worse than no intervention at all.

²⁷Other ways that take surplus away from high b_i^M and are also related to C would also work.

Lemma 2. *In the presence of congestion, social welfare when the government induces efficient productive and socializing efforts within occupations is given by*

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

where

$$G_E(C) = \frac{1}{B^F B^M} \left(\int_0^{B^F} \int_{CB^F}^{B^M} \frac{b_i^{M^2}}{2} \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \bar{b}^{M^2})^2} \right) db_i^M db_i^F \right) (21) \\ - \frac{1}{B^F B^M} \left(\frac{(B^M - CB^F) C^2 B^{F^3}}{2} \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \bar{b}^{M^2})^2} \right) \right).$$

Similarly, welfare in the absence of government intervention is given by

$$w_P(C) = \frac{CB^{F^3}}{2B^M} \left[\frac{d^{F^2} \left(4 + (a^F d^{F^2} \bar{b}^{F^2})^2 \right)}{\left(4 - (a^F d^{F^2} \bar{b}^{F^2})^2 \right)^2} + \frac{d^{M^2} \left(4 + (a^M d^{M^2} \bar{b}^{M^2})^2 \right)}{\left(4 - (a^M d^{M^2} \bar{b}^{M^2})^2 \right)^2} C^2 \right] \\ + 2 \left(1 - C \frac{B^F}{B^M} \right) (CB^F)^2 \frac{d^{M^2} \left(4 + (a^M d^{M^2} \bar{b}^{M^2})^2 \right)}{\left(4 - (a^M d^{M^2} \bar{b}^{M^2})^2 \right)^2} + (1 - v(C)) G_P(C),$$

where

$$G_P(C) = \int_0^{B^F} \int_{CB^F}^{B^M} \frac{2b_i^{M^2}}{B^F B^M} \frac{d^{M^2} \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{\left(4 - \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} db_i^M db_i^F \quad (22)$$

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

Proof. See Appendix A.10 □

To complete the proof of Proposition 13, we only need to find a function $v(C)$ and some parameter values such that no intervention delivers more welfare than the intervention which is focused only on inducing optimal productive and socializing efforts within an occupation. This is done in Lemma 3. It assumes that congestion only has a bite in extremely crowded occupations.

Lemma 3. *Let $v(C) = 1$ for C close to zero, and for C bounded away from zero $v(C) = 0$. Then if $1 - \left(a^M d^{M^2} \frac{B^{M^2}}{3} \right)^2$ (sufficiently small) and B^M big enough, $w(C_E) < w(C_P)$.*

Proof. See Appendix A.11 □

Proposition 13 points to the possibility that intervention at one margin might be worse than no intervention. If there are congestion costs in occupational choice, the absence of any government intervention can be socially better than the existence of local intervention via transfers and taxes that induce the efficient effort levels within the occupations. A parallel result is found in the 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 (2017) have shown that a reduction in class size leads to lower parental effort and hence little (or no) improvement in overall educational performance.

4.3 Inequality and effort choices

We explore now the link between inequality of abilities and productive and socializing efforts. We focus on the Pareto distribution because of its capacity to describe the variation of wages and income (and thus, indirectly, in talent)

across individuals (e.g. Mandelbrot, 1960; Guvenen, Karahan, Ozkan, and Song, 2015). Notice that the shape parameter α_i is an (inverse) measure of the spread of talent. Thus, we can simply associate a general increase of inequality with a reduction of α_i . One difficulty with the Pareto, though, is that reducing α_i increases both mean and dispersion. To circumvent this problem, we look at the effect of a “neutralized” reduction in α_j . More specifically, as α_j falls we impose an equivalent change in a to reduce effort as much as necessary to fix the unconditional mean of the Pareto distribution; which, remember, is $E(b_i^{j^2}) = \alpha_j / (\alpha_j - 2)$, for $j \in \{M, F\}$. This way, we focus exclusively on the effect of changes in the dispersion of talent, which we associate with inequality. We can at this point state:

Proposition 14. *Suppose abilities are distributed independently and follow a Pareto law in $[1, \infty)$ with shape parameter α_j for $j \in \{M, F\}$. Suppose as well that the shape parameter α_j of one of occupations decreases and that a is reduced to exactly compensate for the increase (i.e. $a' = a(\alpha_j - 2)\alpha_j$) in the unconditional mean of squared types. Then, if we hold C_E or C_P constant, both $\overline{b^{M^2}}$ and $\overline{b^{F^2}}$ increase, and thus productive and socializing effort increase in both occupations.*

Proof. See Appendix A.12. □

As a basic intuition, notice that as the dispersion in one of the distributions increases, both occupations receive a better selection of types. The tails of one of the distributions is now larger and comparative advantage forces a selection mostly from the tails. For a more analytical explanation, let us rewrite the expression for $\overline{b^{F^2}}$ to obtain:

$$\begin{aligned} \overline{b^{F^2}} &= E(b_i^{F^2} | b_i^M < C b_i^F) = \frac{\int_1^\infty \int_1^{C b^F} b^{F^2} f_F(b^F) f_M(b^M) db^M db^F}{\int_1^\infty \int_1^{C b^F} f_F(b^F) f_M(b^M) db^M db^F} \\ &= \int_1^\infty b^{F^2} f_F(b^F) \frac{F_M(C b^F)}{\int_1^\infty f_F(b^F) F_M(C b^F) db^F} db^F. \end{aligned}$$

Observe that if α_M decreases, the amount of mass on the tail of the distribution increases. In this way, the weight given to larger values of b^F increased by a (now larger) factor $F_M(C b^F) / \int_1^\infty f_F(b^F) F_M(C b^F) db^F$.

The effect of a decrease of α_F is more direct, as it increases $f_F(b^F)$ for larger values of b^F . But of course, we are compensating for the direct increase by reducing a . But the key difference in the conditional expectation is that

the $F(Cb^F) \int_1^\infty f(b^F) F(Cb^F) db^F$ term, now unchanged, gives more weight to changes that occur for higher values of b^F .

Clearly, the effect of inequality in talent on socializing and productive efforts emerges from the existence of spillovers within occupations. This is a novel empirical implication of our model that stands as a challenge for future empirical work.

5 Conclusion

In this model, we study a model that integrates productive and socializing efforts with occupational choice. Socializing allows for capturing informational spillovers between individuals. We show that the existence of spillovers leads to some interesting implications. It causes more talented individuals to work harder, generating bigger positive externalities within their occupation, but they also have incentives to segregate. We also show that average socializing increases in average group productivity and in network synergies. Also, any increase in within occupation synergies or improvement in final abilities for an occupation causes more people to choose this occupation no matter how abilities in the different occupations are distributed. This result provides interesting testable implications on how sector sizes should vary, for example, after the generalization of new communication technologies, which may differentially be adopted across sectors. Another interesting implication of endogenous spillovers is higher inequality of abilities imply more socialization and productive efforts. Interestingly this effect influences also other occupations. This is something that would not happen in a world without spillovers within occupations.

As one would expect in a model with complementarities, individuals underinvest in productive and social effort. We solve for the optimal subsidy within occupations. However, since occupational choice is endogenous, solving only the within-group investment problem is not sufficient for efficiency, since individuals tend to sort themselves into occupations in a socially inefficient way in the presence of spillovers. Our paper shows that the socially optimal sorting can always be achieved by a linear tax on output no matter how talents are distributed across occupation. Policy-makers have to combine this optimal linear tax with the optimal subsidy for efficient effort within occupations. If they only intervene at one margin, the result might be worse than no intervention at all. We show that inducing optimal effort only within an occupation may

exacerbate misallocations of individuals across occupations. In fact, it can do so to the point that in the presence of congestion costs it may generate an even lower social welfare than no intervention at all.

Since the Pareto distribution is widely used to model wage (and thus indirectly talent distributions), and arguably wages and income are well described by a Pareto law (Guvenen, Karahan, Ozkan, and Song, 2015, for a recent reference), our model provides two rules of thumb to identify potentially over- or underpopulated occupational sectors. The first rule applies to a situation with a low dispersion in the Pareto distribution, implying that there are few superstars in each sector. In such a situation, the sector with a larger impact of synergies is likely to be underpopulated. The second rule applies to a situation where synergies are small in all occupational sectors. If this is the case, overpopulation in the sector with higher synergies is likely to occur.

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.²⁸ Hence in the long-term groups need not be type-biased. We could extend our model to allow for participation in more than one occupation over time and thus ascertain if biases in occupational choice persist over time. Clearly, another extension would be to allow some spillovers between groups and partial participation of agents in several of them. We could also allow for horizontal preferences over occupations which are not necessarily related to individual productivity and for correlated productivities across occupations.

References

- ACS, Z., T. ÅSTEBRO, D. AUDRETSCH, AND D. T. ROBINSON (2016): “Public policy to promote entrepreneurship: a call to arms,” *Small Business Economics*, 47(1), 35–51.
- ALBORNOZ, F., S. BERLINSKI, AND A. CABRALES (2017): “Motivation, re-

²⁸Another 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).

- sources and the organization of the school system,” *Journal of the European Economic Association*, Forthcoming.
- ALBORNOZ, F., A. CABRALES, E. HAUKE, AND P. E. WARNES (2017): “Intergenerational field transitions in economics,” *Economics Letters*, 154, 1–5.
- ARCIDIACONO, P., AND S. NICHOLSON (2005): “Peer effects in medical school,” *Journal of Public Economics*, 89(2), 327–350.
- AZOULAY, P., J. S. G. ZIVIN, AND J. WANG (2010): “Superstar extinction,” *The Quarterly Journal of Economics*, 125, 549–589.
- BABCOCK, P., AND J. R. BETTS (2009): “Reduced-class distinctions: Effort, ability, and the education production function,” *Journal of Urban Economics*, 65(3), 314–322.
- BAYES, A. (2001): “Infrastructure and rural development: insights from a Grameen Bank village phone initiative in Bangladesh,” *Agricultural Economics*, 25(2-3), 261–272.
- BENABOU, R. (1993): “Workings of a City: Location, Education, and Production,” *The Quarterly Journal of Economics*, 108(3), 619–652.
- BENTOLILA, S., C. MICHELACCI, AND J. SUAREZ (2010): “Social contacts and occupational choice,” *Economica*, 77(305), 20–45.
- BOTTAZZI, L., AND G. PERI (2003): “Innovation and spillovers in regions: Evidence from European patent data,” *European Economic Review*, 47(4), 687–710.
- 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.
- BRAYSHAY, M., M. CLEARY, AND J. SELWOOD (2006): “Power geometries: Social networks and the 1930s multinational corporate elite,” *Geoforum*, 37(6), 986–998.
- (2007): “Social networks and the transnational reach of the corporate class in the early-twentieth century,” *Journal of Historical Geography*, 33(1), 144–167.

- BURT, R. (2003): “Freemasonry and business networking during the Victorian period,” *The Economic History Review*, 56(4), 657–688.
- CABRALES, A., A. CALVÓ-ARMENGOL, AND Y. ZENOU (2011): “Social interactions and spillovers,” *Games and Economic Behavior*, 72(2), 339–360.
- CALVO-ARMENGOL, A., AND M. O. JACKSON (2004): “The effects of social networks on employment and inequality,” *The American Economic Review*, 94(3), 426–454.
- CANEN, N., AND F. TREBBI (2016): “Endogenous Network Formation in Congress,” Discussion paper, National Bureau of Economic Research, No. 22756.
- CASSIMAN, B., AND R. VEUGELERS (2002): “R&D cooperation and spillovers: some empirical evidence from Belgium,” *American Economic Review*, 92(4), 1169–1184.
- CHANDRA, A., AND D. O. STAIGER (2007): “Productivity spillovers in health care: evidence from the treatment of heart attacks,” *Journal of Political Economy*, 115(1), 103–140.
- CICALA, S., R. G. FRYER JR, AND J. L. SPENKUCH (2016): “Self-Selection and Comparative Advantage in Social Interactions,” Mimeo.
- CLEMENS, J., J. D. GOTTLIEB, D. HÉMOUS, AND M. OLSEN (2016): “The Spillover Effects of Top Income Inequality,” Mimeo.
- CURRARINI, S., M. O. JACKSON, AND P. PIN (2009): “An economic model of friendship: Homophily, minorities, and segregation,” *Econometrica*, 77(4), 1003–1045.
- EPPLE, D., AND R. E. ROMANO (1998): “Competition between private and public schools, vouchers, and peer-group effects,” *American Economic Review*, pp. 33–62.
- EVANS, D. S., AND B. JOVANOVIĆ (1989): “An estimated model of entrepreneurial choice under liquidity constraints,” *Journal of Political Economy*, 97(4), 808–827.
- FERREYRA, M. M., AND G. KOSENOK (2015): “Charter School Entry and School Choice: The Case of Washington, DC,” .

- GENIUS, M., P. KOUNDOURI, C. NAUGES, AND V. TZOUVELEKAS (2013): “Information transmission in irrigation technology adoption and diffusion: social learning, extension services, and spatial effects,” *American Journal of Agricultural Economics*, pp. 328–344.
- 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.
- GRANOVETTER, M. (1995): *Getting a job: A study of contacts and careers*. University of Chicago Press.
- GUIO, L., L. PISTAFERRI, AND F. SCHIVARDI (2015): “Learning Entrepreneurship From Other Entrepreneurs?,” Discussion paper, National Bureau of Economic Research, No. 21775.
- GUIO, L., AND F. SCHIVARDI (2011): “What determines entrepreneurial clusters?,” *Journal of the European Economic Association*, 9(1), 61–86.
- GUVENEN, F., F. KARAHAN, S. OZKAN, AND J. SONG (2015): “What do data on millions of US workers reveal about life-cycle earnings risk?,” Discussion paper, National Bureau of Economic Research, No. 20913.
- HAUFLER, A., P.-J. NORBÄCK, AND L. PERSSON (2014): “Entrepreneurial innovations and taxation,” *Journal of Public Economics*, 113, 13–31.
- HECKMAN, J. J., AND G. SEDLACEK (1985): “Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market,” *Journal of political Economy*, 93(6), 1077–1125.
- HECKMAN, J. J., AND G. L. SEDLACEK (1990): “Self-selection and the distribution of hourly wages,” *Journal of Labor Economics*, 8(1, Part 2), 329–363.
- HENREKSON, M., AND M. STENKULA (2010): “Entrepreneurship and public policy,” in *Handbook of Entrepreneurship Research*, pp. 595–637. Springer.

- HOANGA, H., AND B. ANTONCIC (2003): “Network-based research in entrepreneurship. A critical review,” *Journal of Business Venturing*, 18, 165–187.
- HSIEH, C.-T., AND P. J. KLENOW (2009): “Misallocation and manufacturing TFP in China and India,” *The Quarterly Journal of Economics*, 124(4), 1403–1448.
- JACKSON, M. O. (2010): *Social and economic networks*. Princeton University Press.
- JAFFE, A. (1986): “Technological Opportunity and Spillovers of R&D: Evidence from Firms’ Patents, Profits, and Market Value,” *American Economic Review*, 76(5), 984–1001.
- JOVANOVIC, B. (1994): “Firm formation with heterogeneous management and labor skills,” *Small Business Economics*, 6(3), 185–191.
- KHABIBULINA, L., AND A. HEFTI (2015): “Talent Allocation, Financial Intermediation and Growth: Evidence and Theory,” Discussion paper, ZBW-Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft.
- KORPI, T. (2001): “Good friends in bad times? Social networks and job search among the unemployed in Sweden,” *Acta Sociológica*, 44(2), 157–170.
- LIO, M., AND M.-C. LIU (2006): “ICT and agricultural productivity: evidence from cross-country data,” *Agricultural Economics*, 34(3), 221–228.
- LOPEZ-MARTIN, B. (2015): “Informal Sector Misallocation,” *Research Department, Banco de Mexico*.
- LUCAS, R. (1978): “On the size distribution of business firms,” *The Bell Journal of Economics*, pp. 508–523.
- MANDELBROT, B. (1960): “The Pareto-Levy law and the distribution of income,” *International Economic Review*, 1(2), 79–106.
- MARA, J., L. DAVIS, AND S. SCHMIDT (2016): “Social Animal House: The Economic and Academic Consequences of Fraternity Membership,” Discussion paper, Union College.

- MARMAROS, D., AND B. SACERDOTE (2002): “Peer and social networks in job search,” *European Economic Review*, 46(4), 870–879.
- MORETTI, E. (2004): “Workers’ education, spillovers, and productivity: evidence from plant-level production functions,” *American Economic Review*, 94(3), 656–690.
- MURPHY, K. M., A. SHLEIFER, AND R. W. VISHNY (1991): “The allocation of talent: Implications for growth,” *The Quarterly Journal of Economics*, 106(2), 503–530.
- NIX, E. (2015): “Learning Spillovers in the Firm,” Discussion paper, Yale.
- OGUS, A. (2002): “The economic basis of legal culture: Networks and monopolization,” *Oxford journal of Legal Studies*, pp. 419–434.
- PIROLO, L., AND M. PRESUTTI (2007): “Towards a dynamic knowledge-based approach to the innovation process: an empirical investigation on social capital inside an industrial cluster,” *International Journal of Learning and Intellectual Capital*, 4(1-2), 147–173.
- POPOV, S. V., AND D. BERNHARDT (2012): “Fraternities and Labor-Market Outcomes,” *American Economic Journal: Microeconomics*, 4(1), 116–141.
- QUADRINI, V. (2009): “Entrepreneurship in macroeconomics,” *Annals of Finance*, 5(3), 295–311.
- RESTUCCIA, D., AND R. ROGERSON (2013): “Misallocation and productivity,” Mimeo.
- ROTHSCHILD, C., AND F. SCHEUER (2012): “Redistributive taxation in the Roy model,” *The Quarterly Journal of Economics*, p. qjs076.
- ROUTON, P. W., AND J. K. WALKER (2014): “The impact of Greek organization membership on collegiate outcomes: Evidence from a national survey,” *Journal of Behavioral and Experimental Economics*, 49, 63–70.
- ROY, A. D. (1951): “Some thoughts on the distribution of earnings,” *Oxford Economic Papers*, 3(2), 135–146.
- SCHEUER, F. (2014): “Entrepreneurial taxation with endogenous entry,” *American Economic Journal: Economic Policy*, 6(2), 126–163.

- SENE, L. M. (2015): “Connectivity as engine for productivity among small-holder peanut farmers in Senegal,” Discussion paper, International Association of Agricultural Economists.
- SHAKHNOV, K. (2014): “The allocation of talent: Finance versus entrepreneurship,” Discussion paper, European University Institute.
- VEGA-REDONDO, F. (2007): *Complex social networks*. Cambridge University Press.
- WALDINGER, F. (2010): “Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany,” *Journal of Political Economy*, 118(4), 787–831.
- WEST, E., D. N. BARRON, J. DOWSETT, AND J. N. NEWTON (1999): “Hierarchies and cliques in the social networks of health care professionals: implications for the design of dissemination strategies,” *Social Science & Medicine*, 48(5), 633–646.
- YANAGIDA, I. (1992): “The business network: A powerful and challenging business tool,” *Journal of Business Venturing*, 7(5), 341–346.
- YEO, G. B., AND A. NEAL (2004): “A multilevel analysis of effort, practice, and performance: effects; of ability, conscientiousness, and goal orientation,” *Journal of Applied Psychology*, 89(2), 231.

A Appendices

A.1 Proof of Propositions 1 and 2

The FOC for the decentralized problem are

$$k_i = db_i + \frac{a}{2}d^2b_i\sqrt{\frac{s_i}{k_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \quad (23)$$

$$s_i = \frac{a}{2}d^2b_i\sqrt{\frac{k_i}{s_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \quad (24)$$

while the FOC for the social planner simplify to

$$k_i^s = db_i + ad^2 b_i \sqrt{\frac{s_i^s}{k_i^s}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j^s s_j^s}}{N^i} dj \text{ for all } i \quad (25)$$

$$s_i^s = ad^2 b_i \sqrt{\frac{k_i^s}{s_i^s}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j^s s_j^s}}{N^i} dj \text{ for all } i \quad (26)$$

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

We divide (23) by (24) to get

$$\frac{k_i}{s_i} = \frac{d + \frac{a}{2} d^2 \sqrt{\frac{s_i}{k_i}} K(\mathbf{b}, \mathbf{k}, \mathbf{s})}{\frac{a}{2} d^2 \sqrt{\frac{k_i}{s_i}} K(\mathbf{b}, \mathbf{k}, \mathbf{s})} = \frac{\sqrt{\frac{k_i}{s_i}} + \frac{a}{2} d K(\mathbf{b}, \mathbf{k}, \mathbf{s})}{\frac{a}{2} d \frac{k_i}{s_i} K(\mathbf{b}, \mathbf{k}, \mathbf{s})} \quad (27)$$

where bold face letters denote vectors and

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

Rearranging (27) gives

$$d \left(\frac{k_i}{s_i} \right)^2 \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = \sqrt{\frac{k_i}{s_i}} + d \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) \quad (28)$$

from which it is immediate that

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

for some $K(\cdot)$ with a unique solution. To see the uniqueness notice that letting $\sqrt{\frac{k_i}{s_i}} = x_i$ (28) can be written as

$$dx_i^4 \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = x_i + d \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) \quad (29)$$

the left hand side of (29) is a convex function taking the value 0 when $x_i = 0$ and the right hand side it is a linear and takes the positive value $d \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s})$ when $x_i = 0$. Hence there is a single crossing point at the positive orthant.

Hence

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

Thus it is clear we can write

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

An analogous proof establishes that also for the centralized problem

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

It remains to determine the common optimal group parameters.

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

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

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

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

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

$$\begin{aligned}
k &= \frac{d}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)^2} \\
s &= \frac{\frac{a}{2} d^3 \bar{b}^2}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)}
\end{aligned}$$

for the decentralized problem and

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

$$\begin{aligned} k^s &= \frac{d}{1 - \left(ad^2 \bar{b}^2\right)^2} \\ s^s &= \frac{ad^3 \bar{b}^2}{1 - \left(ad^2 \bar{b}^2\right)} \end{aligned}$$

The optimal investments follow immediately from solving this system of linear equations. Assuming $\left(ad^2 \bar{b}^2\right)^2 < 1$ guarantees positive investment levels.

Introducing the optimal investment levels into the utility functions gives us

$$\begin{aligned} k &= \frac{d}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)^2} \\ s &= \frac{\frac{a}{2} d^3 \bar{b}^2}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)} \end{aligned}$$

$$\begin{aligned} u_i(b_i) &= db_i^2 k + ad^2 b_i^2 k s \bar{b}^2 - \frac{1}{2} b_i^2 k^2 - \frac{1}{2} b_i^2 s^2 \\ &= db_i^2 \frac{d}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)^2} + ad^2 b_i^2 \frac{d}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)^2} \frac{\frac{a}{2} d^3 \bar{b}^2}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)} \bar{b}^2 \\ &\quad - \frac{1}{2} b_i^2 \left(\frac{d}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)^2} \right)^2 - \frac{1}{2} b_i^2 \left(\frac{\frac{a}{2} d^3 \bar{b}^2}{1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)} \right)^2 \\ &= \frac{b_i^2 d^2}{2} \left(\frac{\left(1 + \left(\frac{a}{2} d^2 \bar{b}^2\right)^2\right)}{\left(1 - \left(\frac{a}{2} d^2 \bar{b}^2\right)^2\right)^2} \right) \end{aligned}$$

for the decentralized solution and

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

for the centralized solution.

A.2 Proof of Proposition 4

Observe that if rather than the original utility we had

$$u_i^n = d^n b_i^n k_i^n + 2ad^n b_i^n (k_i^n)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n (k_j^n)^{1/2} g_{ij}^n(\mathbf{s}) \right) dj - \frac{1}{2} (k_i^n)^2 - \frac{1}{2} (s_i^n)^2$$

by proposition 1 the individual efforts would be socially efficient. Thus a subsidy equal to

$$ad^n b_i^n (k_i^n)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n (k_j^n)^{1/2} g_{ij}^n(\mathbf{s}) \right) dj$$

would be enough to induce the right investments. But note

$$\begin{aligned} &ad^n b_i^n (k_i^n)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n (k_j^n)^{1/2} g_{ij}^n(\mathbf{s}) \right) dj \\ &= d^n b_i^n k_i^n + ad^n b_i^n (k_i^n)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n (k_j^n)^{1/2} g_{ij}^n(\mathbf{s}) \right) dj - d^n b_i^n k_i^n \end{aligned}$$

and using the fact that

$$b_i = \frac{k_i}{k}$$

then given that

$$y_i = d^n b_i^n k_i^n + ad^n b_i^n (k_i^n)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n (k_j^n)^{1/2} g_{ij}^n(\mathbf{s}) \right) dj$$

we have that the optimal subsidy will be

$$y_i - d \frac{(k_i)^2}{k}$$

A.3 Proof of Proposition 5

Define

$$g_A(C) \equiv \frac{d^{F^2}}{d^{M^2}} \frac{\left(1 + \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)}{\left(1 - \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2} \frac{\left(1 - \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(1 + \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)}$$

so that $g(C, \cdot) = g_A(C) - C^2$. Then, given that we assume that $\sup_C \left(ad^2 \overline{b^2}\right)^2 < 1$

$$g(0, \cdot) > 0$$

Then, note that the assumption $\sup_C \left(ad^2 \overline{b^2}\right)^2 < 1$ means that $\overline{b^2}$ is bounded

above, so the numerator of the function $g_A(C)$ is bounded above by $\left(1 + \left(\frac{a^F}{2} d^{F^2} \sup_C \overline{b^{F^2}}\right)^2\right)$.

Similarly the denominator of $g_A(C)$, is bounded below by $\left(1 - \max_C \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2$.

This means that for all C

$$g_A(C) < \frac{d^{F^2}}{d^{M^2}} \frac{\left(1 + \left(\frac{a^F}{2} d^{F^2} \sup_C \overline{b^{F^2}}\right)^2\right)}{\left(1 - \max_C \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2}$$

which implies that if we define \overline{C}_g as

$$\overline{C}_g \equiv \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{\left(1 + \left(\frac{a^F}{2} d^{F^2} \sup_C \overline{b^{F^2}}\right)^2\right)}{\left(1 - \max_C \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)}}$$

we have that for all $C > \overline{C}_g$

$$g(C, \cdot) < 0$$

and thus by the mean value theorem there exists a value $C^* \in (0, \overline{C}_g)$ such that $g(C^*, \cdot) = 0$.

Similarly, let

$$f_A(C) \equiv \frac{d^{F^2}}{d^{M^2}} \frac{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2}$$

so that $f(C, \cdot) = f_A(C) - C^2$. Then, given that we assume that $\sup_C \left(ad^2 \overline{b^2} \right)^2 < 1$

$$f(0, \cdot) > 0$$

The assumption $\sup_C \left(ad^2 \overline{b^2} \right)^2 < 1$ means that $\overline{b^2}$ is bounded above, so the numerator of the function $f_A(C)$ is bounded above by $\left(1 - \left(a^M d^{M^2} \sup_C \overline{b^{M^2}} \right)^2 \right)$.

Similarly the denominator of $f_A(C)$, is bounded below by $\left(1 - \max_C \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2$. This means that for all C

$$f_A(C) < \frac{d^{F^2}}{d^{M^2}} \frac{\left(1 + \left(a^M d^{M^2} \sup_C \overline{b^{M^2}} \right)^2 \right)}{\left(1 - \max_C \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)}$$

which implies that if we define \overline{C}_f as

$$\overline{C}_f \equiv \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{\left(1 + \left(a^M d^{M^2} \sup_C \overline{b^{M^2}} \right)^2 \right)}{\left(1 - \max_C \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)}}$$

we have that for all $C > \overline{C}_f$

$$f(C, \cdot) < 0$$

and thus by the mean value theorem there exists a value $C^* \in (0, \overline{C}_f)$ such that $f(C^*, \cdot) = 0$.

For stability, note that if $g(C) > 0$ we would have

$$\frac{d^{M^2}}{2} \left(\frac{\left(1 + \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{\left(1 - \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} \right) < \frac{d^{F^2}}{2} \left(\frac{\left(1 + \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}} \right)^2 \right)}{\left(1 - \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2} \right)$$

and thus for an individual with $b_i^M = C b_i^F$ would not be indifferent between group M and F but would prefer to move to group F so that C would present a tendency to increase. This leads us to postulate a natural tatônnement-like adjustment dynamic

$$\frac{\partial C(t)}{\partial t} = R(g(C(t, \cdot)))$$

where $R(\cdot)$ is an increasing function that is positive if and only if $g(C, \cdot)$ is positive. It is then easy to see that in any stable equilibrium C^{**} , $g(C, \cdot)$ has to be decreasing at C^{**} as otherwise, a small increase or decrease from C^{**} will push the dynamics away from the equilibrium. An analogous argument proves the result for $f(C, \cdot)$.

A.4 Proof of Proposition 6

From Proposition 5 we know that

$$\begin{aligned}\frac{\partial f(C, \cdot)}{\partial C} &< 0 \\ \frac{\partial g(C, \cdot)}{\partial C} &< 0\end{aligned}$$

So using the implicit function theorem, we only need to check how the functions $f(\cdot)$ and $g(\cdot)$ vary directly with a^M , a^F , d^M and d^F to calculate how C changes with those underlying parameters.

We start by looking at changes in a^M

$$\begin{aligned}\frac{\partial f(C, \cdot)}{\partial a^M} &= \frac{d^{F^2} - 2a^M \left(d^{M^2} \overline{b^{M^2}}\right)^2}{d^{M^2} \left(1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2} < 0 \\ g(C, \cdot) &= \frac{d^{F^2} \left(4 + \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)}{d^{M^2} \left(4 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2} \frac{\left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2} - C_P^2 = 0 \\ \text{sign} \left(\frac{\partial g(C, \cdot)}{\partial a^M} \right) &= \text{sign} \left(\frac{\frac{-4a^M \left(d^{M^2} \overline{b^{M^2}}\right)^2 \left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right) \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)}{\left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2} - \frac{2a^M \left(d^{M^2} \overline{b^{M^2}}\right)^2 \left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2} \right) \\ &< 0\end{aligned}$$

Hence

$$\begin{aligned}\frac{dC_E}{da^M} &= -\frac{\frac{\partial f(C,\cdot)}{\partial a^M}}{\frac{\partial f(C,\cdot)}{\partial C_E}} < 0 \\ \frac{dC_P}{da^M} &= -\frac{\frac{\partial g(C,\cdot)}{\partial a^M}}{\frac{\partial g(C,\cdot)}{\partial C_P}} < 0\end{aligned}$$

If the synergies of the M -group become more important, C decreases, thus more people join the M -group. We now show that the opposite happens when synergies in the F -group increase.

$$\begin{aligned}\frac{\partial f(C,\cdot)}{\partial a^F} &= \frac{d^{F^2}}{d^{M^2}} \frac{\left(1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right) 2a^F \left(d^{F^2} \overline{b^{F^2}}\right)^2}{\left(1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2} > 0 \\ \text{sign}\left(\frac{\partial g(C,\cdot)}{\partial a^F}\right) &= \text{sign}\left(\frac{2a^F \left(d^{F^2} \overline{b^{F^2}}\right)^2 \left(4 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2}{\left(\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2\right)^2} + \frac{4a^F \left(d^{F^2} \overline{b^{F^2}}\right)^2 \left(4 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right) \left(4 + \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)}{\left(\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2\right)^2}\right) \\ &> 0\end{aligned}$$

Hence

$$\begin{aligned}\frac{dC_E}{da^F} &= -\frac{\frac{\partial f(C,\cdot)}{\partial a^F}}{\frac{\partial f(C,\cdot)}{\partial C_E}} > 0 \\ \frac{dC_P}{da^F} &= -\frac{\frac{\partial g(C,\cdot)}{\partial a^F}}{\frac{\partial g(C,\cdot)}{\partial C_P}} > 0\end{aligned}$$

Now we look at changes in d^M .

$$f(C,\cdot) = \frac{d^{F^2}}{d^{M^2}} \frac{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2} - C_E^2 = 0 \quad (30)$$

$$\frac{\partial f(C, \cdot)}{\partial d^M} = -\frac{2d^{F^2}}{d^{M^3}} \frac{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2} - \frac{d^{F^2}}{d^{M^2}} \frac{4d^{M^3} \left(a^M \overline{b^{M^2}}\right)^2}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2} < 0$$

Hence

$$\frac{dC_E}{d^M} = -\frac{\frac{\partial f(C, \cdot)}{\partial d^M}}{\frac{\partial f(C, \cdot)}{\partial C_E}} < 0$$

Similarly

$$\begin{aligned} \text{sign} \left(\frac{\partial g(C, \cdot)}{\partial d^M} \right) &= \text{sign} \left(\frac{\frac{-4d^{M^3} \left(a^M \overline{b^{M^2}}\right)^2 \left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right) \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)}{\left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2} - \frac{4d^{M^3} \left(a^M \overline{b^{M^2}}\right)^2 \left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2} \right) \\ &< 0 \end{aligned}$$

Therefore

$$\frac{\partial C_P}{\partial d^M} = -\frac{\frac{\partial g(C, \cdot)}{\partial d^M}}{\frac{\partial g(C, \cdot)}{\partial C_E}} < 0$$

If d^M increases fewer people join the F -group.

Finally we want to understand how the dividing line is affected by changes in d^F .

$$\frac{\partial f(C, \cdot)}{\partial d^F} = \left(\frac{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2}{d^{M^2}} \right) \left(\frac{2d^F \left(1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right) + 2a^F \left(d^{F^2} \overline{b^{F^2}}\right)^2 d^{F^2}}{\left(1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2} \right) > 0$$

Therefore

$$\frac{\partial C_E}{\partial d^F} = -\frac{\frac{\partial g(C, \cdot)}{\partial d^F}}{\frac{\partial g(C, \cdot)}{\partial C_E}} > 0$$

Similarly

$$\begin{aligned}
& \text{sign} \left(\frac{\partial g(C, \cdot)}{\partial d^F} \right) \\
&= \frac{d^{F^2} \left(4 + \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)}{\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2} \\
&= \text{sign} \left(\begin{aligned} & \frac{\left(8d^F + 6d^{F^5} \left(a^F \overline{b^{F^2}} \right)^2 \right) \left(4 - \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2}{\left(\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2 \right)^2} \\ & + \frac{8d^{F^3} \left(a^F \overline{b^{F^2}} \right)^2 \left(4 - \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right) \left(4d^{F^2} + d^{F^2} \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)}{\left(\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2 \right)^2} \end{aligned} \right) \\
&> 0
\end{aligned}$$

Therefore

$$\frac{\partial C_P}{\partial d^F} = - \frac{\frac{\partial g(C, \cdot)}{\partial d^F}}{\frac{\partial g(C, \cdot)}{\partial C_E}} > 0$$

A.5 Proof of Proposition 7

We first characterize the optimal choices under a linear tax/subsidy on output.

The FOC for the decentralized problem are

$$k_i = dtb_i + \frac{a}{2} d^2 tb_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \text{ for all } i \quad (31)$$

$$s_i = \frac{a}{2} d^2 tb_i \sqrt{\frac{k_i}{s_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \text{ for all } i \quad (32)$$

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

We divide (31) by (32) to get

$$\frac{k_i}{s_i} = \frac{d + \frac{a}{2} d^2 \sqrt{\frac{s_i}{k_i}} K(\mathbf{b}, \mathbf{k}, \mathbf{s})}{\frac{a}{2} d^2 \sqrt{\frac{k_i}{s_i}} K(\mathbf{b}, \mathbf{k}, \mathbf{s})} = \frac{\sqrt{\frac{k_i}{s_i}} + \frac{a}{2} d K(\mathbf{b}, \mathbf{k}, \mathbf{s})}{\frac{a}{2} d \frac{k_i}{s_i} K(\mathbf{b}, \mathbf{k}, \mathbf{s})} \quad (33)$$

where bold face letters denote vectors and

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

Rearranging (33) gives

$$d \left(\frac{k_i}{s_i} \right)^2 \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = \sqrt{\frac{k_i}{s_i}} + d \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) \quad (34)$$

from which it is immediate that

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

for some $K(\cdot)$ with a unique solution. To see the uniqueness notice that letting $\sqrt{\frac{k_i}{s_i}} = x_i$ (34) can be written as

$$dx_i^4 \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) = x_i + d \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s}) \quad (35)$$

the left hand side of (35) is a convex function taking the value 0 when $x_i = 0$ and the right hand side it is a linear and takes the positive value $d \frac{a}{2} K(\mathbf{b}, \mathbf{k}, \mathbf{s})$ when $x_i = 0$. Hence there is a single crossing point at the positive orthant.

Hence

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

Thus it is clear we can write

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

It remains to determine the common optimal group parameters.

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

$\bar{b}^2\sqrt{ks}$ for the individual problem where

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

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

$$\begin{aligned} k &= dt + d^2t \frac{a}{2} \sqrt{\frac{s}{k} \bar{b}^2} \sqrt{ks} = dt + \frac{a}{2} d^2t \bar{b}^2 s \\ s &= \frac{a}{2} d^2t \sqrt{\frac{k}{s} \bar{b}^2} \sqrt{ks} = \frac{a}{2} d^2t \bar{b}^2 k \end{aligned}$$

$$\begin{aligned} k &= \frac{dt}{1 - \left(\frac{a}{2} d^2t \bar{b}^2\right)^2} \\ s &= \frac{\frac{a}{2} d^3t^2 \bar{b}^2}{1 - \left(\frac{a}{2} d^2t \bar{b}^2\right)^2} \end{aligned}$$

The optimal investments follow immediately from solving this system of linear equations. Assuming $\left(ad^2t\bar{b}^2\right)^2 < 1$ guarantees positive investment levels.

Introducing the optimal investment levels into the utility functions gives us

$$\begin{aligned} u_i(b_i) &= db_i^2tk + ad^2b_i^2tk s \bar{b}^2 - \frac{1}{2}b_i^2k^2 - \frac{1}{2}b_i^2s^2 \\ &= b_i^2 \left(dtk + ad^2tk s \bar{b}^2 - \frac{1}{2}k^2 - \frac{1}{2}s^2 \right) \\ &= \frac{b_i^2 d^2 t^2}{2} \left(\frac{2 \left(1 - \left(\frac{a}{2} d^2 t \bar{b}^2 \right)^2 \right) + 4 \left(\frac{a}{2} d^2 t \bar{b}^2 \right)^2 - 1 - \left(\frac{a}{2} d^2 t \bar{b}^2 \right)^2}{\left(1 - \left(\frac{a}{2} d^2 t \bar{b}^2 \right)^2 \right)^2} \right) \end{aligned}$$

so that

$$u_i(b_i) = \frac{b_i^2 d^2 t^2}{2} \left(\frac{1 + \left(\frac{a}{2} d^2 t \bar{b}^2 \right)^2}{\left(1 - \left(\frac{a}{2} d^2 t \bar{b}^2 \right)^2 \right)^2} \right) \quad (36)$$

for the decentralized solution. We similarly obtain that for the centralized solution

$$u_i(b_i) = \frac{b_i^2 d^2 t^2}{2} \frac{1}{\left(1 - \left(a d^2 t \bar{b}^2 \right)^2 \right)^2} \quad (37)$$

From expression (36) we get that the equation that defines C_P implicitly is

$$g(C, t, \cdot) = \frac{d^{F^2} t^{F^2}}{d^{M^2} t^{M^2}} \frac{\left(4 + \left(a^F d^{F^2} t^F \bar{b}^{F^2} \right)^2 \right) \left(4 - \left(a^M d^{M^2} t^M \bar{b}^{M^2} \right)^2 \right)^2}{\left(4 - \left(a^F d^{F^2} t^F \bar{b}^{F^2} \right)^2 \right)^2 \left(4 + \left(a^M d^{M^2} t^M \bar{b}^{M^2} \right)^2 \right)} - C_P^2 = 0 \quad (38)$$

and from expression (37) we get that the equation that defines C_P implicitly is

$$f(C, t, \cdot) = \frac{d^{F^2} t^{F^2}}{d^{M^2} t^{M^2}} \frac{1 - \left(a^M d^{M^2} t^M \bar{b}^{M^2} \right)^2}{1 - \left(a^F d^{F^2} t^F \bar{b}^{F^2} \right)^2} - C_E^2 = 0 \quad (39)$$

from expression (39) and (38) we get that $\lim_{\frac{t^F}{t^M} \rightarrow 0} C_E = \lim_{\frac{t^F}{t^M} \rightarrow 0} C_P = 0$ and $\lim_{\frac{t^M}{t^F} \rightarrow 0} C_E = \lim_{\frac{t^M}{t^F} \rightarrow 0} C_P = \infty$. This, plus continuity of C_E and C_P as a function of t^F , t^M establishes that one can obtain any value of C_E and C_P between 0 and ∞ by appropriately varying $\frac{t^F}{t^M}$.

A.6 Proof of Proposition 8

Existence follows from Proposition 5.

Let first $B^M \geq C B^F$. 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} .

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

So

$$\overline{b^{F^2}} = \frac{B^{F^2}}{2} \quad (40)$$

$$\overline{b^{M^2}} = E\left(b_i^{M^2} \mid b_i^M > Cb_i^F\right) = \frac{\int_0^{B^F} \int_{Cb_F}^{B^M} b^{M^2} db^M db^F}{\int_0^{B^F} \int_{Cb_F}^{B^M} db^M db^F} = \frac{1}{6} \frac{4B^{M^3} - C^3 B^{F^3}}{2B^M - C B^F} \quad (41)$$

So

$$\overline{b^{M^2}} = \frac{1}{6} \left(B^{F^2} C^2 + 2B^F B^M C + 4B^{M^2} - \frac{4B^{M^3}}{2B^M - C B^F} \right) \quad (42)$$

Observe that

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

Since $2C B^{F^2} + 2B^F B^M$ is linear and $\frac{4B^{M^3} B^F}{(2B^M - C B^F)^2}$ is convex then $\frac{\partial \overline{b^{M^2}}}{\partial C} > 0$ provided it is positive for $C = 0$ and for $B^M = C B^F$. 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 $B^M = C B^F$. This means that

$$\frac{\partial \overline{b^{M^2}}}{\partial C} \geq 0 \text{ for } B^M \geq C B^F \text{ with strict equality when } B^M = C B^F \quad (44)$$

a result we will use later.

Using the expressions derived for (40) and (41) we can calculate C_P and C_E . In the case of C_P the expression (16) becomes

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

Rearranging we get

$$\frac{\left(4 - \left(a^F d^{F^2} \frac{B^{F^2}}{2}\right)^2\right)^2}{\left(4 + \left(a^F d^{F^2} \frac{B^{F^2}}{2}\right)^2\right)} C_P^2 = \frac{d^{F^2}}{d^{M^2}} \frac{\left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)} \quad (45)$$

We define

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

and check how it changes with the dividing line C .

$$\begin{aligned} \frac{\partial F(C)}{\partial C} &= \frac{d^{F^2} - 4\overline{b^{M^2}} \left(a_M d^{M^2} \right)^2 \left(4 - \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right) \left(4 + \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right) \frac{\partial \overline{b^{M^2}}}{\partial C}}{d^{M^2} \left(4 + \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} \\ &\quad - \frac{d^{F^2} 2\overline{b^{M^2}} \left(a_M d^{M^2} \right)^2 \left(4 - \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2 \frac{\partial \overline{b^{M^2}}}{\partial C}}{d^{M^2} \left(4 + \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} \\ &= -2 \frac{d^{F^2} \overline{b^{M^2}} \left(a_M d^{M^2} \right)^2 \left(4 - \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{d^{M^2}} \frac{\partial \overline{b^{M^2}}}{\partial C} \frac{\left(12 + \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{\left(4 + \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} \\ &< 0 \end{aligned}$$

where the last inequality is true because we know that $\left(4 - \left(a_M d^{M^2} \overline{b^{M^2}} \right)^2 \right) > 0$ by Assumption 1, and because $\partial \overline{b^{M^2}} / \partial C > 0$ when $B^M > CB^F$ as shown in expression 44. Hence the LHS of (45) is increasing in C while the RHS is decreasing, so equilibrium is unique.

A symmetric argument shows that $\partial F(C) / \partial C < 0$ when $B^M < CB^F$, which establishes the result.

A.7 Proof of Proposition 9

We assume now that returns b follow a Pareto distribution with shape parameter α_i for $i \in \{F, M\}$

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

We will derive the results under the assumption that the C that defines the dividing line $b_i^M = Cb_i^F$ is such that $C \geq 1$.²⁹ Existence follows from

²⁹If $C < 1$, the same results hold with the names of the networks interchanged.

Proposition 5.

We now calculate $\overline{b^{F^2}}$ and $\overline{b^{M^2}}$ for $C \geq 1$.

$$\begin{aligned}\overline{b^{F^2}} &= E\left(b_i^{F^2} \mid b_i^M < C b_i^F\right) = \frac{\int_1^\infty \int_1^{C b_F} b^{F^2} \frac{\alpha_F}{b^{F\alpha_F+1}} \frac{\alpha_M}{b^{M\alpha_M+1}} db^M db^F}{\int_1^\infty \int_1^{C b_F} \frac{\alpha_F}{b^{F\alpha_F+1}} \frac{\alpha_M}{b^{M\alpha_M+1}} db^M db^F} \\ &= \frac{\alpha_F}{\alpha_F - 2} \frac{((\alpha_F + \alpha_M - 2) C^{\alpha_M} - (\alpha_F - 2))}{\left(C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}\right) (\alpha_F + \alpha_M - 2)}\end{aligned}$$

which can be rewritten as

$$\overline{b^{F^2}} = (\alpha_F + \alpha_M) \frac{\alpha_F}{(\alpha_F - 2)} \frac{(\alpha_F + \alpha_M - 2) C^{\alpha_M} - (\alpha_F - 2)}{((\alpha_F + \alpha_M) C^{\alpha_M} - \alpha_F) (\alpha_F + \alpha_M - 2)} \quad (46)$$

while

$$\begin{aligned}\overline{b^{M^2}} &= E\left(b_i^{M^2} \mid b_i^M > C b_i^F\right) \\ &= \frac{\frac{\alpha_M \alpha_F}{(\alpha_M - 2) C^{\alpha_M - 2}} \frac{1}{(\alpha_F + \alpha_M - 2)}}{\frac{\alpha_F}{C^{\alpha_M}} \frac{1}{\alpha_F + \alpha_M}}\end{aligned}$$

which simplifies to

$$\overline{b^{M^2}} = \frac{\alpha_M}{(\alpha_M - 2)} \frac{\alpha_F + \alpha_M}{(\alpha_F + \alpha_M - 2)} C^2 \quad (47)$$

$\overline{b^{M^2}}$ is obviously increasing in C . We now show that $\overline{b^{F^2}}$ is decreasing in C

$$\frac{\partial \overline{b^{F^2}}(C)}{\partial C} = -\frac{(\alpha_F + \alpha_M)}{(\alpha_F + \alpha_M - 2)} \frac{\alpha_F}{(\alpha_F - 2)} \left(\frac{2\alpha_M^2 C^{\alpha_M - 1}}{((\alpha_F + \alpha_M) C^{\alpha_M} - \alpha_F)^2} \right) < 0 \quad (48)$$

We first prove uniqueness of C_E defined by

$$C_E = \frac{d^F}{d^M} \sqrt{\frac{1 - a^M d^M \overline{b^{M^2}}^2}{1 - a^F d^F \overline{b^{F^2}}^2}} \quad (49)$$

Note that the LHS of (49) is increasing in C_E so all we need to show is the RHS is decreasing in C_E so that a unique equilibrium exists. Clearly the numerator of the RHS is decreasing in C_E because $\overline{b^{M^2}}$ is increasing in C_E . Since $\overline{b^{F^2}}$ is decreasing in C_E , the denominator of the RHS is increasing in C_E . And thus the result follows.

We now prove uniqueness of C_P which is defined by

$$C_P = \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{\left(1 + \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right) \left(1 - \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(1 - \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2 \left(1 + \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)}} \quad (50)$$

Again note that the LHS of (50) is increasing in C_P so all we need to show is the RHS is decreasing in C_P so that a unique equilibrium exists. It is again easy to see that $\left(1 - \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)^2 / \left(1 + \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}}\right)^2\right)$ is decreasing in C_P because $\overline{b^{M^2}}$ is increasing in C_P . Also, since we showed in (48) that $\overline{b^{F^2}}$ is decreasing in C_P then $\left(1 + \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right) / \left(1 - \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}}\right)^2\right)^2$ is decreasing in C_P . As a result RHS of (50) is decreasing in C_P and the result follows.

A.8 Proof of Proposition 10

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

$$\begin{aligned} w(C) &= \frac{1}{B^F B^M} \left[\int_0^{B^F} \int_0^{C b_i^F} \frac{b_i^{F^2}}{2} \left(\frac{d^{F^2}}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2} \right) db_i^M db_i^F \right. \\ &\quad \left. + \int_0^{B^F} \int_{C b_i^F}^{B^M} \frac{b_i^{M^2}}{2} \left(\frac{d^{M^2}}{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2} \right) db_i^M db_i^F \right] \\ \frac{\partial w(C)}{\partial C} &= \frac{1}{B^F B^M} \left[\int_0^{B^F} \frac{b_i^{F^3}}{2} \left(\left(\frac{d^{F^2}}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2} \right) - C^2 \left(\frac{d^{M^2}}{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2} \right) \right) db_i^F \right] \\ &\quad + \int_0^{B^F} \int_{C b_i^F}^{B^M} \frac{\partial \frac{b_i^{M^2}}{2} \left(\frac{d^{M^2}}{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2} \right)}{\partial C} db_i^M db_i^F \end{aligned} \quad (51)$$

We already established when proving Proposition 8) that for all C

$$\int_0^{B^F} \int_{C b_i^F}^{B^M} \frac{\partial \frac{b_i^{M^2}}{2} \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \bar{b}^{M^2})^2} \right)}{\partial C} db_i^M db_i^F > 0 \quad (52)$$

by showing the integrand is positive as $\frac{\partial b^{M^2}}{\partial C} > 0$ (44). Letting

$$H(C) = \frac{1}{B^F B^M} \left[\int_0^{B^F} \frac{b_i^{F^3}}{2} \left(\left(\frac{d^{F^2}}{1 - (a^F d^{F^2} \bar{b}^{F^2})^2} \right) - C^2 \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \bar{b}^{M^2})^2} \right) \right) db_i^F \right]$$

$$\frac{\partial H(C)}{\partial C} = \frac{1}{B^F B^M} \left[\int_0^{B^F} \frac{b_i^{F^3}}{2} \left(-2C \left(\frac{d^{M^2}}{1 - a^M d^{M^2} \bar{b}^{M^2}} \right) - C^2 \frac{\partial \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \bar{b}^{M^2})^2} \right)}{\partial C} \right) db_i^F \right]$$

and again by the proof of Proposition 8 we know that

$$\int_0^{B^F} \frac{b_i^{F^3}}{2} \left(\frac{\partial \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \bar{b}^{M^2})^2} \right)}{\partial C} \right) db_i^F > 0$$

so

$$\frac{\partial H(C)}{\partial C} < 0. \quad (53)$$

It is also easy to see that for $C_E = \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{1 - (a^M d^{M^2} \bar{b}^{M^2})^2}{1 - (a^F d^{F^2} \bar{b}^{F^2})^2}}$

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

and hence by (53) we have that $H(C) > 0$ for $C < C_E$ and the result follows for C_E .

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

that social welfare is given by

$$\begin{aligned}
u_i(b_i) &= \frac{b_i^2 d^2}{2} \left(\frac{\left(4 + \left(a d^2 \overline{b^2} \right)^2 \right)}{\left(4 - \left(a d^2 \overline{b^2} \right)^2 \right)^2} \right) \\
w(C) &= \frac{1}{B^F B^M} \left[\int_0^{B^F} \int_0^{C b_i^F} 2b_i^{F^2} \frac{d^{F^2} \left(4 + \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)}{\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2} db_i^M db_i^F \right. \\
&\quad \left. + \int_0^{B^F} \int_{C b_i^F}^{B^M} 2b_i^{M^2} \frac{d^{M^2} \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{\left(4 - \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} db_i^M db_i^F \right] \\
\frac{\partial w(C)}{\partial C} &= \frac{1}{B^F B^M} \left[\int_0^{B^F} 2b_i^{F^3} \left(\frac{d^{F^2} \left(4 + \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)}{\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2} - C^2 \frac{d^{M^2} \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{\left(4 - \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} \right) db_i^F \right] \\
&\quad + \int_0^{B^F} \int_{C b_i^F}^{B^M} \frac{\partial \left(2b_i^{M^2} \frac{d^{M^2} \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{\left(4 - \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} \right)}{\partial C} db_i^M db_i^F
\end{aligned}$$

We already established when proving Proposition 8) that for all C

$$\int_0^{B^F} \int_{C b_i^F}^{B^M} \frac{\partial \left(2b_i^{M^2} \frac{d^{M^2} \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)}{\left(4 - \left(a^M d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2} \right)}{\partial C} db_i^M db_i^F > 0 \quad (54)$$

by showing the integrand is positive as $\frac{\partial \overline{b^{M^2}}}{\partial C} > 0$ (44). Letting

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

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

and again by the proof of Proposition 8 we know that

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

So

$$\frac{\partial H_P(C)}{\partial C} < 0. \quad (55)$$

It is also easy to see that for $C_P = \sqrt{\frac{\frac{d^{F^2}}{d^{M^2}} \left(1 + \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}} \right)^2 \right) \left(1 - \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2}{\left(1 - \left(\frac{a^F}{2} d^{F^2} \overline{b^{F^2}} \right)^2 \right)^2 \left(1 + \left(\frac{a^M}{2} d^{M^2} \overline{b^{M^2}} \right)^2 \right)^2}}$

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

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

A.9 Proof of Proposition 11

We will prove Proposition for $\alpha_F = \alpha_M$. We will first show that

Lemma 4. $C_E > 1 \Leftrightarrow a^{M^2} d^{M^2} < a^{F^2} d^{F^2}$

Note also that if $a^{M^2} d^{M^2} = a^{F^2} d^{F^2}$ the solution of (49) is at $C^E = 1$. An increase of $a^{M^2} d^{M^2}$ with respect to $a^{F^2} d^{F^2}$ displaces the RHS to the left so that the new equilibrium entails $C^E < 1$.

We will now show that for $C_E > 1$ there might be too few $\frac{\partial w(C)}{\partial C} \Big|_{C=C_E} > 0$ or too many people $\frac{\partial w(C)}{\partial C} \Big|_{C=C_E} < 0$ in the F group compared to the social

optimum.³⁰ The F group will be underpopulated if and only if

$$\frac{a^{M^2} d^{M^2}}{a^{F^2} d^{F^2}} > \frac{((2\alpha - 2) C^\alpha - (\alpha - 2))^2 C^\alpha}{\alpha^2 (2C^\alpha - 1)^3 C^2} \quad (56)$$

We will check how a decentralized group choice deviates from the efficient group 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 socializing effort.

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

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

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

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

³⁰The assumption $C_E > 1$ is without loss of generality subject to relabeling. It implies that synergies are bigger in occupation F .

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

Therefore

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

and

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

By Lemma 4 since $C_E > 1 \Leftrightarrow a^{M^2}d^{M^2} < a^{F^2}d^{F^2}$, hence $\frac{a^{M^2}d^{M^2}}{a^{F^2}d^{F^2}} < 1$.

We will now show that

$$1 > \frac{((2\alpha-2)C^\alpha - (\alpha-2))^2 C^\alpha}{\alpha^2 (2C^\alpha-1)^3 C^2} = \frac{((\alpha-1)(2C^\alpha-1)+1)^2 C^\alpha}{\alpha^2 (2C^\alpha-1)^3 C^2} \quad (57)$$

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

$$\frac{((\alpha-1)(2C^\alpha-1)+1)^2 C^\alpha}{\alpha^2 (2C^\alpha-1)^3 C^2} < \frac{C^\alpha}{(2C^\alpha-1)C^2} < \frac{1}{C} < 1 \quad (58)$$

where the last two inequalities hold since $C > 1$, noting that in that case $2C^\alpha-1 > C^\alpha$. Thus equation (58) establishes (57).

The next two lemmas establish that overpopulation can occur in both sectors and depends on the underlying parameters. Lemma 5 shows the existence of parameter values that $\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} < 0$ while Lemma 6 shows the existence of parameter values that $\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} > 0$.

Lemma 5. *Let $\frac{a^{M^2}d^{M^2}}{a^{F^2}d^{F^2}} = r < 1$. For a fixed α and r there exists an a^{F^2} low*

enough that

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

Since

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

we have that

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

thus

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

Lemma 6. Let $\frac{a^{M^2} d^{M^2}}{a^{F^2} d^{F^2}} = r < 1$. For a fixed a^{F^2} and r such that C_E exists, there is an α high enough that

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

For a bounded C_E

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

Hence

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

and thus

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

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

$$\frac{-\left(1 - a^{F^2} d^{F^2}\right) + \sqrt{(1 - a^{F^2} d^{F^2})^2 + 4ra^{F^2} d^{F^2}}}{2ra^{F^2}} > \frac{1}{2r} \quad (59)$$

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

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

Proposition 11 immediately follows from these Lemmas.

A.10 Proof of lemma 2

Under congestion, the welfare of the group F remains unchanged while the welfare of group M is given by

$$W_E^M(C) = \frac{1}{B^F B^M} \int_0^{B^F} \int_{Cb_i^F}^{B^M} f(C, b_i^M) \frac{b_i^{M^2}}{2} \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \bar{b}^{M^2})^2} \right) db_i^M db_i^F \quad (60)$$

when the government induces efficient socializing and productive efforts within a group and by

$$W_P^M(C) = \frac{1}{B^F B^M} \int_0^{B^F} \int_{Cb_i^F}^{B^M} f(C, b_i^M) 2b_i^{M^2} \frac{d^{M^2} \left(4 + (a^M d^{M^2} \bar{b}^{M^2})^2 \right)}{\left(4 - (a^M d^{M^2} \bar{b}^{M^2})^2 \right)^2} db_i^M db_i^F \quad (61)$$

These expressions (60) and (61) can be decomposed in the welfare of those member of M below $b_i^M < CB^F$ for whom congestion does not matter and those above $b_i^M \geq CB^F$ for whom congestion impinges. For the case where the government induces efficient efforts

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

The second line captures welfare of those for whom congestion matters. After

some calculations this second line becomes

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

which decomposes the welfare of people beyond the $b_i^M \geq C_E^* B^F$ boundary in two parts. First the welfare for types exactly at the boundary is $\frac{(C B^F)^2}{2} \left(\frac{d^{M^2}}{1 - (a^M d^{M^2} \overline{b^{M^2}})^2} \right)$ times the fraction of people in that area is $\frac{(B^M - C B^F) B^F}{B^F B^M}$, which gives the first line. And the second line is the surplus welfare for those types, in addition to what the boundary types get, and on which the congestion impinges. Using (21) we can now write total welfare when the government induces efficient socializing and productive efforts as

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

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

A.11 Proof of lemma 3

We first show that it suffices to have $1 - \left(a^M d^{M^2} \frac{B^{M^2}}{3} \right)^2 \approx 0$ (sufficiently small), to have $C_E = \varepsilon \approx 0$ (very small). Observe that when $C_E \approx 0$

$$\overline{b^{M^2}} = \frac{1}{6} \frac{4 B^{M^3} - C^3 B^{F^3}}{2 B^M - C B^F} \simeq \frac{B^{M^2}}{3} \quad (62)$$

Hence

$$C_E = \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{1 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2}{1 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2}} \quad (63)$$

in which case C_E small by having

$$C_E^2 \simeq \frac{d^{F^2}}{d^{M^2}} \frac{1 - \left(a^M d^{M^2} \frac{B^{M^2}}{3}\right)^2}{1 - \left(a^F d^{F^2} \frac{B^{F^4}}{4}\right)^2} \quad (64)$$

and thus it suffices to have

$$1 - \left(a^M d^{M^2} \frac{B^{M^2}}{3}\right)^2 \simeq \varepsilon^2 \quad (65)$$

small to have C_E small.

Next we show that $C_P > 0$ for these parameter values (65). Recall that

$$C_P = \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{\left(4 + \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right) \left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2 \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)}} \quad (66)$$

Assume for contradiction that $C_P = \varepsilon \approx 0$. Since $\overline{b^{M^2}} = \frac{1}{6} \frac{4B^{M^3} - C^3 B^{F^3}}{2B^M - C B^F}$, in this case $\overline{b^{M^2}}^2 = \left(\frac{B^{M^2}}{3}\right)^2$. But using (65) we get

$$4 - \left(a^M d^{M^2} \frac{B^{M^2}}{3}\right)^2 = 3 + \underbrace{1 - \left(a^M d^{M^2} \frac{B^{M^2}}{3}\right)^2}_{\approx \varepsilon^2} \simeq 3$$

$$4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2 = 5 - \underbrace{\left(1 - \left(a^M d^{M^2} \frac{B^{M^2}}{3}\right)^2\right)}_{\approx \varepsilon^2} \simeq 5$$

and so

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

$$C_P = \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{\left(4 + \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right) \left(4 - \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}{\left(4 - \left(a^F d^{F^2} \overline{b^{F^2}}\right)^2\right)^2 \left(4 + \left(a^M d^{M^2} \overline{b^{M^2}}\right)^2\right)^2}} \simeq \sqrt{\frac{d^{F^2}}{d^{M^2}} \frac{\left(4 + a^{F^2} \frac{B^{F^4}}{4}\right) \frac{9}{5}}{\left(4 - a^{F^2} \frac{B^{F^4}}{4}\right)^2 \frac{9}{5}}} > 0$$
(67)

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

Observe that rewriting (18) as

$$\frac{d^{M^2}}{\left(1 - \left(a^M d^{M^2} \frac{B^{M^2}}{3}\right)^2\right)^2} C_E^2 \simeq \frac{d^{F^2}}{1 - \left(a^F d^{F^2} \frac{B^{F^2}}{2}\right)^2}$$
(68)

we can express welfare when optimal socializing and productive efforts are induced in the group (Lemma 2) when (65) holds and hence C_E is very small as

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

$$+ \left(1 - C \frac{B^F}{B^M}\right) \frac{(B^F)^2}{2} \frac{d^{F^2}}{1 - \left(a^F d^{F^2} \frac{B^{F^2}}{2}\right)^2} + (1 - v(C)) G_E(C)$$

which reduces to

$$w_E(C_E \approx 0) \approx \frac{(B^F)^2}{2} \frac{d^{F^2}}{1 - \left(a^F d^{F^2} \frac{B^{F^2}}{2}\right)^2}.$$
(69)

To calculate welfare without any government intervention $w_P(C_P)$ recall that

$C_P \neq 0$ and hence $v(C_P) = 0$. Hence

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

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

Note that

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

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

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

We know by (65) that $\left(a^M d^{M^2} \frac{B^{M^2}}{3}\right)^2 \simeq 1$, so then

$$w(C_P) \geq \frac{d^{M^2}}{B^F B^M} \left[\frac{20}{9} \left(\frac{B^{M^3} B^F}{6} - \frac{B^{F^4}}{24} \right) \right]$$

$$\text{Using (69)} w_E(C_E \approx 0) \approx \frac{(B^F)^2}{2} \frac{d^{F^2}}{1 - \left(a^F d^{F^2} \frac{B^{F^2}}{2}\right)^2}$$

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

which is negative for B^M big enough.

A.12 Proof of Proposition 14

Suppose we normalize by the expected second moment $\frac{\alpha_M}{(\alpha_M - 2)}$, then

$$\overline{b^{M^2}}_{NORM} = \frac{\alpha_F + \alpha_M}{(\alpha_F + \alpha_M - 2)} C^2 = \left(1 + \frac{2}{\alpha_F + \alpha_M - 2} \right) C^2$$

Clearly, this is decreasing in α_F and α_M .

Suppose we normalize $\overline{b^{F^2}}$ by the expected second moment $\frac{\alpha_F}{(\alpha_F - 2)}$, then

$$\overline{b^{F^2}}_{NORM} = \frac{\left(1 - \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \frac{1}{C^{\alpha_M}} \right)}{\left(1 - \frac{\alpha_F}{\alpha_F + \alpha_M} \frac{1}{C^{\alpha_M}} \right)}$$

$$\begin{aligned}
\frac{\partial \overline{b^{F^2}}_{NORM}}{\partial \alpha_F} &= \frac{1}{\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} - 1} \left(\frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M - 2)} - \frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M - 2)^2} \right) \\
&\quad - \frac{1}{\left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} - 1 \right)^2} \left(\frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M)} - \frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M)^2} \right) \\
&\quad \times \left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} - 1 \right) \\
&= -\frac{1}{1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M}} \left(\frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M - 2)^2} \right) \\
&\quad + \frac{1}{\left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} - 1 \right)^2} \left(\frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M)^2} \right) \left(1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \right)
\end{aligned}$$

we want to show that

$$\frac{\partial \overline{b^{F^2}}_{NORM}}{\partial \alpha_F} < 0$$

this is true if and only

$$\frac{1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2}}{1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M}} < \frac{(\alpha_F + \alpha_M)^2}{(\alpha_F + \alpha_M - 2)^2}$$

or

$$\left(1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \right) (\alpha_F + \alpha_M - 2)^2 < \left(1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} \right) (\alpha_F + \alpha_M)^2 \quad (70)$$

We will now show that

$$G(\alpha_F) \equiv (\alpha_F + \alpha_M)^2 \left(1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} \right)$$

is increasing in α_F and given that $\partial \overline{b^{F^2}}_{NORM} / \partial \alpha_F < 0$ is equivalent to expression (70) the result follows. Then

$$\begin{aligned}
\frac{\partial G(\alpha_F)}{\partial \alpha_F} &= - \left(\frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M)} - \frac{1}{C^{\alpha_M} (\alpha_F + \alpha_M)^2} \right) (\alpha_F + \alpha_M)^2 \\
&\quad - \left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} - 1 \right) (2\alpha_F + 2\alpha_M) \\
&= \frac{1}{C^{\alpha_M}} (2C^{\alpha_M} \alpha_F + 2C^{\alpha_M} \alpha_M - 2\alpha_F - \alpha_M) > 0
\end{aligned}$$

Now note that

$$\begin{aligned} \frac{\partial \overline{b^{F^2}}_{NORM}}{\partial \alpha_M} &= \frac{\left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} - 1 \right)}{\left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} - 1 \right)^2} \left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{(\alpha_F + \alpha_M)^2} + \frac{1}{C^{\alpha_M}} \alpha_F \frac{\ln C}{\alpha_F + \alpha_M} \right) \\ &\quad - \frac{1}{\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} - 1} \left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{(\alpha_F + \alpha_M - 2)^2} + \frac{1}{C^{\alpha_M}} (\ln C) \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \right) \end{aligned}$$

then

$$\frac{\partial \overline{b^{F^2}}_{NORM}}{\partial \alpha_M} < 0$$

requires

$$\begin{aligned} &\frac{1}{1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M}} \left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{(\alpha_F + \alpha_M - 2)^2} + \frac{1}{C^{\alpha_M}} (\ln C) \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \right) \\ &< \frac{\left(1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \right)}{\left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} - 1 \right)} \left(\frac{1}{C^{\alpha_M}} \frac{\alpha_F}{(\alpha_F + \alpha_M)^2} + \frac{1}{C^{\alpha_M}} \alpha_F \frac{\ln C}{\alpha_F + \alpha_M} \right) \end{aligned}$$

which is equivalent to

$$\begin{aligned} &\frac{1}{\left(1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \right)} \left(\frac{\alpha_F - 2}{(\alpha_F + \alpha_M - 2)^2} + (\ln C) \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2} \right) \\ &< \frac{1}{1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M}} \left(\frac{\alpha_F}{(\alpha_F + \alpha_M)^2} + \alpha_F \frac{\ln C}{\alpha_F + \alpha_M} \right) \end{aligned}$$

Clearly this is true as $\alpha_F / (\alpha_F + \alpha_M)^2$, $\alpha_F / (\alpha_F + \alpha_M)$ and $1 / \left(1 - \frac{1}{C^{\alpha_M}} \frac{\alpha_F}{\alpha_F + \alpha_M} \right)$ all increase in α_F