Occupational choice with endogenous spillovers[*]

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

June 29, 2018

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 under-invest in productive and socializing effort, and sort themselves inefficiently into occupations. We derive the optimal subsidy to reach the first best sorting into different occupations. Finally, we derive a rule to identify overpopulated sectors and establish the connection between inequality of talents, socializing and productive efforts and occupation size.

JEL-Classification: D85, H21, H23, J24

Key-words: occupational choice, socialization, social interactions, endogenous spillovers

[†]University of Nottingham and IIIEP-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 with whom they establish connections. 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 with whom to interact, and then choosing the strength of these interactions, together with their productive effort. However, the literature has explored these two dimensions of socializing separately.[1] We develop a framework to study the joint determination of both dimensions of socializing in the context of occupational choice. We show how spillovers emerge endogenously from individual productive efforts and socializing decisions, investigate the associated effects on welfare and derive novel implications for policy and empirical research.

Formally, we study a model in which individuals are endowed with different (occupation-specific) abilities and socializing is multidimensional. First, each individual decides which occupation to join. Once individuals sort themselves into occupations, they choose their productive effort and the intensity of their social interactions. Socializing allows individuals to benefit from the endogenous spillovers emerging from the productive efforts of those in the same occupation. Despite the complexity imposed by these features, our framework allows for a comprehensive equilibrium and welfare analysis in a simple and intuitive way. Our analysis generates new insights and empirical implications for occupational choice with great ease. Furthermore, as we briefly discuss in the Conclusion, our framework is compelling since it is sufficiently flexible to study many social and economic phenomena.

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

---

[1]The 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).

cally the importance of social connections for entrepreneurs,[2] for professionals[3] and even for the unemployed[4] has been widely established. Second, because spillovers matter: Guiso and Schivardi (2011) find that spillovers rather than heterogeneous entry costs are the explanation for differences in entrepreneurial activities across Italian regions.[5] Finally, 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.[6]

Analyzing endogenous spillovers leads to important insights. Our framework allows us to characterize the optimal policy to achieve efficiency in occupational choice. We provide the explicit form of the the first-best allocation for the social planner which can be achieved in equilibrium combining a linear tax with a particular subsidy for any type of distribution of talents. This result shows that the optimal policy needs to have two dimensions to correct both the within sector inefficiency caused by the externality of spillovers and the misallocations across occupations. This finding rationalizes why many measures implemented by governments to correct inefficient occupational choices, in particular measures to boost entrepreneurship, have failed.[7] There are other examples, where a policy is designed with the purpose of affecting the interaction environment of individuals. As Carrell, Sacerdote, and West (2013) show, doing this without considering the incentives for interaction, and the fact that group formation is endogenous can lead to counterproductive effects of the intervention.

One of the most novel contributions of our analysis is to study the implications of spillovers for allocations of individuals across sectors. We assume that individual abilities are given according to a Pareto distribution, since wages and income, at least at the top of the distribution, are well described by a Pareto.[8] With a Pareto distribution of talents our equilibrium is unique,

---

[2]See for example, Guiso and Schivardi (2011); Guiso, Pistaferri, and Schivardi (2015); Hoanga and Antoncic (2003).

[3]See for example West, Barron, Dowsett, and Newton (1999) for the medical and Ogus (2002) for the legal profession.

[4]Korpi (2001).

[5]The 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).

[6]For example, Guiso and Schivardi (2011); Cicala, Fryer Jr, and Spenkuch (2016); Chandra and Staiger (2007).

[7]See for example, Henrekson and Stenkula (2010); Acs, Åstebro, Audretsch, and Robinson (2016).

[8]See Guvenen, Karahan, Ozkan, and Song (2015) for a recent reference.

which is a useful feature of our framework. We find that the type of allocative inefficiency (overpopulation or underpopulation of one occupation) depends on the Pareto shape parameter and the strength of the synergies. 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 each sector is characterized by few superstars. In such a situation, the sector with stronger synergies is going to be underpopulated. On the contrary, when synergies are weak in all occupational sectors, the sector with stronger synergies is overpopulated. These results are important because they provide guidance for both policy and empirical work.

We also explore how increasing inequality in talent affects allocative inefficiency. In our framework, increasing the dispersion of a distribution enhances socializing and productive efforts. Interestingly, this effect is not confined to the occupation where inequality increases, but also takes place in the other occupation. Distributional spillovers across occupations imply that more inequality leads to a better selection of types in both occupations inducing higher productive and socializing intensities, thereby connecting two phenomena that are generally considered as independent from each other. This is another novel result with useful implications for policy and empirical analysis.[9]

We also provide a secondary set of results concerning individual decisions about productive and socializing effort within a given occupation for general distributions of talents. We show that more talented individuals do not only work harder but also generate more spillovers.[10] Furthermore, our model predicts that on average individuals in more productive occupations work harder and socialize more.[11] 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.[12] Insofar as synergies capture institutional and technological aspects of socializ-

---

[9]Inequality 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.

[10]This 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.

[11]The 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).

[12]This result is observed by Nix (2015) for the case of Sweden.

ing, 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 socializing are greater for highly productive workers; a feature that rationalizes the existence of fraternities and elite societies (e.g. Popov and Bernhardt, 2012).

This paper is organized as follows. We begin by discussing our contribution to the literature (Section 2). In Section 3, we spell out the model. Section 4 contains the equilibrium analysis and the general results valid for any occupation specific ability distribution. In Section 5, we study the case of a Pareto distribution of talents. We conclude in Section 6. Most proofs are gathered in an online Appendix.

2 Contribution to the literature

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).[13] 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 partic-

---

[13]Early examples are Heckman and Sedlacek (1985, 1990) and Jovanovic (1994).

ipants in specific occupations. Many countries make it a priority to spur entrepreneurship. Shakhnov (2014) finds that financial markets are over-crowded 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 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 very large 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.

3   A model of occupational choice

We consider an economy with a continuum of heterogeneous individuals. Occupational choice is modeled as a two-stage game. In the first stage, individuals simultaneously choose their occupation. They can either be employed in occupation M or in occupation F. For illustrative purposes we will often refer to occupation F as entrepreneurs and to occupation M as employees, but our results also apply to different occupation sectors.[14] In the second stage, all agents within the same occupation simultaneously decide their direct productive effort k_i^n and their socializing effort s_i^n.

---

[14]Indeed, when we cite evidence we will make use of the following distinctions: financial markets versus entrepreneurship, financial markets versus manufacturing, formal versus informal sector and different academic fields.

3.1 The pay-offs

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.[15]

The payoff within a particular occupation n is the sum of two components, a private component 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

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

where d^n is an occupation-specific parameter and is multiplicative in individual ability in occupation 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.[16] The synergistic returns are given by

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

where $\mathcal{N}_i$ denotes the occupational group to which individual i belongs; the parameter a captures the overall strength of synergies, $\mathbf{s}^n$ is the profile of all socializing efforts within the occupation (and which we assume has no effect between occupations) and $g_{ij}^n(s_i^n, s_j^n)$ is the link intensity of individual i and j, defined as follows:[17]

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

Each occupational group is composed by a continuum of individuals $\mathcal{N}^n \subset \mathbb{R}$ for $n \in \{M, F\}$, where the measure of the set $\mathcal{N}^n$ is N^n.

The payoff of individual i in an occupational group n is the combination

---

[15]For the time being we make no specific assumptions on how these abilities are distributed, which also implies that the distribution of talents across occupations might follow any correlation structure or be independent. The specific case of a Pareto distribution is studied in Section 5. Results for the uniform distribution can be found in the working paper version (Albornoz, Cabrales, and Hauk, 2017).

[16]Of course, socializing could also have a negative effect, say, because of revealing secrets to competitors. Assuming complementarity implies that we focus on situations where the benefits of socializing are larger than the costs.

[17]We provide a micro-foundation for this functional form in online Appendix A.1.

of i's private returns and its synergistic component:

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

We interpret the positive part of this individual payoff as the *value of individual output* and denote it by y_i^n To be more precise:

$$
y_i^n \equiv d^n b_i^n k_i^n + a d^n b_i^n \left(k_i^n\right)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n \left(k_j^n\right)^{1/2} g_{ij}^n(s_i^n, s_j^n) \right) dj.
\tag{3}
$$

3.2 Discussion of the main assumptions

Our set-up can be thought of describing occupational choice in a competitive frictionless labor market with endogenous spillovers. There is a bounded set of workers in the market, namely $\mathcal{N}^M \cup \mathcal{N}^F$. Their productivity is perfectly observed. There are no matching frictions. Firms have a linear technology and there is free entry from an unbounded set of them, and there are no fixed costs. In this environment, the positive component of individual payoffs y_i^n corresponds to the value of the output produced by this worker in equilibrium.

In our model, socializing within each occupation is undirected, but it is directed across occupations.[18] 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, individuals choose the occupational group where they socialize. Many real world examples fit this way of socializing: 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.

---

[18]Undirected 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.

We assume that productive efforts are complementary. In particular, we model synergistic returns as multiplicative in individual productivity parameters and in the square root of productive efforts additively separable by pairs.[19] 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.

We restrict individuals to belong to one single group only. 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.[20] 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.[21]

Individual ability in our model is always multiplied by the occupation-specific parameter d^n, so that the "effective" individual ability of individual i in occupation n is captured by $d^n b_i^n$. This is a purely technical assumption which amounts to a normalization of the distribution of "effective" abilities $d^n b_i^n$. It 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 "effective" abilities $d^n b_i^n$.

---

[19]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.

[20]The 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.

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

4 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.

4.1 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^{sp} and s_i^{sp} to maximize the sum of individual utilities given by

$$\int_{i \in \mathcal{N}^M \cup \mathcal{N}^F} u_i(b_i) di = \int_{i \in \mathcal{N}^M \cup \mathcal{N}^F} \left(d^{\mathcal{N}_i} b_i k_i + a d^{\mathcal{N}_i} b_i \sqrt{k_i s_i} \int_{j \in \mathcal{N}_i} \frac{d^{\mathcal{N}_i} b_j \sqrt{k_j s_j}}{N^i} dj - \frac{1}{2} k_i^2 - \frac{1}{2} s_i^2 \right) di. \tag{4}$$

Denote by $\overline{b^2} = \int_{j \in \mathcal{N}_i} b_j^2 dj$. We first define some functions that are going to be useful in the description of the equilibrium values.

$$k = \frac{d}{1 - \left(\frac{a}{2} d^2 \overline{b^2}\right)^2}, s = \frac{a}{2} d^2 \overline{b^2} k, \tag{5}$$

and

$$k^{sp} = \frac{d}{1 - \left(a d^2 \overline{b^2}\right)^2}, s^{sp} = a d^2 \overline{b^2} k^{sp}. \tag{6}$$

To avoid unbounded equilibrium choices we assume:[22]

Assumption 1. $\sup_{\mathcal{N}_i} \left(a d^2 \int_{j \in \mathcal{N}_i} b_j^2 dj\right)^2 < 1,$

which guarantees that the k and k^{sp} are always well defined. 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*

---

[22]Note that output is of the order k^2 because of the complementarities, like the costs.

each individual is equal to her own productivity multiplied by a function that is identical for all individuals in the group.[23] *That is*

$$k_i = b_i k \text{ and } s_i = b_i s \text{ for all } i \quad (7)$$

$$k_i^{sp} = b_i k^{sp} \text{ and } s_i^{sp} = b_i s^{sp} \text{ for all } i \quad (8)$$

for the individual choice problem, and for the social planner, respectively.

Proof. See online Appendix A.2. □

Proposition 1 has important empirical consequences. Since individual productivity b_i is complementary to effort, it follows that

Empirical Implication 1. *More talented individuals work harder.*

The correlation between talent and effort has been observed in education; a sector for which we have good data on both ability and effort.[24] But these individual features also translate to the group level, something that allows to make intergroup comparisons as well. On the one hand, highly 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 had a negative effect on the productivity of the Ph.D. students left behind.

Proposition 1 also shows that average socializing is increasing in average group productivity $\bar{b} = \int_{j \in \mathcal{N}_i} b_j dj$.[25] Therefore,

Empirical Implication 2. *Individuals within more productive occupational groups socialize more on average.*

---

[23]The 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.

[24]See e.g. Yeo and Neal (2004) and Babcock and Betts (2009).

[25]The common group functions are increasing in the average group squared productivity $\overline{b^2}$ and in average group productivity $\bar{b}$. Since individual socializing is $s_i = b_i s$, average socializing is $\bar{b}s$.

10

This 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 their 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.

Proposition 1 also reveals that average socializing and hence learning spillovers are increasing in network synergies a. Thus,

Empirical Implication 3. *Occupations with fewer synergies should experience lower interactions and fewer spillovers.*

This 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.

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).

Using the optimal efforts derived in Proposition 1, we can calculate the associated individual utilities.

Proposition 2. *Equilibrium individual utilities are*

$$u_i(b_i) = \frac{b_i^2 d^2}{2} \left(\frac{\left(1 + \left(\frac{a}{2}d^2\overline{b^2}\right)^2\right)}{\left(1 - \left(\frac{a}{2}d^2\overline{b^2}\right)^2\right)^2} \right), \tag{9}$$

in the individual choice problem and

$$u_i^{sp}(b_i) = \frac{\frac{1}{2}d^2 b_i^2}{\left(1 - \left(ad^2\overline{b^2}\right)^2\right)}. \tag{10}$$

for the social planner solution.

Proof. See online Appendix A.2. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

From Proposition 2 we observe that while all individuals benefit from being in a more productive occupational group,[26] higher types benefit even more from a given level of within-occupation externalities.[27] Since productivity is independent of occupational group size for a given average spillover,[28] it follows that

Empirical Implication 4. *High types have an incentive to segregate from low types if possible.*

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.[29] 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.[30]

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

---

[26] Since $\partial u_i(b_i)/\partial\overline{b^2} > 0$, and $\partial u_i^{sp}(b_i)/\partial\overline{b^2} > 0$, utility increases in $\overline{b^2}$ for everyone.

[27] That is, $\partial^2 u_i(b_i)/\partial b_i \partial\overline{b^2} > 0$, $\partial^2 u_i^{sp}(b_i)/\partial b_i \partial\overline{b^2} > 0$, so that individual type and group type are complementary.

[28] As shown by Lemma 1 in the online Appendix.

[29] Carrell, Sacerdote, and West (2013) study interaction in the context of a military academy. They show that even without physical separation within groups, individuals of different ability types tend to socialize with one another in homogeneous subgroups.

[30] There 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%.

decentralized solution (9) is lower than the individual utility resulting from the social planner solution (10). In other words,

Empirical Implication 5. *Individuals underinvest in both productive and socializing effort ($k^{sp} > k$ and $s^{sp} > s$).*

This underinvestment is specifically severe in professional activities where learning spillovers are important for productivity[31] and provides a rational for subsidizing these activities. Entrepreneurship has emerged as a key issue in the policy arena in the last few decades.[32] 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.

Proposition 3. *A subsidy that achieves efficient effort within an occupation (taking as given the selection into occupations) is given by:*

$$y_i - \frac{1}{2} d \frac{(k_i)^2}{k^{sp}}. \tag{11}$$

Proof. See online Appendix A.3 ☐

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.[33] However, it takes as given the selection into occupations. For this reason, it is only part of an optimal policy. Individuals choose their

---

[31]This 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.

[32]The Economist on 14th March 2009 published a special report on entrepreneurship with the title "Global Heroes".

[33]Of course, it also assumes that the distribution of talent and other common parameters are known. But these are things that can in principle be estimated from aggregate data and observable output.

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 efficiency.

4.2 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). 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 subsection.

We show 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 by a cutoff value C, such that individuals for whom the ratio $b_i^M/b_i^F < C$ choose group F, while individuals for whom for whom $b_i^M/b_i^F > C$ choose group M.

Such a solution implies that

$$\overline{b^{M^2}} = E\left(b_i^{M^2} \,\middle|\, b_i^M > Cb_i^F\right), \tag{12}$$

$$\overline{b^{F^2}} = E\left(b_i^{F^2} \,\middle|\, b_i^M < Cb_i^F\right). \tag{13}$$

In other words, comparative advantage determines the choice of occupation in a particularly simple way. 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.

Proposition 4. *For any underlying distribution of abilities, if assumption 1 is satisfied, both C_P and C_E exist and are decreasing in a^M and d^M and increasing in a^F and d^F.*[34]

Proof. See online 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

---

[34]Since we do not establish equilibrium uniqueness, this proposition has to be understood in a local way, assuming the parameter change does not change equilibrium selection.

the F-group). Thus, according to Proposition 4, 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 can be caused by the introduction of new or improved communication technology facilities. Shifts in d could be technological changes that affect the productivity of every individual in a given occupation. Or they could be due to institutional features.

The effect of communication technologies (changes in a) on productivities has been widely acknowledged.[35] To our knowledge, there is no study linking the relative sizes of economic sectors with their differential adoption of communication technologies. This paper provides a clear prediction for this linkage.

Empirical Implication 6. *The differential adoption of communication technologies in different sectors should be accompanied by an increase in the relative size of the sector after the technology is adopted.*

This prediction can be tested in future research and exhibits 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. This is especially relevant if we interpret our model as choosing to work in the formal or informal sector. In some institutional settings very large (or very small) firms are extremely regulated, while in others there are too many loopholes for politically connected firms. A looser control of informal activities induces a high d in the informal sector. The d in the formal sector would suffer from high taxation. Therefore,

Empirical Implication 7. *More people will work in the informal sector at the expense of working in the formal sector the looser the controls of informal activities and the higher formal sector taxation.*

Lopez-Martin (2015) finds plentiful evidence consistent with this implication.

---

[35]Some 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.

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. However, independently of the direction of inefficiency, we can show that every relative sector size can be achieved using a linear tax/subsidy no matter the underlying talent distributions and hence the planner can also achieve the social optimal sorting into occupations. Furthermore:

Proposition 5. *The first-best allocation for the social planner, including the socially optimal C, can be achieved in equilibrium using a linear tax/subsidy on output t, plus a subsidy equal to* $t\left(y_i - \frac{1}{2}d\frac{(k_i)^2}{k^{sp}}\right)$.

Proof. See online Appendix A.5. □

Proposition 5 establishes that a first-best allocation can be achieved by combining a linear tax with a particular subsidy for any type of distribution of talents. However, since there might be multiple equilibria in the occupational choice stage, it abstracts from equilibrium selection issues. Also, in order to determine the correct linear tax or subsidy, which depends on whether an unregulated occupational sector is too big or too small, we will need to make specific assumptions about the underlying talent distributions. In what follows we will focus on the Pareto distribution which is empirically relevant and - as we will show - leads to a unique equilibrium in our model.[36]

5 The case of a Pareto distribution of talent

The empirical relevance of the Pareto distribution for describing variations of wages and income across individuals has been well established.[37] Since its shape parameter is an inverse measure of the spread of talent, the Pareto distribution therefore also captures the empirical distribution of talents. Assuming that individual talents in each occupation follow a Pareto distribution and that talent is occupation-specific leads to a unique equilibrium in our model.[38]

Proposition 6. *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 (28) and C_E, defined by (30) exist and are unique.*

---

[36] The working paper version also includes results for the uniform distribution (Albornoz, Cabrales, and Hauk, 2017).

[37] For example, Mandelbrot (1960); Guvenen, Karahan, Ozkan, and Song (2015).

[38] Notice that we assume uncorrelated talents for expositional simplicity. As discussed below, our main results in this section are robust to the correlation structure.

Proof. See online Appendix A.6. □

Consequently, there is also a unique first best solution, which requires an intervention at both margins, i.e. by inducing optimal effort within the occupation as well as choices leading to the optimal occupational choice. Without an intervention on location, one sector will be overpopulated while the other sector will be underpopulated. There is plenty of evidence of excessive or insufficient size of specific occupations. Shakhnov (2014) shows that financial markets are too large with respect to the entrepreneurship sector with a model that 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.

The following results provide some insights on the direction of overpopulation when abilities follow a Pareto distribution and investments in productive and socializing effort are optimally determined by a social planner.[39]

Proposition 7. *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. Without loss of generality assume that sector F has higher overall strength of synergies, i.e. $a^F d^F > a^M d^M$. Then social welfare may increase by adding (i.e. $\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} > 0$) or by decreasing (i.e. $\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} < 0$) the number of workers in occupation F.*

In particular,

- *The sector with the overall higher strength of synergies (the F-sector) is underpopulated for distributions with relatively low dispersion.[40]*

---

[39]Proposition 7 established that intervening only locally within a group leads to suboptimal choice of occupations. This poses the question whether an uncoordinated intervention can be worse than no intervention. In the working paper version (Albornoz, Cabrales, and Hauk, 2017), we provide an example of such a result when the talent distribution is uniform and in the presence of a special type of congestion costs. In this setup the equilibrium is also unique.

[40]The exact technical condition is the following: 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 $\left. \frac{\partial w(C)}{\partial C} \right|_{C=C_E} > 0$).

- *When overall synergies are sufficiently small in both occupations, the F-sector is overpopulated for sufficiently low synergies a^F.[41]*

Proof. See online Appendix A.7. □

If the overall strength of synergies is higher in the F-sector, then reallocating M-types that are close to indifferent to occupation F leads 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.[42] The overall effect on social welfare is therefore ambiguous. Notice that Proposition 7 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 implies that the number of superstars is very small.[43] Thus, if both sectors have a low number of very able individuals, the welfare can increase by augmenting the size of the sector, which has the larger impact of synergies.

Empirical Implication 8. *If the number of very able individuals in each sector is small, the sector with the larger impact of synergies will be too small.*

Proposition 7 also establishes a second rule of thumb:

Empirical Implication 9. *The size of the occupation with higher overall synergies is sub-optimally large when synergies are sufficiently small in both occupations.[44]*

To the extent that dispersion of talents and the strength of spillovers within a particular occupation are observable, Empirical Implications 8 and 9 provide potentially useful rules of thumb to detect local underpopulation or overpopulation of different occupational sectors. Verifying these rules is left as an empirical challenge for future work.

---

[41]The exact condition is the following: 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 the F-sector is overpopulated ($\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} < 0$).

[42]This is true because the average type in occupation F decreases with C, while the average type in occupation M increases with C.

[43]Because high α implies low dispersion so the tails of the distribution are thin.

[44]To see this, notice that occupation F is overpopulated when a^F is very low. Since F has higher overall synergies ($a^{M^2} d^{M^2} < a^{F^2} d^{F^2}$) this also implies a sufficiently low a^M

Inequality and effort choices

The Pareto distribution is also appropriate to study the link between inequality of abilities and productive and socializing efforts. Since - as we explained already - the shape parameter α_i is an (inverse) measure of the spread of talent, 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 that keeps the unconditional mean of the Pareto distribution constant.[45] This way, we focus exclusively on the effect of changes in the dispersion of talent, which we associate with inequality.

Proposition 8. *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 in the unconditional mean of squared types.[46] Then, if we hold C_E or C_P constant, both $\overline{b^{M2}}$ and $\overline{b^{F2}}$ increase, and thus productive and socializing effort increase in both occupations.*

Proof. See Appendix A.8. □

Proposition 8 simply states that if the dispersion of talents in one occupation increases, both occupations receive a better selection of types and hence productive and socializing effort increase in both occupations. The basic intuition is that the tails of one of the distribution is now larger and comparative advantage forces a selection mostly from the tails.[47]

Clearly, the effect of inequality in talent on socializing and productive efforts emerges from the existence of spillovers within occupations. This is a

---

[45]More specifically, as α_j falls we impose an equivalent change is a to reduce effort as much as necessary to fix the unconditional mean of the Pareto distribution; which is $E\left(b_i^{j2}\right) = \alpha_j / (\alpha_j - 2)$, for $j \in \{M, F\}$).

[46]i.e. $a' = a\,(\alpha_j - 2)\,\alpha_j$.

[47]For a more analytical explanation, note that the expression for $\overline{b^{F2}}$ can be written as:

$$\overline{b^{F2}} = \int_1^\infty b^{F2} f_F\left(b^F\right) \frac{F_M\left(Cb^F\right)}{\int_1^\infty f_F\left(b^F\right) F_M\left(Cb^F\right) db^F} db^F.$$

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 $F_M\left(Cb^F\right) \int_1^\infty f_F\left(b^F\right) F_M\left(Cb^F\right) db^F$.

The effect of a decrease of α_F is more direct, as it increases $f_F\left(b^F\right)$ 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 $F\left(Cb^F\right) \int_1^\infty f\left(b^F\right) F\left(Cb^F\right) db^F$ term, now unchanged, gives more weight to changes that occur for higher values of b^F.

novel empirical implication of our model that stands as a challenge for future empirical work.

Empirical Implication 10. *An increase in inequality in talent leads to higher production and more socialization in all occupations.*

Admittedly, assuming that productivity parameters are independent across occupations requires some degree of heroism. A natural question is whether our results hinge on this assumption. We find it reassuring that our propositions 6 and 8 for the Pareto distribution are robust to the following correlation structure: with probability p the two values of b_i^j for $j \in \{F, M\}$ are independent of one another. With probability $(1 - p)$, they coincide, namely $b_i^M = b_i^F$ and they are distributed with shape parameter α_F. Indeed, both propositions 6, and 8 are proved under this assumption which includes the independence assumption for $p = 1$.[48]

6 Conclusion

In this paper, 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 introduction of new communication technologies, which may be adopted differentially across sectors. Another interesting implication of endogenous spillovers is that a higher inequality of abilities in one occupation imply more socialization and productive efforts in both occupations. This is something that would not happen in a world without spillovers within occupations.

Our framework can be applied to investigate a range of different contexts, where individuals choose which group to belong to and then decide how much

---

[48]We have not been able to show whether Proposition 7 is also robust to correlation, but we have not found a counterexample either.

to invest in the productive and socializing efforts. Education choices share many features with the case we studied in this paper. We intend to use this model in future work to study the demand and supply for different subjects and skills (e.g. high-level Science, Technology, Engineering, and Math). Interestingly, the case of education choices requires endogenizing ability. In Albornoz, Cabrales, and Hauk (2017), we take a first step in this direction, where parents can initially invest in their children's abilities. We show that parental educational investment can mitigate or reinforce distributional inefficiencies. Our model could also be useful to study residential choice, as the benefits of living in a community often depend on social interactions within them. The choices of leisure activities are another potential fruitful avenue of application of our ideas. A more intriguing area for the development of this kind of model refers to aspects more connected to an individual's identity. The national, religious, or ethnic identification of a person is sometimes a matter of choice, and is connected to the decisions of others. For example, whether a person feels she is European, British or Welsh, and to which degree, could be influenced by her efforts and those of others in pursuit of their own identity. We think that our contribution is an important step towards understanding the determinants and effects of socializing.

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.[49] 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.

---

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

References

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., A. CABRALES, AND E. HAUK (2017): "Which club should I attend, Dad?: Targeted socialization and production," Discussion paper, CEPR DP10321.

ALBORNOZ, F., A. CABRALES, E. HAUK, 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*, 25, 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.

CARRELL, S. E., B. I. SACERDOTE, AND J. E. WEST (2013): "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 81(3), 855–882.

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. JOVANOVIC (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.

GUISO, L., L. PISTAFERRI, AND F. SCHIVARDI (2015): "Learning Entrepreneurship From Other Entrepreneurs?," Discussion paper, National Bureau of Economic Research, No. 21775.

GUISO, 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.

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 smallholder 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 Online appendices

A.1 A microfoundation for the g_{ij}^n function

Lemma 1. *Suppose that, for all* $\mathbf{s} \neq \mathbf{0}$*, the link intensity satisfies the following assumptions*[50]*:*

(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: $\int_{j \in \mathcal{N}_i} g_{ij}^n(s_i^n, s_j^n) dj = \frac{1}{N^n} \int_{j \in \mathcal{N}_i} \left(s_i^n\right)^{1/2} \left(s_j^n\right)^{1/2} dj$;

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

Lemma 2. *Then, the link intensity is given by*

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

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

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

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

Notice that given $(A2)$ and 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 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 $1 / \left(N^n\right)^{\beta}$ for some β different from 1.

A.2 Proof of Propositions 1 and 2

The FOC for the decentralized problem are

---

[50]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.

$$k_i = db_i + \frac{a}{2}d^2b_i\sqrt{\frac{s_i}{k_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \qquad (14)$$

$$s_i = \frac{a}{2}d^2b_i\sqrt{\frac{k_i}{s_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \qquad (15)$$

while the FOC for the social planner simplify to

$$k_i^{sp} = db_i + ad^2b_i\sqrt{\frac{s_i^{sp}}{k_i^{sp}}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_j^{sp}s_j^{sp}}}{N^i}dj \text{ for all } i \qquad (16)$$

$$s_i^{sp} = ad^2b_i\sqrt{\frac{k_i^{sp}}{s_i^{sp}}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_j^{sp}s_j^{sp}}}{N^i}dj \text{ for all } i \qquad (17)$$

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

We divide (14) by (15) to get

$$\frac{k_i}{s_i} = \frac{d + \frac{a}{2}d^2\sqrt{\frac{s_i}{k_i}}K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right)}{\frac{a}{2}d^2\sqrt{\frac{k_i}{s_i}}K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right)} = \frac{\sqrt{\frac{k_i}{s_i}} + \frac{a}{2}dK\left(\mathbf{b},\mathbf{k},\mathbf{s}\right)}{\frac{a}{2}d\frac{k_i}{s_i}K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right)} \qquad (18)$$

where bold face letters denote vectors and

$$K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right) = \int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj$$

Rearranging (18) gives

$$d\left(\frac{k_i}{s_i}\right)^2\frac{a}{2}K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right) = \sqrt{\frac{k_i}{s_i}} + d\frac{a}{2}K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right) \qquad (19)$$

from which it is immediate that

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

for some $K\left(.\right)$ with a unique solution. To see the uniqueness notice that letting $\sqrt{\frac{k_i}{s_i}} = x_i$ (19) can be written as

$$dx_i^4\frac{a}{2}K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right) = x_i + d\frac{a}{2}K\left(\mathbf{b},\mathbf{k},\mathbf{sp}\right) \qquad (20)$$

2

the left hand side of (20) 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 $d\frac{a}{2}K(\mathbf{b}, \mathbf{k}, \mathbf{sp})$ when $x_i = 0$. Hence there is a single crossing point at the positive orthant.

Hence

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

Thus it is clear we can write

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

An analogous proof establishes that also for the centralized problem

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

It remains to determine the common optimal group parameters.

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

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

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

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

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

3

$$k = \frac{d}{1 - \left(\frac{a}{2}d^2\overline{b^2}\right)^2}$$

$$s = \frac{\frac{a}{2}d^3\overline{b^2}}{1 - \left(\frac{a}{2}d^2\overline{b^2}\right)}$$

for the decentralized problem and

$$k^{sp} = d + ad^2\sqrt{\frac{s^{sp}}{k^{sp}}}\overline{b^2}\sqrt{k^{sp}s^{sp}} = d + ad^2\overline{b^2}s^{sp} \tag{21}$$

$$s^{sp} = ad^2\sqrt{\frac{k^{sp}}{s^{sp}}}\overline{b^2}\sqrt{k^{sp}s^{sp}} = ad^2\overline{b^2}k^{sp} \tag{22}$$

$$k^{sp} = \frac{d}{1 - \left(ad^2\overline{b^2}\right)^2}$$

$$s^{sp} = \frac{ad^3\overline{b^2}}{1 - \left(ad^2\overline{b^2}\right)}$$

for the centralized problem.

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

Introducing the optimal investment levels into the utility functions gives us

$$k = \frac{d}{1 - \left(\frac{a}{2}d^2\overline{b^2}\right)^2}$$

$$s = \frac{\frac{a}{2}d^3\overline{b^2}}{1 - \left(\frac{a}{2}d^2\overline{b^2}\right)}$$

$$u_i(b_i) = db_i^2 k + ad^2 b_i^2 k s \overline{b^2} - \frac{1}{2} b_i^2 k^2 - \frac{1}{2} b_i^2 s^2$$

$$= db_i^2 \frac{d}{1 - \left(\frac{a}{2} d^2 \overline{b^2}\right)^2} + ad^2 b_i^2 \frac{d}{1 - \left(\frac{a}{2} d^2 \overline{b^2}\right)^2} \frac{\frac{a}{2} d^3 \overline{b^2}}{1 - \left(\frac{a}{2} d^2 \overline{b^2}\right)} \overline{b^2}$$

$$- \frac{1}{2} b_i^2 \left(\frac{d}{1 - \left(\frac{a}{2} d^2 \overline{b^2}\right)^2} \right)^2 - \frac{1}{2} b_i^2 \left(\frac{\frac{a}{2} d^3 \overline{b^2}}{1 - \left(\frac{a}{2} d^2 \overline{b^2}\right)} \right)^2$$

$$= \frac{b_i^2 d^2}{2} \left(\frac{\left(1 + \left(\frac{a}{2} d^2 \overline{b^2}\right)^2\right)}{\left(1 - \left(\frac{a}{2} d^2 \overline{b^2}\right)^2\right)^2} \right)$$

for the decentralized solution and

$$u_i^{sp}(b_i) = db_i^2 k^{sp} + ad^2 b_i^2 k^{sp} s^{sp} \overline{b^2} - \frac{b_i^2}{2} (k^{sp})^2 - \frac{b_i^2}{2} (s^{sp})^2$$

$$= \frac{\frac{1}{2} d^2 b_i^2}{\left(1 - \left(ad^2 \overline{b^2}\right)^2\right)}.$$

for the centralized solution.

A.3 Proof of Proposition 3

Let the subsidy be

$$y_i - \frac{1}{2} d \frac{(k_i)^2}{k^{sp}}$$

then the resulting utility with this subsidy is

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

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

This leads to first order conditions

$$k_i = 2db_i + ad^2 b_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj - d \frac{k_i}{k^{sp}} \text{ for all } i \qquad (23)$$

$$s_i = ad^2 b_i \sqrt{\frac{k_i}{s_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \text{ for all } i \qquad (24)$$

5

letting $k_i = b_i k^{sp}$ and $s_i = b_i s^{sp}$, we get

$$k^{sp} = d + ad^2 \sqrt{\frac{s^{sp}}{k^{sp}}} \overline{b^2} \sqrt{k^{sp} s^{sp}} = d + ad^2 \overline{b^2} s^{sp} \tag{25}$$

$$s^{sp} = ad^2 \sqrt{\frac{k^{sp}}{s^{sp}}} \overline{b^2} \sqrt{k^{sp} s^{sp}} = ad^2 \overline{b^2} k^{sp} \tag{26}$$

Clearly the system $(25), (26)$ is the same as $(21), (22)$ and thus $(23), (24)$ also solves the same system as $(16),(17)$ and the result follows.

A.4   Proof of Proposition 4

We first establish existence and a useful technical result for the rest of the proposition.

Lemma 3. *For any underlying distribution of abilities, if assumption 1 is satisfied, there exist mappings $f(C)$ and $g(C)$ such that a zero of the mappings $f(C)$ and $g(C)$ is an equilibrium of, respectively, the decentralized and centralized problems. Furthermore an equilibrium always exists, and in any stable equilibrium, $\frac{\partial f(C)}{\partial C} < 0$ and $\frac{\partial g(C)}{\partial C} < 0$.*

Proof. 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

$$\frac{b_i^{M^2} d^{M^2}}{2} \left(\frac{\left(1 + \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}^2\right)^2\right)}{\left(1 - \left(\frac{a_M}{2} d^{M^2} \overline{b^{M^2}}^2\right)^2\right)^2} \right) > \frac{b_i^{F^2} d^{F^2}}{2} \left(\frac{\left(1 + \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}^2\right)^2\right)}{\left(1 - \left(\frac{a_F}{2} d^{F^2} \overline{b^{F^2}}^2\right)^2\right)^2} \right). \tag{27}$$

If the dividing line exists, its slope is defined when the expressions on either side of the inequality in (27) are equal. In other words, the dividing line is defined by the following expression:

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

Hence C_P is the fixed point of the mapping

$$C_P = \frac{d^F}{d^M} \sqrt{\frac{\left(1 + \left(\frac{a_F}{2} d^{F2} \overline{b^{F2}}^2\right)^2\right) \left(1 - \left(\frac{a_M}{2} d^{M2} \overline{b^{M2}}^2\right)^2\right)^2}{\left(1 - \left(\frac{a_F}{2} d^{F2} \overline{b^{F2}}^2\right)^2\right)^2 \left(1 + \left(\frac{a_M}{2} d^{M2} \overline{b^{M2}}^2\right)^2\right)}}, \qquad (28)$$

where the right hand of (28) depends on C_P through $\overline{b^{M2}}$ and $\overline{b^{F2}}$, which are defined by equations (12) and (13) respectively. Put differently, C_P is implicitly defined by a zero of the mapping

$$g(C, \cdot) \equiv \frac{d^{F2}}{d^{M2}} \frac{\left(1 + \left(\frac{a^F}{2} d^{F2} \overline{b^{F2}}\right)^2\right) \left(1 - \left(\frac{a^M}{2} d^{M2} \overline{b^{M2}}\right)^2\right)^2}{\left(1 - \left(\frac{a^F}{2} d^{F2} \overline{b^{F2}}\right)^2\right)^2 \left(1 + \left(\frac{a^M}{2} d^{M2} \overline{b^{M2}}\right)^2\right)} - C^2. \qquad (29)$$

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

$$C_E = \frac{d^F}{d^M} \sqrt{\frac{\left(1 - \left(a d^{M2} \overline{b^{M2}}\right)^2\right)}{\left(1 - \left(a d^{F2} \overline{b^{F2}}\right)^2\right)}}, \qquad (30)$$

and is implicitly defined by a zero of the mapping

$$f(C, \cdot) \equiv \frac{d^{F2}}{d^{M2}} \frac{1 - \left(a^M d^{M2} \overline{b^{M2}}\right)^2}{1 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2} - C^2. \qquad (31)$$

Define

$$g_A(C) \equiv \frac{d^{F2}}{d^{M2}} \frac{\left(1 + \left(\frac{a^F}{2} d^{F2} \overline{b^{F2}}\right)^2\right) \left(1 - \left(\frac{a^M}{2} d^{M2} \overline{b^{M2}}\right)^2\right)^2}{\left(1 - \left(\frac{a^F}{2} d^{F2} \overline{b^{F2}}\right)^2\right)^2 \left(1 + \left(\frac{a^M}{2} d^{M2} \overline{b^{M2}}\right)^2\right)}$$

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

1

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

Then, note that the assumption $\sup_C \left(a d^2 \overline{b^2} \right)^2 < 1$ 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 $C^* \in \left(0, \overline{C}_g \right)$ 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 $\sup_C \left(a d^2 \overline{b^2} \right)^2 < 1$

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

The assumption $\sup_C \left(a d^2 \overline{b^2} \right)^2 < 1$ 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 $C^* \in \left(0, \overline{C}_f\right)$ such that $f(C^*, \cdot) = 0$.

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

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

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

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

where $R(.)$ 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 $f(C, \cdot)$.

$\square$

By Lemma 3, $\partial f(C)/\partial C < 0$ and $\partial g(C)/\partial C < 0$, hence to establish com-

parative static results, one only needs to check 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 $n \in \{M, F\}$. So using the implicit function theorem, we only need to check how the functions $f(.)$ and $g(.)$ 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

$$\frac{\partial f(C, \cdot)}{\partial a^M} = \frac{d^{F2}}{d^{M2}} \frac{-2a^M \left(d^{M2}\overline{b^{M2}}\right)^2}{1 - \left(a^F d^{F2}\overline{b^{F2}}\right)^2} < 0$$

$$g(C, \cdot) = \frac{d^{F2}}{d^{M2}} \frac{\left(4 + \left(a^F d^{F2}\overline{b^{F2}}\right)^2\right) \left(4 - \left(a^M d^{M2}\overline{b^{M2}}\right)^2\right)^2}{\left(4 - \left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)^2 \left(4 + \left(a^M d^{M2}\overline{b^{M2}}\right)^2\right)} - C_P^2 = 0$$

$$sign \left(\frac{\partial g(C, \cdot)}{\partial a^M}\right)$$

$$= sign \left(\frac{-4a^M \left(d^{M2}\overline{b^{M2}}\right)^2 \left(4 - \left(a^M d^{M2}\overline{b^{M2}}\right)^2\right) \left(4 + \left(a^M d^{M2}\overline{b^{M2}}\right)^2\right)}{\left(4 + \left(a^M d^{M2}\overline{b^{M2}}\right)^2\right)^2} - \frac{2a^M \left(d^{M2}\overline{b^{M2}}\right)^2 \left(4 - \left(a^M d^{M2}\overline{b^{M2}}\right)^2\right)^2}{\left(4 + \left(a^M d^{M2}\overline{b^{M2}}\right)^2\right)^2} \right)$$

$$< 0$$

Hence

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

$$\frac{dC_P}{da^M} = -\frac{\frac{\partial g(C, \cdot)}{\partial a^M}}{\frac{\partial g(C, \cdot)}{\partial C_P}} < 0$$

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.

$$\frac{\partial f(C,\cdot)}{\partial a^F} = \frac{d^{F2}}{d^{M2}} \frac{\left(1 - \left(a^M d^{M2} \overline{b^{M2}}\right)^2\right) 2a^F \left(d^{F2} \overline{b^{F2}}\right)^2}{\left(1 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2\right)^2} > 0$$

$$sign\left(\frac{\partial g(C,\cdot)}{\partial a^F}\right)$$

$$= sign\left(\frac{2a^F \left(d^{F2} \overline{b^{F2}}\right)^2 \left(4 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2\right)^2}{\left(\left(4 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2\right)^2\right)^2} + \frac{4a^F \left(d^{F2} \overline{b^{F2}}\right)^2 \left(4 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2\right)\left(4 + \left(a^F d^{F2} \overline{b^{F2}}\right)^2\right)}{\left(\left(4 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2\right)^2\right)^2}\right)$$

$$> 0$$

Hence

$$\frac{dC_E}{da^F} = -\frac{\frac{\partial f(C,\cdot)}{\partial a^F}}{\frac{\partial f(C,\cdot)}{\partial C_E}} > 0$$

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

Now we look at changes in d^M.

$$f(C,\cdot) = \frac{d^{F2}}{d^{M2}} \frac{1 - \left(a^M d^{M2} \overline{b^{M2}}\right)^2}{1 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2} - C_E^2 = 0 \tag{32}$$

$$\frac{\partial f(C,\cdot)}{\partial d^M} = -\frac{2d^{F2}}{d^{M3}} \frac{1 - \left(a^M d^{M2} \overline{b^{M2}}\right)^2}{1 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2} - \frac{d^{F2}}{d^{M2}} \frac{4d^{M3} \left(a^M \overline{b^{M2}}\right)^2}{1 - \left(a^F d^{F2} \overline{b^{F2}}\right)^2} < 0$$

Hence

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

Similarly

$$sign\left(\frac{\partial g(C,\cdot)}{\partial d^M}\right) = sign\left(\frac{-4d^{M^3}\left(a^M\overline{b^{M^2}}\right)^2\left(4-\left(a^Md^{M^2}\overline{b^{M^2}}\right)^2\right)\left(4+\left(a^Md^{M^2}\overline{b^{M^2}}\right)^2\right)}{\left(4+\left(a^Md^{M^2}\overline{b^{M^2}}\right)^2\right)^2} - \frac{4d^{M^3}\left(a^M\overline{b^{M^2}}\right)^2\left(4-\left(a^Md^{M^2}\overline{b^{M^2}}\right)^2\right)^2}{\left(4+\left(a^Md^{M^2}\overline{b^{M^2}}\right)^2\right)^2}\right)$$

$$< 0$$

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.

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

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

12

$$sign\left(\frac{\partial g(C,\cdot)}{\partial d^F}\right)$$

$$= \frac{d^{F2}\left(4+\left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)}{\left(4-\left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)^2}$$

$$= sign\left(\begin{array}{c}\dfrac{\left(8d^F+6d^{F5}\left(a^F\overline{b^{F2}}\right)^2\right)\left(4-\left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)^2}{\left(\left(4-\left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)^2\right)^2}\\[4mm]+\dfrac{8d^{F3}\left(a^F\overline{b^{F2}}\right)^2\left(4-\left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)\left(4d^{F2}+d^{F2}\left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)}{\left(\left(4-\left(a^F d^{F2}\overline{b^{F2}}\right)^2\right)^2\right)^2}\end{array}\right)$$

$$> 0$$

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 5

We first establish

Lemma 4. *Any $C \in [0,\infty)$ can be obtained in equilibrium using a linear tax/subsidy on output.*

Proof. We characterize the optimal choices under a linear tax/subsidy on output. The FOC for the decentralized problem are

$$k_i = dtb_i + \frac{a}{2}d^2tb_i\sqrt{\frac{s_i}{k_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \tag{33}$$

$$s_i = \frac{a}{2}d^2tb_i\sqrt{\frac{k_i}{s_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \tag{34}$$

We first prove that $\frac{k_i}{s_i} = \frac{k_j}{s_j}$ for all i and j.
We divide (33) by (34) to get

$$\frac{k_i}{s_i} = \frac{d + \frac{a}{2}d^2\sqrt{\frac{s_i}{k_i}}K(\mathbf{b},\mathbf{k},\mathbf{sp})}{\frac{a}{2}d^2\sqrt{\frac{k_i}{s_i}}K(\mathbf{b},\mathbf{k},\mathbf{sp})} = \frac{\sqrt{\frac{k_i}{s_i}} + \frac{a}{2}dK(\mathbf{b},\mathbf{k},\mathbf{s})}{\frac{a}{2}d\frac{k_i}{s_i}K(\mathbf{b},\mathbf{k},\mathbf{sp})} \tag{35}$$

where bold face letters denote vectors and

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

Rearranging (35) gives

$$d\left(\frac{k_i}{s_i}\right)^2\frac{a}{2}K(\mathbf{b},\mathbf{k},\mathbf{sp}) = \sqrt{\frac{k_i}{s_i}} + d\frac{a}{2}K(\mathbf{b},\mathbf{k},\mathbf{sp}) \tag{36}$$

from which it is immediate that

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

for some $K(.)$ with a unique solution. To see the uniqueness notice that letting $\sqrt{\frac{k_i}{s_i}} = x_i$ (36) can be written as

$$dx_i^4\frac{a}{2}K(\mathbf{b},\mathbf{k},\mathbf{sp}) = x_i + d\frac{a}{2}K(\mathbf{b},\mathbf{k},\mathbf{sp}) \tag{37}$$

the left hand side of (37) 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 $d\frac{a}{2}K(\mathbf{b},\mathbf{k},\mathbf{sp})$ when $x_i = 0$. Hence there is a single crossing point at the positive orthant.

Hence

$$k_i = dtb_i + \frac{a}{2}d^2tb_i\frac{K(\mathbf{b},\mathbf{k},\mathbf{sp})}{\sqrt{F(K(\mathbf{b},\mathbf{k},\mathbf{sp}))}} \text{ for all } i$$

$$s_i = \frac{a}{2}d^2tb_i\sqrt{F(K(\mathbf{b},\mathbf{k},\mathbf{sp}))}K(\mathbf{b},\mathbf{k},\mathbf{sp}) \text{ for all } i$$

Thus it is clear we can write

$$k_i = b_ik(\mathbf{b},\mathbf{k},\mathbf{s},t) \text{ for all } i$$

$$s_i = b_is(\mathbf{b},\mathbf{k},\mathbf{s},t) \text{ for all } i$$

We now determine the common optimal group parameters.

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

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

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

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

which lead to

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

Assuming $\left(a d^2 t \overline{b^2}\right)^2 < 1$ guarantees positive investment levels.

Introducing the optimal investment levels into the utility functions gives us

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

so that

$$
u_i(b_i) = \frac{b_i^2 d^2 t^2}{2} \left(\frac{1 + \left(\frac{a}{2} d^2 t \overline{b^2}\right)^2}{\left(1 - \left(\frac{a}{2} d^2 t \overline{b^2}\right)^2\right)^2} \right) \tag{38}
$$

15

for the decentralized solution.

Similarly for the centralized solution the first order conditions are

$$k_i = dtb_i + ad^2tb_i\sqrt{\frac{s_i}{k_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \tag{39}$$

$$s_i = ad^2tb_i\sqrt{\frac{k_i}{s_i}}\int_{j\in\mathcal{N}_i}\frac{b_j\sqrt{k_js_j}}{N^i}dj \text{ for all } i \tag{40}$$

and using analogous arguments as before we have that using $k_i = b_ik^{sp}$ and $s_i = b_is^{sp}$ the expressions (39) and (40) can be written as

$$k = dt + d^2ta\sqrt{\frac{sp}{k}}\overline{b^2}\sqrt{ks} = dt + ad^2t\overline{b^2}s$$

$$s = ad^2t\sqrt{\frac{k}{sp}}\overline{b^2}\sqrt{ks} = ad^2t\overline{b^2}k$$

where

$$k = \frac{dt}{1 - \left(ad^2t\overline{b^2}\right)^2}$$

$$s = \frac{ad^3t^2\overline{b^2}}{1 - \left(ad^2t\overline{b^2}\right)^2}$$

This leads, after some manipulations, as before, to

$$u_i(b_i) = \frac{b_i^2d^2t^2}{2}\frac{1}{\left(1 - \left(ad^2t\overline{b^2}\right)^2\right)^2} \tag{41}$$

From expression (38) we get that the equation that defines C_P implicitly is

$$g(C, t, \cdot) = \frac{d^{F^2}t^{F^2}}{d^{M^2}t^{M^2}}\frac{\left(4 + \left(a^Fd^{F^2}t^F\overline{b^{F^2}}\right)^2\right)\left(4 - \left(a^Md^{M^2}t^M\overline{b^{M^2}}\right)^2\right)^2}{\left(4 - \left(a^Fd^{F^2}t^F\overline{b^{F^2}}\right)^2\right)^2\left(4 + \left(a^Md^{M^2}t^M\overline{b^{M^2}}\right)^2\right)} - C_P^2 = 0 \tag{42}$$

16

and from expression (41) we get that the equation that defines C_P implicitly is

$$f(C, t, \cdot) = \frac{d^{F^2} t^{F^2}}{d^{M^2} t^{M^2}} \frac{1 - \left(a^M d^{M^2} t^M \overline{b^{M^2}}\right)^2}{1 - \left(a^F d^{F^2} t^F \overline{b^{F^2}}\right)^2} - C_E^2 = 0 \qquad (43)$$

From expression (43) and (42) we get that $\lim_{\frac{t^F}{t^M} \to 0} C_E = \lim_{\frac{t^F}{t^M} \to 0} C_P = 0$ and $\lim_{\frac{t^M}{t^F} \to 0} C_E = \lim_{\frac{t^M}{t^F} \to 0} C_P = \infty$. 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 $\frac{t^F}{t^M}$.

$\square$

Having established Lemma 4 we now proceed with the remainder of the proof of Proposition 5

Given the tax and subsidy scheme proposed we can write the utility of the agent as

$$
\begin{aligned}
u_i^n &= t^n \left(d^n b_i^n k_i^n + a d^n b_i^n (k_i^n)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n (k_j^n)^{1/2} g_{ij}^n(s_i^n, s_j^n) \right) dj \right) - \frac{1}{2} (k_i^n)^2 - \frac{1}{2} (s_i^n)^2 \\
&\quad + t^n \left(y_i - \frac{1}{2} d \frac{(k_i)^2}{k} \right) \\
&= 2 t^n \left(d^n b_i^n k_i^n + a d^n b_i^n (k_i^n)^{1/2} \int_{j \in \mathcal{N}_i} \left(d^n b_j^n (k_j^n)^{1/2} g_{ij}^n(s_i^n, s_j^n) \right) dj \right) - \frac{1}{2} (k_i^n)^2 - \frac{1}{2} (s_i^n)^2 \\
&\quad - t^n \left(\frac{1}{2} d \frac{(k_i)^2}{k} \right)
\end{aligned}
$$

The FOC for the decentralized problem for all i are:

$$
\begin{aligned}
k_i &= d t^n b_i + \frac{a}{2} d^2 t^n b_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj + t^n \left(d b_i + \frac{a}{2} d^2 b_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj - d \frac{k_i}{k} \right) \\
s_i &= \frac{a}{2} d^2 t^n b_i \sqrt{\frac{k_i}{s_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj + t^n \left(\frac{a}{2} d^2 b_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \right).
\end{aligned}
$$

Letting $k_i = b_i k^{sp}$ and $s_i = b_i s^{sp}$, we have

$$k_i = dt^n b_i + ad^2 t^n b_i \sqrt{\frac{s_i}{k_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \text{ for all } i \tag{44}$$

$$s_i = ad^2 t^n b_i \sqrt{\frac{k_i}{s_i}} \int_{j \in \mathcal{N}_i} \frac{b_j \sqrt{k_j s_j}}{N^i} dj \text{ for all } i \tag{45}$$

Notice that the expressions (44) and (45) are identical to (39) and (40) Hence we will have that

$$k^{sp} = \frac{dt^n}{1 - \left(\frac{a}{2} d^2 t^n \overline{b^2}\right)^2}$$

$$s^{sp} = \frac{\frac{a}{2} d^3 t^{n2} \overline{b^2}}{1 - \left(\frac{a}{2} d^2 t^n \overline{b^2}\right)^2}$$

and

$$u_i(b_i) = \frac{b_i^2 d^2 t^2}{2} \frac{1}{\left(1 - \left(ad^2 t \overline{b^2}\right)^2\right)^2} \tag{46}$$

The remainder of the proof follows from Lemma 4.

A.6 Proof of Proposition 6

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

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

We will prove a slightly more general statement of the proposition, allowing for the following correlation structure. With probability p the two values of b_i^j for $j \in \{F, M\}$ are independent of one another. With probability $(1 - p)$, $b_i^M = b_i^F$ and they are distributed with shape parameter α_F.

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 $C \ge 1$.[51] Existence follows from

---

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

Proposition 3.

We now calculate $\overline{b^{F2}}$ and $\overline{b^{M2}}$ for $C \geq 1$. Note that in the correlated part, $b_i^M = b_i^F$ implies that $b_i^M < Cb_i^F$ and thus if one player gets a correlated draw she forms part of the F network).

$$
\begin{aligned}
\overline{b^{F2}} &= E\left(b_i^{F2}\,\middle|\,b_i^M < Cb_i^F\right) = p\frac{\int_1^\infty \int_1^{Cb_F} b^{F2}\frac{\alpha_F}{b^{F\alpha_F+1}}\frac{\alpha_M}{b^{M\alpha_M+1}}db^M\,db^F}{\int_1^\infty \int_1^{Cb_F}\frac{\alpha_F}{b^{F\alpha_F+1}}\frac{\alpha_M}{b^{M\alpha_M+1}}db^M\,db^F} + (1-p)\int_1^\infty b^{F2}\frac{\alpha_F}{b^{F\alpha_F+1}}db^F \\
&= p\frac{\alpha_F}{\alpha_F-2}\frac{((\alpha_F+\alpha_M-2)\,C^{\alpha_M}-(\alpha_F-2))}{\left(C^{\alpha_M}-\frac{\alpha_F}{\alpha_F+\alpha_M}\right)(\alpha_F+\alpha_M-2)} + (1-p)\int_1^\infty \frac{\alpha_F}{b^{F\alpha_F-1}}db^F \\
&= p\frac{\alpha_F}{\alpha_F-2}\frac{((\alpha_F+\alpha_M-2)\,C^{\alpha_M}-(\alpha_F-2))}{\left(C^{\alpha_M}-\frac{\alpha_F}{\alpha_F+\alpha_M}\right)(\alpha_F+\alpha_M-2)} + (1-p)\frac{\alpha_F}{\alpha_F-2}
\end{aligned}
$$

which can be rewritten as

$$
\overline{b^{F2}} = p\,(\alpha_F+\alpha_M)\frac{\alpha_F}{(\alpha_F-2)}\frac{(\alpha_F+\alpha_M-2)\,C^{\alpha_M}-(\alpha_F-2)}{((\alpha_F+\alpha_M)\,C^{\alpha_M}-\alpha_F)(\alpha_F+\alpha_M-2)}+(1-p)\frac{\alpha_F}{\alpha_F-2}
\tag{47}
$$

while

$$
\begin{aligned}
\overline{b^{M2}} &= E\left(b_i^{M2}\,\middle|\,b_i^M > Cb_i^F\right) \\
&= p\frac{\frac{\alpha_M\alpha_F}{(\alpha_M-2)C^{\alpha_M-2}}\frac{1}{(\alpha_F+\alpha_M-2)}}{\frac{\alpha_F}{C^{\alpha_M}}\frac{1}{\alpha_F+\alpha_M}}
\end{aligned}
$$

which simplifies to

$$
\overline{b^{M2}} = p\frac{\alpha_M}{(\alpha_M-2)}\frac{\alpha_F+\alpha_M}{(\alpha_F+\alpha_M-2)}C^2
\tag{48}
$$

$\overline{b^{M2}}$ is obviously increasing in C. We now show that $\overline{b^{F2}}$ is decreasing in C

$$
\frac{\partial \overline{b^{F2}}(C)}{\partial C} = -\frac{(\alpha_F+\alpha_M)}{(\alpha_F+\alpha_M-2)}\frac{\alpha_F}{(\alpha_F-2)}\left(\frac{2\alpha_M^2 C^{\alpha_M-1}}{((\alpha_F+\alpha_M)\,C^{\alpha_M}-\alpha_F)^2}\right) < 0
\tag{49}
$$

We first prove uniqueness of C_E defined by

$$
C_E = \frac{d^F}{d^M}\sqrt{\frac{1-a^M d^M \overline{b^{M2}}^2}{1-a^F d^F \overline{b^{F2}}^2}}
\tag{50}
$$

Note that the LHS of (50) 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 $\overline{b^{M2}}$ is increasing in C_E. Since $\overline{b^{F2}}$ 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

$$
C_P = \sqrt{\frac{d^{F2}}{d^{M2}} \frac{\left(1 + \left(\frac{a^F}{2} d^{F2}\overline{b^{F2}}\right)^2\right)\left(1 - \left(\frac{a^M}{2} d^{M2}\overline{b^{M2}}\right)^2\right)^2}{\left(1 - \left(\frac{a^F}{2} d^{F2}\overline{b^{F2}}\right)^2\right)^2 \left(1 + \left(\frac{a^M}{2} d^{M2}\overline{b^{M2}}\right)^2\right)}} \tag{51}
$$

Again note that the LHS of (51) 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 $\left(1 - \left(\frac{a^M}{2} d^{M2}\overline{b^{M2}}\right)^2\right)^2 / \left(1 + \left(\frac{a^M}{2} d^{M2}\overline{b^{M2}}\right)^2\right)$ is decreasing in C_P because $\overline{b^{M2}}$ is increasing in C_P. Also, since we showed in (49) that $\overline{b^{F2}}$ is decreasing in C_P then $\left(1 + \left(\frac{a^F}{2} d^{F2}\overline{b^{F2}}\right)^2\right) / \left(1 - \left(\frac{a^F}{2} d^{F2}\overline{b^{F2}}\right)^2\right)^2$ is decreasing C_P. As a result RHS of (51) is decreasing in C_P and the result follows.

A.7   Proof of Proposition 7

We will prove the proposition for $\alpha_F = \alpha_{M.}$. We will first show that

Lemma 5. $C_E > 1 \Leftrightarrow a^{M2} d^{M2} < a^{F2} d^{F2}$

Note also that if $a^{M2} d^{M2} = a^{F2} d^{F2}$ the solution of (50) is at $C^E = 1$. An increase of $a^{M2} d^{M2}$ with respect to $a^{F2} d^{F2}$ 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 $\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} > 0$ or too many people $\left.\frac{\partial w(C)}{\partial C}\right|_{C=C_E} < 0$ in the F group compared to the social optimum.[52] The F group will be underpopulated if and only if

$$
\frac{a^{M2} d^{M2}}{a^{F2} d^{F2}} > \frac{\left((2\alpha - 2) C^\alpha - (\alpha - 2)\right)^2 C^\alpha}{\alpha^2 \left(2C^\alpha - 1\right)^3 C^2} \tag{52}
$$

---

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

We will check how a decentralized group choice deviates from the efficient group choice C^{sp} 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

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

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

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

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

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

Therefore

$$\frac{\partial w(C)}{\partial C}\bigg|_{C=C_E} > 0 \iff -a^{F^2}d^{F^2}\frac{((2\alpha-2)\,C^\alpha-(\alpha-2))^2}{(2C^\alpha-1)^3\,C}+a^{M^2}d^{M^2}\frac{\alpha^2}{C^{\alpha-1}} > 0$$

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

and

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

By Lemma 5 since $C_E > 1 \iff a^{M^2}d^{M^2} < a^{F^2}d^{F^2}$, hence $\frac{a^{M^2}d^{M^2}}{a^{F^2}d^{F^2}} < 1$.

We will now show that

$$1 > \frac{((2\alpha-2)\,C^\alpha-(\alpha-2))^2\,C^\alpha}{\alpha^2\,(2C^\alpha-1)^3\,C^2} = \frac{((\alpha-1)\,(2C^\alpha-1)+1)^2\,C^\alpha}{\alpha^2\,(2C^\alpha-1)^3\,C^2} \qquad (53)$$

Note that

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

since that expression is equivalent to

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

thus

$$\frac{((\alpha-1)\,(2C^\alpha-1)+1)^2\,C^\alpha}{\alpha^2\,(2C^\alpha-1)^3\,C^2} < \frac{C^\alpha}{(2C^\alpha-1)\,C^2} < \frac{1}{C} < 1 \qquad (54)$$

where the last two inequalities hold since $C > 1$, noting that in that case $2C^\alpha - 1 > C^\alpha$. Thus equation (54) establishes (53).

The next two lemmas establish that overpopulation can occur in both sectors and depends on the underlying parameters. Lemma 6 shows the existence of parameter values that $\frac{\partial w(C)}{\partial C}\big|_{C=C_E} < 0$ while Lemma 7 shows the existence of parameter values that $\frac{\partial w(C)}{\partial C}\big|_{C=C_E} > 0$.

Lemma 6. *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

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

we have that

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

thus

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

Lemma 7. *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*

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

For a bounded C_E

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

Hence

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

and thus

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

Now since

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

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

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

23

or

$$\frac{-\left(1 - a^{F^2} d^{F^2}\right) + \sqrt{\left(1 - a^{F^2} d^{F^2}\right)^2 + 4r a^{F^2} d^{F^2}}}{2r a^{F^2}} > \frac{1}{2r} \qquad (55)$$

$$\sqrt{\left(1 - a^{F^2} d^{F^2}\right)^2 + 4r a^{F^2} d^{F^2}} > 1$$
$$a^{F^2} d^{F^2} \left(a^{F^2} d^{F^2} + 4r - 2\right) > 0$$

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

Proposition 7 immediately follows from these Lemmas.

A.8   Proof of Proposition 8

We will prove this Proposition for a slightly more general case, using the same correlation structure as in the proof of Proposition. 6. With probability p the two values of b_i^j for $j \in \{F, M\}$ are independent of one another. With probability $(1 - p)$, $b_i^M = b_i^F$ and they are distributed with shape parameter α_F. With this correlation structure $\overline{b^{F^2}}$ and $\overline{b^{M^2}}$ have been calculated by (47) and (48) respectively, as

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

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

We now normalize $\overline{b^{M^2}}$ by the expected second moment $\frac{\alpha_M}{(\alpha_M - 2)}$. Hence

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

Clearly, this is decreasing in α_F and α_M.

We normalize $\overline{b^{F^2}}$ by the expected second moment $\frac{\alpha_F}{(\alpha_F - 2)}$. Hence

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

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

24

Noting that

$$
\begin{aligned}
\frac{\partial \overline{b^{F2}}_{NORM}}{\partial \alpha_F} &= p\frac{1}{C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}}\left(\frac{-1}{\alpha_F + \alpha_M - 2} + \frac{\alpha_F - 2}{(\alpha_F + \alpha_M - 2)^2}\right)\\
&\quad - p\frac{1}{\left(C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}\right)^2}\left(\frac{-1}{\alpha_F + \alpha_M} + \frac{\alpha_F}{(\alpha_F + \alpha_M)^2}\right)\\
&\quad \times \left(C^{\alpha_M} - \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2}\right)\\
&= p\frac{1}{C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}}\left(\frac{-\alpha_M}{(\alpha_F + \alpha_M - 2)^2}\right) - p\frac{C^{\alpha_M} - \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2}}{\left(C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}\right)^2}\left(\frac{-\alpha_M}{(\alpha_F + \alpha_M)^2}\right)\\
&= p\frac{\alpha_M}{C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}}\left(\left(\frac{-1}{(\alpha_F + \alpha_M - 2)^2}\right) + \frac{C^{\alpha_M} - \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2}}{\left(C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}\right)}\left(\frac{1}{(\alpha_F + \alpha_M)^2}\right)\right)
\end{aligned}
$$

we want to show that

$$
\frac{\partial \overline{b^{F2}}_{NORM}}{\partial \alpha_F} < 0.
$$

A sufficient condition is

$$
\left(\frac{-1}{(\alpha_F + \alpha_M - 2)^2}\right) + \frac{C^{\alpha_M} - \frac{\alpha_F - 2}{\alpha_F + \alpha_M - 2}}{\left(C^{\alpha_M} - \frac{\alpha_F}{\alpha_F + \alpha_M}\right)}\left(\frac{1}{(\alpha_F + \alpha_M)^2}\right) < 0
$$

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

or

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

$$
\tag{56}
$$

We will now show that

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

is increasing in α_F and given that $\partial \overline{b^{F2}}_{NORM}/\partial \alpha_F < 0$ is equivalent to expres-

sion (56) the result follows. Then

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

Now note that

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

Now

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

then

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

requires

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

which is equivalent to

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

26

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