The Allocation of Scientific Talent*

Andrea Canidio

IMT School of Advanced Studies, Lucca (Italy) and INSEAD, Fontainebleau (France).
Email: andrea.canidio@imtlucca.it.

Abstract

I consider a model in which firms produce new knowledge by building labs and hiring researchers in a competitive market. I show that, for given distribution of labs, the allocation of researchers to firms may be efficient or inefficient depending on how fast the firms' marginal return on the knowledge produced decreases with the amount of knowledge produced. I then argue that the allocation of researchers to labs is likely to be inefficient if firms invest in R&D primarily to increase their *absorptive capacity*, that is, their ability to use the stock of publicly available knowledge. When the distribution of labs is endogenous, a second source of inefficiency arises: firms' underinvestment in labs. Policies subsidizing the investment in labs are ineffective at restoring the first best, unless policies aimed at reallocating researchers to firms are also put in place.

Keywords: Labor Market for Researchers, Matching, R&D Productivity, Organization of Scientific Research, Absorptive Capacity, Innovation Policy.
JEL Numbers: J44, J21, O32, L22, O31, L10.

*I'm grateful to Martin Cripps, Lucia Esposito, Jeffrey Furman, Thomas Gall, Alfonso Gambardella, Joshua Gans, Miklos Koren, Patrick Legros, Megan MacGarvie, Dilip Mookherjee, Andrew Newman, Zvika Neeman, Julie Wulf, Scott Stern, Tomas Sjöström, Adam Szeidl, Tim Van Zandt and three anonymous referees for helpful discussions and constructive comments. This paper initially circulated under the title "The Production of Science."

1 Introduction

Despite headline-grabbing stories about elite researchers leaving one country for another or one firm for another,[1] very little is known about the efficiency properties of the labor market for researchers. A firm that hires a very talented researcher increases its R&D output and profits, but may cause other firms to settle for someone with lower scientific talent. A frictionless labor market matches each researcher to the firm that values him or her the most. However, in the presence of positive externalities such as knowledge spillovers, there is no presumption that the equilibrium allocation of researchers to firms also maximizes the positive externality that each firm generates. This paper discusses whether and under what conditions a trade-off exists between the assignment of researchers to firms maximizing the private value of knowledge production and the one maximizing the social value of knowledge production.

I consider a theoretical model in which several firms produce new knowledge by building labs and hiring researchers in a frictionless labor market. Researchers are a non-homogeneous input in the production of knowledge and are characterized by their *ability*. Each lab is an aggregate of all other inputs in the production of new knowledge (e.g., machines, technicians, raw material), and is characterized by its *size* that is, the size of the investment required to build it. The knowledge produced within a firm increases with the ability of the researcher hired and the size of the lab built.[2] Furthermore, I assume that researcher's ability and lab size are complements in the knowledge production function, so that total knowledge produced in the economy is maximized under a Positive Assortative Matching (PAM) rule, assigning the most productive researcher to work in the largest lab. Because of knowledge spillovers, each firm benefits from the knowledge produced by all other firms.

As a first step to my analysis, I take the distribution of labs as given and show that the market allocation of researchers to firms depends on how fast the firms' marginal return

[1] For example, see "Steal This Scientist" M. Liu, *Newsweek,* 14th of November 2009 (accessed from http://www.thedailybeast.com/newsweek/2009/11/13/steal-this-scientist.html). See also "Climbing Mount Publishable: The Old Scientific Powers Are Starting to Lose Their Grip" *The Economist,* 11th of November 2010 (accessed from http://www.economist.com/node/17460678).

[2] Lab size may be determined by the lab physical size, by the number of people (technicians, post-docs) working in it, but also by the quality and productivity of its machines and staff. However, for the sake of clarity, I use terms related to quantity (i.e., size) when referring to labs, and to quality (i.e., ability) when referring to researchers.

on its research activities decreases with the amount of knowledge produced. When the marginal return decreases slowly, researchers and labs are complements from the firm's point of view. It follows that the market allocation of researchers to labs is PAM, which also maximizes aggregate knowledge production. On the other hand, if the marginal return decreases fast, the two inputs in the production of new knowledge—labs and researchers—may be substitutes in the firms' objective function. In this case, the market allocation of researchers to labs is Negative Assortative Matching (NAM), in which the most productive researcher in the economy is hired by the firm with the smallest lab. It follows that, for any given distribution of labs, the private sector minimizes the amount of knowledge produced: the decentralized allocation of researchers to labs is inefficient.

Furthermore, if firms' return on knowledge production is an iso-elastic function, then the equilibrium allocation is NAM or PAM depending on whether this function is bounded or unbounded above. Based on this observation, I argue that the allocation of researchers to firms is likely to be inefficient if firms' R&D activities are motivated by *absorptive capacity*, which is the ability to find, understand, and use available public knowledge. For example, it is well understood that R&D provides "a ticket of admission to an information network" (Rosenberg, 1990; p. 170), and allows firms to become immediately aware of the latest discoveries and inventions. Also, new knowledge may be difficult to understand and use in a timely manner, unless a firm employs researchers who are familiar with the current scientific and technological frontier. Producing new knowledge, therefore, causes an increase in a firms' absorptive capacity (Cohen and Levinthal, 1989).[3] If this increase in absorptive capacity is the only benefit generated by R&D activities, then, quite naturally, the private benefit from producing knowledge has an upper bound given by the benefit of absorbing all publicly available knowledge. In this case, assuming that the return on the production of knowledge is iso-elastic, we should expect the allocation of researchers to labs to be inefficient.[4]

When the distribution of labs is endogenous, an additional source of inefficiency arises. Each firm invests in its lab taking as given the researcher it will be matched with and the

[3] On the empirical relevance of absorptive capacity, see Tilton (1971); Cockburn and Henderson (1998); Gambardella (1992); Griffith, Redding, and Van Reenen (2003, 2004).

[4] Other mechanisms can generate a bounded private return on knowledge production and a misallocation of scientific talent in equilibrium. For example, when experimenting in a lab failures can be socially valuable because of learning. However, only successes are valuable from the firm point of view, generating a private return on knowledge that can be very curved (see the conclusion of the paper for further details). Here I focus on absorptive capacity because there is ample empirical evidence showing its importance for firms.

stock of public knowledge available in the economy. This investment generates a positive externality on other firms, which implies that, in equilibrium, firms build labs that are smaller than the socially optimal size. Subsidies to firms' investment in labs, however, cannot restore efficiency because they do not affect the labor market for researchers. The first best can be achieved by simultaneously subsidizing firms' investments in labs and reallocating researchers to firms.

The fact that the best researchers may not be hired by the firm with the largest labs finds empirical support in the literature on firms' size and R&D productivity, which shows that small firms are often more productive than large firms in their R&D efforts. In the model, the productivity of a lab is determined by the ability of the researcher hired. Hence, when absorptive capacity is the dominant motive for science production the model predicts a negative relationship between lab size and lab productivity. If large firms have a comparative advantage at building labs (because of, for example, economies of scopes among different research lines), they should build larger labs compared to smaller firms. If the equilibrium matching pattern is NAM, we should therefore expect a negative relationship between lab productivity and firm size. By looking at the number of patents produced in firms of different sizes, several authors find results that support this prediction.[5] More directly, Elfenbein, Hamilton, and Zenger (2010) examine the allocation of researchers to firms and show that productive R&D workers are more likely to work for small firms than for big firms.

Finally, I extend the model by assuming that producing new knowledge has an additional benefit: it increases a researcher's reputation. Because of a cash constraint, researchers cannot transfer the value of reputation back to firms. I show that, if reputation concerns are strong enough, the equilibrium in the private sector may switch from NAM to PAM. Intuitively, when researchers' reputation concerns are strong enough, they may receive most of their compensation in the form of reputation rather than a monetary payment. Because reputation increases with the amount of knowledge produced, researchers prefer to work for firms with large labs. Hence, reputation may cause productive researchers to work for firms with large labs, therefore increasing the knowledge produced in the economy and total welfare.

[5] See Scherer (1965), Acs and Audretsch (1987), and Cohen and Klepper (1996) who review the empirical evidence.

1.1 Literature

Several papers have argued that equilibrium matching patterns may be inefficient and that re-matching policies may increase welfare. However, in most of the existing literature, inefficient matching patterns emerge because peer effects and cash constraints prevent agents from efficiently sharing the surplus generated within each match (see, for example, Estevan, Gall, Legros, and Newman, 2017, who consider the problem of students sorting into schools). In my model, instead, surplus can be freely shared among firms and researchers by using wages. Nonetheless, the presence of widespread externalities in the form of knowledge spillovers may lead to an inefficient matching pattern. Furthermore, in the presence of widespread externalities, restricting the ability of agents to share surplus may lead to a welfare improvement (see Section 4, where I introduce reputation concerns).

Matching markets with widespread externalities have been studied theoretically by Hammond, Kaneko, and Wooders (1989), Hammond (1995), and Kaneko and Wooders (1996). Hammond et al. (1989) and Kaneko and Wooders (1996) show the existence of the competitive equilibrium in large economies with widespread externalities. Hammond (1995) show existence of the equilibrium in economies with widespread externalities and different types of taxes or subsidies. My paper studies the effect of widespread externalities in an applied, policy-relevant matching problem. Also, to the best of my knowledge, it is the first one to study a matching model in which there are both widespread externalities and a pre-matching investment phase similar to Cole, Mailath, and Postlewaite (2001).

There is a large literature examining the determinants and the efficiency consequences of the allocation of talent to different tasks and professions.[6] Closest to my paper are Acemoglu, Akcigit, and Celik (2014) and Vandenbussche, Aghion, and Meghir (2006). In Acemoglu et al. (2014), managers can pursue either radical innovation or incremental innovation. In Vandenbussche et al. (2006), skilled and unskilled workers can be employed either in innovative activities (i.e. pushing the knowledge frontier) or in the adoption of new technology (i.e. reducing the distance to the knowledge frontier). Here, instead, I analyze the allocation of scientific talent within the same sector and the same activity, and not the allocation of scientific talent across sectors and activities. In particular, following the literature on absorptive capacity in my model creating new knowledge and reducing the distance

[6] See for example Murphy et al. (1991); Baumol (1996); Acemoglu, Aghion, and Zilibotti (2006) and, related to knowledge production, Jovanovic and Rob (1989) and Lucas Jr and Moll (2014).

to the knowledge frontier are byproducts of the same activity. Therefore, at a firm level there is no trade-off between innovation and adoption of new knowledge, because they are both increasing in the amount of R&D carried out. Despite this, a trade-off between the two activities emerges endogenously in the assignment of researchers to labs.

Hammerschmidt (2009), Kamien and Zang (2000) and Leahy and Neary (2007) study the theoretical implications of absorptive capacity, and show that it generates a new set of strategic considerations for firms. I also study the theoretical implications of absorptive capacity, but I abstract away from strategic considerations because each firm is assumed to be small relative to the number of firms active on the labor market for researchers. My work is also part of a growing literature recognizing that the inputs in the research process are non-homogeneous, and that the way these inputs are combined is a key determinant of research outcomes. Most papers in this literature are concerned with the determinants of research team size (see, for example, Bikard, Murray, and Gans, 2015). Here I keep the number of inputs in the production of knowledge fixed, and I argue that the sorting of research inputs matters for aggregate research output.

The remainder of the paper proceeds as follows. In the next section, I present a version of the model in which the distribution of labs is exogenously given. In the third section, I derive the distribution of labs endogenously. In the fourth section I introduce reputation concerns. In the last section, I conclude. All proofs missing from the text are in appendix.

2 The Model

The economy is populated by a continuum of heterogeneous firms and a continuum of heterogeneous researchers. Firms differ in the size (or productivity) of their lab L which is continuously distributed over $\mathcal{L} = [\underline{L}, \bar{L}]$ with $p.d.f.$ $h(L)$. The distribution of labs is taken as given here, but it is derived endogenously in Section 3. Researchers differ in their ability a which is continuously distributed over $\mathcal{A} = [0, \bar{a}]$ with $p.d.f.$ $z(a)$.

Firms' objective is to maximize profits, while researchers' objective is to maximize their wage. All agents have the same outside option, which is assumed to be zero. The economy runs for two periods. In the first period each firm hires one researcher and produces new knowledge; in the second period profits are realized and wages are paid.

Period $t = 1$: producing new knowledge. In period $t = 1$, each firm hires one researcher to work in the firm's lab. The amount of knowledge produced within each lab is:

$$k = af(L),$$

where $f(L) \geq 0$, $f(0) = 0$, $f'(L) > 0$, and $f''(L) \leq 0$.

The reader should interpret the lab size L as inclusive of all inputs that can increase the chance of a scientific or technological discovery for given researcher's ability, including machines (e.g. a bigger telescope, a more powerful microscope and a state-of-the-art DNA sequencing machine). Also, researchers typically work in teams. This fact can be incorporated into the model by defining a as the research team's average quality. A previous matching stage determines how researchers form research teams, and how the distribution of a is determined from the distribution of individual ability. To keep the model as simple as possible, I do not pursue this interpretation further.

Period $t = 2$: wages and profits. Call V the stock of public knowledge, available in the economy at the beginning of period 2, and taken as given by firms and researchers. The surplus generated withing a firm depends on the knowledge produced in-house and on the stock of public knowledge, so that:

$$\Phi_V(a, L) = \phi(k, V)$$

$$\text{s.t. } k = af(L),$$

where $\phi(,)$ is increasing in both arguments, with $\frac{\partial^2 \phi(k,V)}{\partial k^2} < 0$ and $\phi(k = 0, V) = 0$ for all V. For example, $\phi(k, V)$ may represent the revenues earned by inventing new products and bringing them to market. The surplus generated is then split between a wage to the researcher, and profits to the firm:

$$\phi(k, V) = w(a) + \pi(L)$$

where wages for each ability level $w(a) : \mathcal{A} \rightarrow \mathbb{R}^+$ and profits as a function of lab size $\pi(L) : \mathcal{L} \rightarrow \mathbb{R}^+$ are endogenous and determined in equilibrium.

Note that this specification abstracts away from competition in the product market. The reason is that firms may apply the same knowledge to very different products. For example, all biotech firms benefit from the same scientific knowledge. However, some firms develop DNA sequencing machines, some develop drugs and others develop bacteria that can produce biofuel out of garbage. Some firms compete with each other, some complement each other, and some belong to very different product markets. For this reason, I do not consider firms' competition on the product market and focus on their competition for researchers.

Endogenous Science. The stock of public knowledge is taken as given by firms and researchers but is determined endogenously, by aggregating all the knowledge produced within each firm:

$$V = \nu \int_{\underline{L}}^{\overline{L}} m(L) f(L) h(L) dL, \tag{1}$$

where $m(L) : \mathcal{L} \cup 0 \to \mathcal{A} \cup 0$ assigns labs to researchers and is determined in equilibrium, with $m(L) = 0$ representing an unmatched lab and $m(0) = a$ representing an unmatched researcher. $\underline{L} > 0$ is the smallest lab that is matched, and is also determined in equilibrium.[7] Finally, the parameter $\nu \geq 0$ captures both the size of knowledge spillovers, and the possibility of complementaries between knowledge produced within different firms.

Note that, in general, it is possible that $m(L)$ is a correspondence: two researchers with different abilities are assigned to the same lab. We argue below that, in equilibrium, this is possible only for a lab of size 0. That is, in equilibrium, two researchers of different abilities may not be matched, but conditional on being matched there is a one-to-one correspondence between researchers and labs. The notation used anticipates the fact that $m(L)$ is a function over the range of integration of (1).

The above specification implies that ability a and lab size L are complements in the knowledge production function. In turn, this implies that the allocation of researchers to labs that maximizes the aggregate production of knowledge is Positive Assortative Matching (PAM), in which the most productive researcher is allocated to the biggest lab. This is a key assumption that I will maintain throughout the paper. It is justified by the observation that, when research funds are allocated by grant-giving institutions having the explicit goal

[7] If the distribution of labs is exogenous, it is possible that some labs are unmatched and therefore idle in equilibrium. However, when the distribution of labs is derived endogenously (see Section 3) no labs will be unmatched in equilibrium, because a firm that anticipates not being able to hire a researcher will not build a lab.

of producing new knowledge, researchers who can demonstrate higher ability and better past performance typically have a higher chance of receiving funds and are allocated larger grants.[8]

Finally, the knowledge produced by a given firm may interact with the knowledge produced by other firms and determine the stock of public knowledge in complex ways. As a consequence, a more general functional form such as $V = \int_{\underline{L}}^{\overline{L}} \varrho\left(m(L)f(L)\right)h(L)dL$ may be desirable. As long as $\varrho()$ is not too curved, the two inputs in the knowledge-production process remain complements. However, because the main results of the paper do not depend on the specific shape of $\varrho()$, for ease of exposition I only consider the linear case.[9]

Welfare. The total welfare of the economy is equal to the sum of all profits and all wages earned, which is also equal to the sum of the surplus generated in each match:

$$\int_{\underline{L}}^{\overline{L}} \phi(m(L)f(L), V)h(L)dL. \tag{2}$$

The key observation is that, whereas firms and researchers take the total stock of knowledge V as given, from the social point of view V is determined endogenously and depends on the equilibrium matching pattern.

2.1 Equilibrium

I first consider the equilibrium on the labor market for researchers for given V.

Definition 1 (Partial Equilibrium)**.** For given V, the functions $m(L)$, $\pi(L)$ and $w(a)$ constitute a partial equilibrium if the following two properties hold:

[8] For example, Arora and Gambardella (2005) analyze the funding allocation decisions of NSF and show that the reputation (past publication record) is positively correlated with the probability of being awarded the grant and with the size of the grant. Arora, David, and Gambardella (1998) analyze the funding allocation decisions of the Italian CNR (equivalent to NSF) and find similar results. They also show that the elasticity of research output with respect of research budget is higher then average for prestigious scientists, which implies complementarity between ability and funds.

[9] The results do not depend on $\varrho()$ in the sense that, for any $\varrho()$, if the marginal private-return on knowledge production decreases fast enough, a and L are substitutes in the private surplus function (see Lemma 1), and the equilibrium allocation may be inefficient (see Proposition 1). However, the exact curvature of $\phi(.,.)$ determining whether a and L are substitutes in the private surplus function depends on $\varrho()$.

- Feasibility: $\pi(L) + w(a) \leq \Phi_V(a, L)$ $\forall L \in \mathcal{L} \cup 0$ and $a \in m(L)$.
- Stability: $\pi(L) + w(a') \geq \Phi_V(a', L)$ $\forall L, L' \in \mathcal{L} \cup 0$ and $a' \in m(L')$.[10]

To understand the implication of this definition, note that feasibility and stability imply that

$$\Phi_V(m(L), L) - w(m(L)) \geq \Phi_V(a', L) - w(a') \ \forall L, L' \in \mathcal{L} \cup 0 \text{ and } a' \in m(L').$$

In other words, given V and the equilibrium wage function $w()$, each firm "shops for researchers" in a competitive market.[11] Profits are equal to the difference between the total surplus and the wage paid. In equilibrium, each firm prefers to hire the researcher specified by $m()$ rather than any other researcher. For any given V, the existence and uniqueness of the equilibrium allocation $m(L)$ is a standard result in matching theory. Also, because L and a have continuous distributions, the equilibrium matching function $m(L)$ is a bijection for $m(L) > 0$: each L matched is matched with one and only one a. If in equilibrium some researchers and some firms are not matched, then also $w(a)$ and $\pi(L)$ are unique. If instead all researchers and all firms are matched, then there are multiple equilibrium $w(a)$ and $\pi(L)$, depending on the payoff of the least productive researcher and of the firm with the smallest lab. The payoff functions $w(a)$ and $\pi(L)$ are continuous (for all these results see, for example, Kamecke, 1992). Finally, for future reference note that $m(L) > 0$ is a bijection from a closed interval (the set of labs that are matched) to a closed interval (the set of researchers that are matched). Hence, conditioned on $L > 0$, if $m(L)$ is monotonic it must also be continuous.

Feasibility and stability define the equilibrium in the labor market for researchers for given V. However, the stock on knowledge available in the economy is itself a function of the matching function $m(L)$, which implies the following definition.

Definition 2 (General Equilibrium)**.** A given V, $m(L)$, $\pi(L)$ and $w(a)$ constitute a general equilibrium if feasibility and stability hold, and

$$V = \nu \int m(L) f(L) h(L) dL.$$

[10] Remember that, in general, $m(L)$ could be a correspondence: two researchers with different abilities are matched with the same lab. In equilibrium, this will happen for lab of size 0.

[11] Note that stability and feasibility also imply that $\Phi_V(m(L), L) - \pi(L) \geq \Phi_V(m(L), L') - \pi(L')$. In other words, we could equivalently imagine that researchers shop for firms.

The existence of the general equilibrium is shown in Kaneko and Wooders (1996), who consider a two-sided matching model with a continuum of players and widespread externalities. The general equilibrium may not be unique. The reason is that when a large V is expected, more firms and more researchers will match, and the total knowledge produced will be high. Instead, when a low V is expected, fewer firms and researchers will match, and the total knowledge produced will be low. Note also that equilibria with larger V Pareto dominate equilibria with lower V. When discussing the efficiency property of the equilibrium, I will always implicitly assume that general equilibrium of the economy is the Pareto preferred one.

2.2 Solution

Despite the fact that ability and lab size are complements in the knowledge production function, whether they are complements or substitutes in the private surplus function depends on the sign of

$$\frac{\partial^2 \phi(af(L),V)}{\partial a \partial L} = \frac{\partial \phi(af(L),V)}{\partial k} f'(L) + \frac{\partial^2 \phi(af(L),V)}{\partial k^2} af(L)f'(L).$$

The first term on the RHS of the above expression is positive because ability and lab size are complements in the production of knowledge. The second term is negative whenever there are decreasing returns to knowledge creation. Thus, if the marginal return on knowledge decreases fast enough, ability and lab size are substitutes in the surplus function.

To understand why, consider a simplified version of the model in which there are only two firms L', L'' with $L' > L''$, and two researchers a', a'' with $a' > a''$. Because a and L are complements in the knowledge production function, the increase in knowledge production when hiring a' instead of a'' is larger for firm L' than firm L''. However, if the marginal benefit from producing knowledge decreases fast enough, the increase in knowledge will be more beneficial to L'', which is the firm producing less knowledge when matched with either researcher. It follows that the firm with the smallest lab benefits the most from matching with a productive researcher. If instead the marginal benefit from producing knowledge decreases slowly enough, the firm with the largest lab benefits the most from matching with a productive researcher.

The following lemma shows that a and L are local substitutes or local complements depending on whether $\frac{\partial \phi}{\partial k}$ decreases faster or slower than $\frac{\partial \log(k)}{\partial k}$.

Lemma 1. *Consider the equilibrium matching function $m(L)$ and a given $\hat{L}$ such that $m(L)$ is differentiable at $\hat{L}$. Define $\hat{k} = m(\hat{L})f(\hat{L})$.*

- *the equilibrium is local NAM (i.e. $m'(\hat{L}) < 0$) if and only if for k in the neighborhood of $\hat{k}$, $\phi(k, V) = \chi(log(k), V)$ where $\chi(.,.) : \mathbb{R}^2 \to \mathbb{R}$ is strictly increasing and* ***concave*** *in $log(k)$.*

- *the equilibrium is local PAM (i.e. $m'(\hat{L}) > 0$) if an only if for k in the neighborhood of $\hat{k}$, $\phi(k, V) = \chi(log(k), V)$ where $\chi(.,.) : \mathbb{R}^2 \to \mathbb{R}$ is strictly increasing and* ***convex*** *in $log(k)$.*

The following proposition derives conditions under which a and L are global substitutes or global complements in the private surplus function.

Proposition 1. *Define $\overline{k} \equiv \overline{a} f(\overline{L})$ as the maximum amount of knowledge that can be produced by any firm.*

1. *If $\phi(k, V) = \chi(log(k), V)\, \forall k \leq \overline{k}$ with $\chi(.,.) : \mathbb{R}^2 \to \mathbb{R}$ strictly increasing and* ***concave*** *in $\log(k)$, then in equilibrium $m'(L) < 0$ for every L that is matched (i.e.* ***global NAM****.) If ν is sufficiently large this equilibrium is inefficient: social welfare can be increased by implementing PAM somewhere.*

2. *If $\phi(k, V) = \chi(log(k), V)\, \forall k \leq \overline{k}$ with $\chi(.,.) : \mathbb{R}^2 \to \mathbb{R}$ strictly increasing and* ***convex*** *in $\log(k)$, then in equilibrium $m'(L) > 0$ for every L that is matched (i.e.* ***global PAM****.) This equilibrium is efficient: it is not possible to re-allocate researchers to labs and increase welfare.*

Hence, if the marginal return on knowledge decreases everywhere faster than a logarithmic function, the equilibrium matching pattern is global NAM, in which the least productive researcher among the ones who are hired is allocated to the largest lab. In this case, if ν is sufficiently large the equilibrium matching pattern is inefficient: welfare can be improved by reallocating some researchers.[12] On the other hand, when ν is close to zero, it is possible

[12] More precisely: because multiple equilibria are possible, if ν is sufficiently large any equilibrium matching pattern is inefficient.

that both in equilibrium and in the first best no researcher is hired and no knowledge is produced.

Conversely, if the marginal return on knowledge decreases everywhere slower than a logarithmic function, the equilibrium matching pattern is global PAM: the best researcher is allocated to the largest lab. In this case, firms and researchers are simultaneously maximizing their individual payoffs and the positive externality they generate. It follows that if PAM is the equilibrium allocation, it must also be the efficient allocation.

Note that, in general, the marginal return on knowledge may somewhere decrease faster than a logarithmic function and somewhere slower, and therefore we may be outside the cases considered by Proposition 1. When this happens, a and L are somewhere local complements, and somewhere local substitutes in $\phi(af(L), V)$. The equilibrium matching pattern will depend on the distribution of labs and researchers' ability, and may be PAM, NAM, or PAM over some range and NAM over some other range. In these cases, it is possible to derive the equilibrium only via numerical methods.

To summarize, the fact that knowledge is a public good may or may not imply that the allocation of researchers to labs is inefficient. The key determinant of whether welfare can be improved by re-allocating researchers to firms is the speed at which the marginal private return on k decreases with k.

2.3 First best

A natural question is how to restore efficiency. We can apply here the results in Hammond (1995, Section 6), who considers a large economy with widespread externalities and shows a combination of a Pigouvian subsidy and Lindahl taxes can achieve the first best allocation. In our framework, this result implies that the social planner should impose a Lindahl tax on the public knowledge consumed by firm i equal to $T_i = \frac{\partial \phi(V, af(L))}{\partial V}|_{V=V^*}$, that is firm i's marginal benefit of public knowledge V evaluated at its first best level V^*. Simultaneously, the social planner should impose a Pigouvian subsidy on firms' contribution to public knowledge equal to the average Lindahl tax $\tau = \int T_i h(L) d(L)$.

To understand the effect of T_i and τ, it is useful to think of the equilibrium as a situation in which firms "shop for researchers" in a competitive market. In this case, each firm's profit-maximization problem is

$$\{m(L), V\} = \operatorname{argmax}_{a,V} \{\phi(V, af(L)) + \tau \cdot \nu \cdot af(L) - T_i V - w(a)\}.$$

The Lindahl tax is the price at which the firm can purchase public knowledge V, and implies that

$$\frac{\partial \phi(V, af(L))}{\partial V} = T_i = \frac{\partial \phi(V, af(L))}{\partial V}|_{V=V^*}$$

or $V = V^*$. That is, the "price" for public knowledge T_i is such that each firm will demand exactly its first best-level V^*. It follows that, in the competitive equilibrium with Pigouvian subsidy and Lindahl taxes, it must be that $V = V^*$.[13]

Given this, and using the expression for the Pigouvian subsidy τ, the marginal benefit of knowledge production $k = af(L)$ is

$$\frac{\partial \phi(V, k)}{\partial k}|_{V=V^*} + \tau v = \frac{\partial \phi(V, k)}{\partial k}|_{V=V^*} + v \int \left(\frac{\partial \phi(V, af(L))}{\partial V}|_{V=V^*} \right) h(L) d(L)$$

which is equal to the marginal *social* benefit of knowledge production *evaluated at* V^* (see equation 2). It follows that each firm internalizes the benefit that its own investment in knowledge generates on other firms, under the expectation that all other firms will contribute the first-best level of knowledge.[14] Finally, note that this scheme is budget-neutral, because the amount paid as Pigouvian subsidy equals the amount collected by the Lindahl taxes.

The equilibrium matching pattern now depends on whether a and L are complements or substitutes in

$$\phi(V, af(L)) + \tau \cdot \nu \cdot af(L).$$

If a and L are complements in $\phi(V, af(L))$, then they are complements also in the above expression and, as we already know, PAM is the efficient allocation. If a and L are substitutes in $\phi(V, af(L))$, from the social point of view a and L may be somewhere complements and somewhere substitutes. The optimal allocation of researchers to labs exists and is unique,

[13] The existence of the competitive equilibrium with Pigouvian subsidy and Lindahl taxes is in Hammond (1995).

[14] Note an important point. The Pigouvian subsidy guarantees that for given V the equilibrium allocation is efficient conditional on V. But this gives rise to the possibility of multiple equilibria, some of them with an inefficient level of V. Lindahl taxes guarantee that the equilibrium is unique at $V = V^*$. See also Hammond (1995, Section 7) for a full discussion on this point.

but can only be derived numerically and may involve implementing PAM over some range and NAM over some other range. Intuitively, the social planner may balance production of knowledge and profit maximization by implementing PAM across some firm and NAM across some other firms. Of course, the social planner can achieve the first best also by directly allocating researchers to firms.

2.4 Discussion: absorptive capacity

Is it reasonable to expect the equilibrium matching pattern in the labor market of researchers to be NAM? To explore this question, I restrict my attention to $\phi(k, V)$ that are isoelastic for all V, so that

$$\frac{\partial \phi(k, V)}{\partial k} \cdot \frac{k}{\phi(k, V)} \equiv \sigma \ \forall k, V. \tag{3}$$

Under this restriction, if $\sigma < 0$ the equilibrium allocation is NAM, while if $\sigma > 0$ the equilibrium allocation is PAM. Importantly, this does not depend on the distribution of labs and researchers. Hence, (3) guarantees that the shape of $\phi(k, V)$ is the only determinant of the equilibrium matching pattern. Note also that $\sigma < 0$ is equivalent to $\phi(k, V)$ being bounded above in k, while $\sigma > 0$ is equivalent to $\phi(k, V)$ being unbounded above in k. Hence, under (3) the above question can be restated as: are there reasons to believe that firms' private benefit of knowledge production is bounded above?

One possible reason is absorptive capacity. According to Rosenberg (1990, p. 170):

> *Knowledge is regarded by economists as being "on the shelf" and costlessly available to all comers once it has been produced. But this model is seriously flawed because it frequently requires a substantial research capability to understand, interpret and to appraise knowledge that has been placed upon the shelf.*

A large literature has since argued that firms' absorptive capacity—that is, their ability to use the available stock of science—increases with the amount of knowledge produced internally (see Cohen and Levinthal, 1989, and the papers discussed in Section 1). This view implies that, if absorptive capacity is the only motive for knowledge production, then $\phi(k, V)$ has an upper bound in k given by the limit case in which the stock of knowledge V is costlessly available to the firm. In this case, therefore, we should expect the equilibrium matching pattern in the labor market of researchers to be NAM.

Of course, producing knowledge may be beneficial to a firm beyond the increase in its absorptive capacity. For example, the knowledge produced may be directly beneficial to the firm, in which case there is no presumption that $\phi(k, V)$ is bounded above. In general, the shape of $\phi(k, V)$ depends on several elements that are not part of the model, such as each firm's market share, the IP regimes, the market structure. Whether in a given sector $\phi(k, V)$ is bounded above is ultimately an empirical question. However, the literature discussed in the introduction has found evidence of NAM in the labor market for researchers. In light of the model, absorptive capacity is a plausible mechanism to explain this empirical fact.

3 Endogenous labs

In this section I add a period in which firms invest in labs. I will show that, taking as given the researchers they will be matched with, firms build labs that are smaller than the socially optimal size. Furthermore, because the level of investment depends on the researcher each firm expects to hire, if the equilibrium allocation of researchers to labs is inefficient there is an interesting interaction between the inefficiency arising in the matching stage and the inefficiency arising in the investment phase. I will also derive conditions for the existence of the competitive equilibrium with endogenous labs.

Initially, firms differ in their productivity p which is continuously distributed over $\mathcal{P} = [0, \bar{p}]$ with $p.d.f.$ $\gamma(p)$. In period $t = 0$, firms build labs at cost $c(p, L)$ continuous, positive, with continuous first and second derivative, strictly increasing and strictly convex in L with $\lim_{L\to\infty} \frac{\partial c(p,L)}{\partial L} = \infty\ \forall p$, strictly decreasing in p with $\lim_{p\to 0} \frac{\partial c(p,L)}{\partial L} = \infty\ \forall L$, and with $c(p, 0) = 0\, \forall p$.

Therefore, p represents a firm's comparative advantage in investing in labs. For example, although not explicitly considered here, some firms may be active in multiple scientific fields and labor markets for researchers. It is possible that those firms benefit from economies of scale when building labs, and therefore have higher p compared to firms that are active in only one scientific field.[15]

Following the investment in labs, events unfold as described in the previous section. In

[15] Whenever firms are identical, the results that are the same as the ones presented here, but the derivations are more convoluted. The reason is that, in this case, identical firms behave differently depending on what researcher they expect to be matched with. Hence, the equilibrium investment in labs is a correspondence and not a function.

period 1 each firm hires one researcher and produces new knowledge $k = af(L)$; in the last period surplus $\Phi_V(a, L) = \phi(k, V)$ is generated, which is then split between profits and wages. See Figure 1.

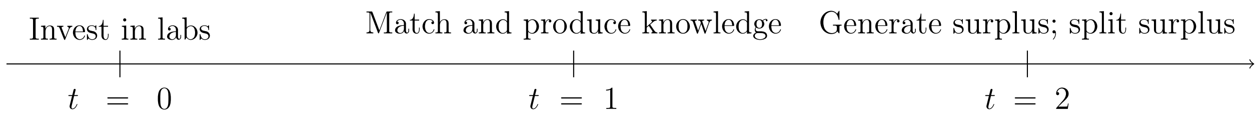

Fig. 1: Timeline

Welfare For a given $m(L)$ and given investment in labs by each firm $i(p)$ (which will be determined in equilibrium), total welfare now has to account for the cost of building labs and is therefore given by:

$$SW = \int_{\underline{p}}^{\overline{p}} \left[\phi(m(i(p))f(i(p)), V) - c(p, i(p))\right] \gamma(p) dP, \tag{4}$$

where $\underline{p}$ is the least productive firm building a lab of strictly positive size, also determined in equilibrium. Again, whereas firms and researchers take the total stock of knowledge V as given, from the social point of view V is determined endogenously and depends both on firms' investment decisions and on the equilibrium matching pattern.

3.1 Equilibrium with investment in labs

I introduce the following notation:

- $i(p) : P \rightrightarrows \mathbb{R}^+$, the investment made by firm p, determining the size of this firm's lab.
- $\tilde{m}(p) \equiv m(i(p)) : P \rightrightarrows A$, the matching rule on the equilibrium path (for labs built by some firms) mapping firms to researchers.
- $\tilde{\pi}(p) \equiv \pi(i(p)) : P \rightarrow \mathbb{R}^+$, firms' profits on the equilibrium path.
- $l(a) \equiv i(\tilde{m}^{-1}(a))$ the lab a researcher of ability a receives in equilibrium.

The definition of equilibrium that I use is similar to the one in Cole, Mailath, and Postlewaite (2001) with the difference that, here, the investment in labs generates a positive externality on the other firms.

Definition 3 (Equilibrim with investment in labs)**.** A given $\{V, i(.), m(.), \pi(.), w(.)\}$ constitutes an equilibrium if the following conditions are fulfilled:

1. The investment is optimal:

$$i(p) = \arg\max_{L \geq 0}\{\pi(L) - c(p, L)\}$$

2. Ex post, the matching is feasible and stable, where:

   - Feasibility: $\tilde{\pi}(p) + w(a) \leq \Phi_V(a, i(p))\ \forall p \in P$ and $a \in \tilde{m}(p)$.[16]
   - Stability: $\tilde{\pi}(p) + w(a') \geq \Phi_V(a', i(p))\ \forall p, p' \in P$ and $a' \in \tilde{m}(p')$.

3. For $L \notin \{L : L = i(p)$ for some $p \in P\}$ (investments off the equilibrium path):

$$\pi(L) = \max_a \{\Phi_V(a, L) - w(a)\}$$

4. $V = \nu \int af(l(a))z(a)da$ (remember that $z(a)$ is the p.d.f. of a).

To understand the definition, assume that there is an equilibrium, and consider a deviation made by a single firm. Because there is a continuum of firms, any change in the investment level of a single firm does not affect the distribution of labs, and has no impact on the equilibrium $w(a)$. In other words, firms are price takers in the market for researchers, both for on-equilibrium and off-equilibrium investment levels. Therefore, when choosing the investment level, each firm takes $w(a)$ as given. In addition, regardless of the size of the lab built, each firm can match with any researcher a provided that it pays the market wage $w(a)$.

[16] The general definition of feasibility is more complicated (see Cole et al., 2001). However, here this simpler version can be used because all distributions are continuous.

3.2 Solution

I can now solve for the equilibrium investment in labs and equilibrium stock of public knowledge.

Lemma 2. *In equilibrium, for $L > 0$:*

$$\pi'(L) = \frac{\partial \Phi_V(a, L)}{\partial L}|_{a=\tilde{m}(p)},$$

Proof. By applying the envelope theorem to point 1 and 3 of the definition of the equilibrium. □

The above lemma implies that whenever firms make positive investment, it must be that

$$\frac{\partial c(p, L)}{\partial L} = \frac{\partial \Phi_V(a, L)}{\partial L}|_{a=\tilde{m}(p)}. \tag{5}$$

Hence, firms' investments maximize surplus taking as given V and the researchers they will hire in equilibrium. Because the social planner would take into account the impact of the individual investment in L on the stock of public knowledge, Lemma 2 implies that, for given researcher's ability a, the investment is inefficient.

Furthermore, the lemma also implies that the matching pattern expected to emerge in the following period affects firms' investment decisions. If $\frac{\partial^2 \Phi_V(a,L)}{\partial L \partial a} < 0$, then the equilibrium matching pattern is NAM, and the investment in labs is distorted for two reasons: because the firm does not internalize the effect of its investment on V, and because the firm anticipates that the researcher hired is not the first best one. If the researcher hired is more productive than the first best researcher, then the two effects work in the same direction: the misallocation or researchers causes this firm to underinvest relative to the social optimum. If the researcher hired is less productive than the first best researcher, then the two effects work in opposite directions: the mismatch in the labor market causes this firm to overinvest in labs, while for given researcher this firm is underinvesting in labs.

To show the existence of the general equilibrium, I restrict my attention to the two cases considered in Proposition 1, so that the matching pattern is either NAM or PAM, independently from the distribution of labs and V. I also will use the following fact:

Lemma 3. *$i(p)$ is a continuous function.*

Call $\underline{p}(V)$ the least productive firm building a lab of positive size as a function of V. This firm earns zero profits in equilibrium, and therefore

$$\underline{p}(V) \equiv \min \{\overline{p}, p : \ \tilde{\pi}(p) = c(p, i(p))\}$$

Note that the higher is V, the larger the surplus to be shared and the payoff of each firm. It follows that $\underline{p}(V)$ is decreasing in V. However, because the marginal cost of investing becomes arbitrarily large for arbitrarily small p, we have that $\underline{p}(V) > 0 \ \forall V$. Finally, $\underline{p}(V)$ is continuous because $\pi(L)$, $m(L)$ and $i(p)$ are continuous.

Similarly, call $V(\underline{p})$ the total knowledge produced as a function of the least productive firm building a lab of positive size, that is

$$V(\underline{p}) \equiv V : \ V = \int_{\underline{p}}^{\overline{p}} \nu \cdot \tilde{m}(p) f(i(p)) \gamma(p) dp, \tag{6}$$

By equation (5), $i(p)$ depends on V, which implies that (6) could have multiple or no solutions. I therefore need to impose some additional restrictions:

Assumption 1. *L solution to* (5)

- *is strictly greater than zero at $V = 0$ for some $p > 0$,*
- *is either decreasing in V or strictly concave in V.*

Note that these restrictions relate to the optimal investment made by firms *conditional on investing.* That is, it is possible that there is a corner solution and the profit maximizing level of investment at $V = 0$ is zero, and at the same time the first part of the assumption is satisfied because the investment level that satisfies equation (5) is positive. By the implicit function theorem, L solution to (5) has slope

$$\frac{-\frac{\partial^2 \Phi_V(a,L)}{\partial L \partial V}}{\frac{\partial^2 \Phi_V(a,L)}{\partial L^2} a f'(L)}.$$

Because, by assumption $\frac{\partial^2 \Phi_V(a,L)}{\partial L^2} < 0$, whenever L and V are substitutes in $\Phi_V(a, L)$ then the above expression is negative and the second part of the Assumption 1 is satisfied. If instead L and V are strict complements in $\Phi_V(a, L)$, Assumption 1 is satisfied only if the above expression is decreasing in V.

Lemma 4. *If Assumption 1 holds, then $V(\underline{p})$ exists, is unique and is continuous in $\underline{p}$.*

Under the assumption 1, $V(\underline{p})$ reaches a maximum level at $\underline{p} = 0$, decreasing in $\underline{p}$, and is zero at $\underline{p} = \overline{p}$. Note also that $V(\underline{p})$ shifts upward whenever ν increases. The curves $V(\overline{p})$ and $\overline{p}(V)$ are plotted in Figures 2 and 3, which differ for the fact that ν (and therefore $V(\overline{p})$) is higher in Figure 3 than in Figure 2. The shapes and continuity of $V(\underline{p})$ and $\underline{p}(V)$ imply that they must cross at least one, leading to the following Proposition.

Proposition 2. *Assume that we are in one of the two cases considered in Proposition 1. Under Assumption 1 the competitive equilibrium of the economy always exists. If the value of research ν is high enough, equilibria where a positive amount of knowledge is produced exist.*

Proof. In the text. □

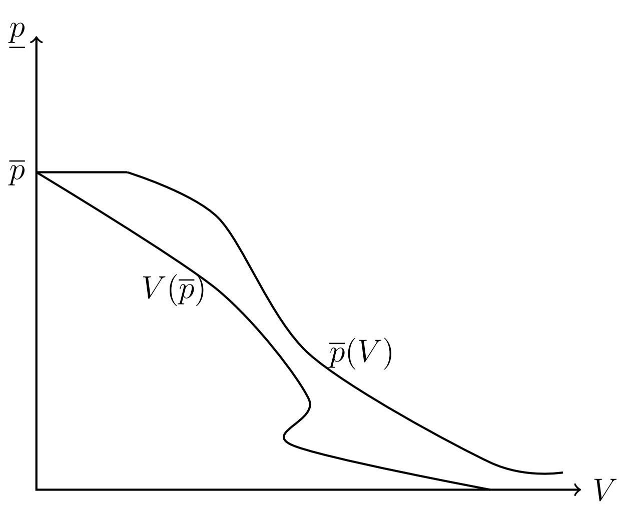

Fig. 2: Low ν, only equilibrium at $V = 0$.

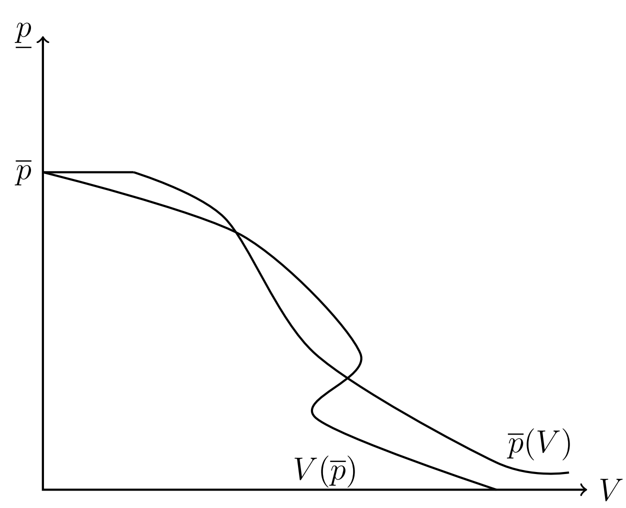

Fig. 3: High ν, equilibria with $V > 0$.

Finally, note that the combination of a Pigouvian tax and Lidahl subsidies discussed in Section 2.3 aligns the social benefit of building labs with its private benefit, and therefore achieves the first best also when the distribution of labs is endogenous. The social planner can achieve the first best also by imposing the efficient allocation of researchers to labs, and imposing an appropriate subsidy to the investment in labs. Only imposing a subsidy in labs, however, cannot restore efficiency, as the discussion in Section 2.3 shows.

4 Extension: Reputation Concerns.

In the benchmark model, I implicitly assumed that the creation of new knowledge generates only monetary benefits, which can be freely split between firms and researchers. However, since the work of Merton (1957), it is well known that researchers care about reputation. Merton calls scientists' concern for reputation the *race for priority*: researchers want to be recognized as the first to discover something. The role of reputation in the production of knowledge was explored in the economic literature by Dasgupta and David (1985). They argue that, on the one hand, reputation motivates researchers, which is important because an incentive scheme based exclusively on the quality of scientific output would be hard to implement. On the other hand, reputation fosters openness, which guarantees the circulation of ideas and generates a faster pace of scientific progress.

In this section, I introduce reputation into the model as an additional output of R&D, and I show that reputation may affect the production of knowledge in a new way: by changing the equilibrium allocation of researchers to firms. The reason is that reputation accrues to the researcher but cannot be transferred back to the firm because of liquidity constraints, and therefore transforms the matching problem in a Non-Transferrable Utility (NTU) problem (see Legros and Newman, 2007).[17] I show that NTU has the usual effect on the equilibrium matching: if it is strong enough it will generate a PAM allocation. More interestingly, here NTU may actually improve welfare because PAM is the knowledge-maximizing allocation of researchers to labs.

Formally, I assume that the researchers' utility is:

$$U(a) = w(a) + \rho\Phi_V(a, L)$$

where $w(a) \geq 0$ is the monetary payment received and $\rho \in (0, 1)$ is the fraction of surplus

[17] NTU can arise in many other cases. For example, if researchers' effort impact the value of the science produced and firms cannot implement incentive schemes depending on the value of output, the particular surplus split agreed ex ante determines total surplus.

that accrue to the researcher as *reputation*. By assumption $\rho\Phi_V(a,L)$ cannot be shared with the firm. The remaining fraction of surplus $(1-\rho)\Phi_V(a,L)$ accrues to the firm, and can be shared with the researcher.[18]

Lemma 5. *If ρ is sufficiently large, the equilibrium matching is PAM.*

Intuitively, the "reputation component" of the researcher's utility is increasing in the knowledge produced. Hence researchers are willing to give up their cash payments to produce more knowledge, the more so the larger is ρ.[19] When ρ is large enough cash payments are zero, and the payoff for both sides (firms and researchers) increases with the other side's type, and the equilibrium matching is PAM.

For given distribution of labs, the fact that reputation concerns change the equilibrium matching from NAM to PAM has two effects on welfare. For given V firms' private surplus will decrease, but the stock of public knowledge V will increase. The net effect will depend on the parameter ν (measuring the value of a unit of new knowledge), and is positive as long as ν is sufficiently large.

Finally, reputation has an effect on the investment made by firms. As ρ increases, both the benefit and marginal benefit from investing in labs may decrease, because the fraction of surplus that is in monetary form and can be transferred to firms decreases. However, when a firm's investment is sufficiently inelastic, the overall welfare effect of reputation concerns is positive. For example, setting $c(p,L) = \left(\frac{L}{p}\right)^{\Delta}$ for Δ large enough guarantees that firms will invest $L \approx p$ regardless of the value of ρ. To summarize: if firms investment is sufficiently inelastic and ν is sufficiently large the introduction of reputation concerns increases total welfare.

[18] In an alternative specification the researcher's utility could depend directly on k. Also in this case, the fact that part of the surplus is not sharable will push the equilibrium toward PAM. In addition to that, if the researcher cares directly about k, then there is a supermodular element in the researcher's payoff function, which again pushes the equilibrium toward PAM. Hence, both specifications lead to the same equilibrium allocation. However, the specification chosen here allows for a more immediate comparison of the welfare properties of the equilibrium with and without reputation concerns, because reputation does not change the total surplus generated in a match, but only how this surplus is allocated.

[19] This finding is consistent with Stern (2004). In his paper "Do Scientists Pay to be Scientists?" the author collects data on job offers received by a sample of biology Ph.D. job market candidates. He finds that firms engaged in scientific research offer wages 25% lower than firms that are not engaged in scientific research.

5 Conclusions

In this paper, I build a model of knowledge production in which firms invest in labs and hire researchers on a competitive market, and I analyze the competitive allocation of researchers. I show that whenever the marginal private benefit of firms' R&D activities decreases sufficiently fast, scientific talent may be misallocated in equilibrium. In case of misallocation, subsidies to the investment in labs cannot restore full efficiency as they do not affect the equilibrium matching pattern. There is scope for policies that distort the market allocation of researchers, because reallocating researchers from one firm to another firm may generate an aggregate welfare gain (in addition to a local gain and a local loss).

I also show that the presence of reputation concerns may change the equilibrium matching pattern to PAM: good researchers work in big labs because they benefit directly from their research output. It follows that reputation concerns may increase the aggregated knowledge produced and total welfare. It also follows that the empirical matching patterns between researchers and labs should depend on the strength of researchers' reputation concerns. For example, assuming that reputation concerns are stronger for young researchers, we should observe productive young researchers working for firms with large labs, but productive old researchers working for firms with small labs.

Finally, I characterize the optimal allocation of researchers to labs only partially. I show that whenever the allocation of researchers to firms is inefficient, it is possible to increase total welfare by re-allocating productive researchers from small labs to larger labs. However, the first-best allocation of researchers to firms depends on the exact distribution of labs and researchers' ability, and may involve PAM over some range and NAM over some other range. Deriving the first-best re-allocation policy and describing which researcher should move to what lab is left for future work.

Appendix: mathematical derivations.

Proof of Lemma 1

Proof. By standard matching theory, local substitutability implies local NAM, and local complementarity implies local PAM. Substitutability or complementarity at a given $\hat{a} =$

$m(\hat{L}), \hat{L}$ depends on the sign of:

$$\frac{\partial\phi(af(L),V)}{\partial a\partial L}|_{a=m(\hat{L}),L=\hat{L}}$$

that itself depends on the sign of

$$\frac{\partial\chi(\log(af(L)),V)}{\partial a\partial L}|_{a=m(\hat{L}),L=\hat{L}}$$

Using the chain rule, it is possible to show that the above expression is negative if

$$\frac{\partial^2\chi(x,V)}{\partial x^2}|_{x=\log(m(\hat{L})\hat{L})} < 0$$

(i.e. $\chi(,)$ is concave in its first argument) and positive otherwise. □

Proof of Proposition 1

Proof. The fact the the equilibrium is global PAM/NAM depending on whether $\chi(,)$ is globally convex/concave in its first argument follows immediately from Lemma 1.

To show that global NAM is inefficient if ν is large, I prove the following stronger statement: that *local* NAM is inefficient for ν sufficiently large. Consider an equilibrium matching function $m(L)$ and assume that $m'(\hat{L}) < 0$ at some $\hat{L}$. Consider an alternative matching function $m_\epsilon(L)$ with $m_\epsilon(L) = m(L)$ for $L \notin [\hat{L}-\epsilon, \hat{L}+\epsilon]$, and $m'_\epsilon(L) > 0$ for all $L \in [\hat{L}-\epsilon, \hat{L}+\epsilon]$ for some $\epsilon > 0$ small. In other words, the alternative matching function is a perturbation of the the equilibrium matching function, and implements PAM instead of NAM in the interval $[\hat{L}-\epsilon, \hat{L}+\epsilon]$. Call V the stock of aggregate knowledge produced under the equilibrium matching function $m(L)$, and V_ϵ stock of aggregate knowledge produced under the matching function $m_\epsilon(L)$.

When the matching function is perturbed in the way described above, the change in the surplus generated by a firm with lab L is[20]

$$\frac{d\phi(m(L)L,\nu\int m(L)Lh(L)dL)}{d\epsilon} = \underbrace{\frac{\partial\phi(m(L)L,\nu\int m(L)Lh(L)dL)}{\partial k}}_{\geq 0} \underbrace{\frac{\partial[m(L)L]}{\partial\epsilon}}_{<0 \text{ for some firms}} + \underbrace{\frac{\partial\phi(m(L)L,\nu\int m(L)Lh(L)dL)}{\partial V}\cdot\frac{\partial[\int m(L)Lh(L)dL]}{\partial\epsilon}\cdot\nu}_{>0}$$

Note that for some firms $\frac{\partial[m(L)L]}{\partial\epsilon} < 0$: under the alternative matching function some firms end up matched with a worse researcher compared to the equilibrium. On the other hand, $\frac{\partial[\int m(L)Lh(L)dL]}{\partial\epsilon} > 0$ for all $\epsilon > 0$: the total stock of public knowledge is greater when local NAM is transformed into local PAM. The parameter ν measures the value of the new knowledge produced. When it is sufficiently large, the effect of a perturbation in the matching function on the surplus generated by a given firm is positive, even for firms that are matched with a worse researcher.

Hence, for ν is sufficiently large, implementing the alternative matching function increases the surplus generates by each firm in the economy. It follows that total surplus is greater under the alternative matching function rather than the equilibrium matching function.

To conclude the proof, I need to show that if PAM is the equilibrium, then the equilibrium is efficient. This follows simply from the fact that the equilibrium maximizes surplus for given V. At the same time PAM maximizes V. Because surplus is increasing in V, PAM also maximizes surplus.

$\square$

Proof of Lemma 3

Proof. By point 1 of the definition of equilibrium, $i(p) > 0$ solves $\pi'(L) = \frac{\partial c(p,L)}{\partial L}$. By Lemma 2, $\pi''(L) = \frac{\partial^2\Phi_V(a,L)}{\partial L^2} < 0$, while by definition $\frac{\partial^2 c(p,L)}{\partial L^2} > 0$ with $\lim_{L\to\infty}\frac{\partial c(p,L)}{\partial L} = \infty\ \forall p$. Hence, there is always a unique L that satisfies $\pi'(L) = \frac{\partial c(p,L)}{\partial L}$. Continuity follows from the fact that $\pi(L)$ and $c(p,L)$ are continuous with continuous first derivatives, and therefore firm's optimal investment is also continuous. $\square$

[20] Remember that, by definition $k = m(L)L$ and $V = \nu\int m(L)Lh(L)dL$.

Proof of Lemma 4

Proof. Here I take as given that firms in $[\underline{p}, \overline{p}]$ build labs of strictly positive size. Because their investment satisfy (5), the two parts of Assumption 1 imply, respectively, that

- $\int_{\underline{p}}^{\overline{p}} \nu \cdot \tilde{m}(p) f(i(p))\gamma(p) dp$ is positive at $V = 0$ for all $\underline{p} > 0$,
- $\int_{\underline{p}}^{\overline{p}} \nu \cdot \tilde{m}(p) f(i(p))\gamma(p) dp$ is either decreasing or strictly concave in V.

Which imply that $V = \int_{\underline{p}}^{\overline{p}} \nu \cdot \tilde{m}(p) f(i(p))\gamma(p) dp$ has a unique solution, which is continuous because $i(.)$, $m(.)$, $\gamma(.)$, $f(.)$ are continuous. □

Proof of Lemma 5

Proof. Consider a given V and a given distribution of labs. Suppose that, in equilibrium, the equilibrium is PAM and researchers receive no cash payments: their full payoff is in form of reputation. In this equilibrium, the only possible deviation for a research is to move to a firm with a smaller lab and capture a positive monetary payoff. Consider the best researcher $\overline{a}$, that is matched with the firm with the largest lab $\overline{L}$. This researcher could work for the firm with the smallest lab and receive a cash payment equal to the entire surplus generated within this match. This deviation is not profitable as long as:

$$\rho \Phi(\overline{a}, \overline{L}) > \Phi(\overline{a}, \underline{L})$$

If this deviation is not profitable for the most productive researcher it is not profitable for any researcher. Furthermore, for any V and any distribution of labs, we always have $\Phi(\overline{a}, \overline{L}) > \Phi(\overline{a}, \underline{L})$. Hence, if ρ is sufficiently close to 1, the equilibrium is PAM for any V and any distribution of labs. □

References

Acemoglu, D., P. Aghion, and F. Zilibotti (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association 4*(1), 37–74.

Acemoglu, D., U. Akcigit, and M. A. Celik (2014). Young, restless and creative: Openness to disruption and creative innovations. Technical report, National Bureau of Economic Research.

Acs, Z. and D. Audretsch (1987). Innovation in large and small firms. *Economics Letters 23*(1), 109–112.

Arora, A., P. David, and A. Gambardella (1998). Reputation and competence in publicly funded science: estimating the effects on research group productivity. *Annales d'Economie et de Statistique*, 163–198.

Arora, A. and A. Gambardella (2005). The impact of NSF support for basic research in economics. *Annales d'Economie et de Statistique*, 91–117.

Baumol, W. J. (1996). Entrepreneurship: Productive, unproductive, and destructive. *Journal of Business Venturing 11*(1), 3–22.

Bikard, M., F. Murray, and J. S. Gans (2015). Exploring trade-offs in the organization of scientific work: Collaboration and scientific reward. *Management Science*.

Cockburn, I. and R. Henderson (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *Journal of Industrial Economics 46*(2), 157–182.

Cohen, W. and S. Klepper (1996). A reprise of size and R&D. *The Economic Journal 106*(437), 925–951.

Cohen, W. M. and D. A. Levinthal (1989). Innovation and learning: The two faces of R&D. *Economic Journal 99*(397), 569–96.

Cole, H., G. Mailath, and A. Postlewaite (2001). Efficient non-contractible investments in large economies. *Journal of Economic Theory 101*(2), 333–373.

Dasgupta, P. and P. David (1985). Information disclosure and the economics of science and technology. *CEPR Discussion Papers*.

Elfenbein, D., B. Hamilton, and T. Zenger (2010). The small firm effect and the entrepreneurial spawning of scientists and engineers. *Management Science 56*(4), 659–681.

Estevan, F., T. Gall, P. Legros, and A. F. Newman (2017). College admission and high school integration. *Working Paper*.

Gambardella, A. (1992). Competitive advantages from in-house scientific research: the us pharmaceutical industry in the 1980s. *Research Policy 21*(5), 391–407.

Griffith, R., S. Redding, and J. Van Reenen (2003). R&D and absorptive capacity: Theory and empirical evidence. *The Scandinavian Journal of Economics 105*(1), 99–118.

Griffith, R., S. Redding, and J. Van Reenen (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. *Review of Economics and Statistics 86*(4), 883–895.

Hammerschmidt, A. (2009, September). No pain, no gain: An R&D model with endogenous absorptive capacity. *Journal of Institutional and Theoretical Economics (JITE) 165*(3), 418–437.

Hammond, P. (1995). Four characterizations of constrained pareto efficiency in continuum economies with widespread externalities. *Japanese Economic Review 46*(2), 103–124.

Hammond, P., M. Kaneko, and M. Wooders (1989). Continuum economies with finite coalitions: Core, equilibria, and widespread externalities. *Journal of Economic Theory 49*, 113–134.

Jovanovic, B. and R. Rob (1989). The growth and diffusion of knowledge. *The Review of Economic Studies 56*(4), 569–582.

Kamecke, U. (1992). On the uniqueness of the solution to a large linear assignement problem. *Journal of Mathematical Economics 21*, 509–21.

Kamien, M. and I. Zang (2000). Meet me halfway: research joint ventures and absorptive capacity. *International Journal of Industrial Organization 18*(7), 995–1012.

Kaneko, M. and M. Wooders (1996). The nonemptiness of the f-core of a game without side payments. *International Journal of Game Theory 25*(2), 245–258.

Leahy, D. and J. Neary (2007). Absorptive capacity, R&D spillovers, and public policy. *International Journal of Industrial Organization 25*(5), 1089–1108.

Legros, P. and A. F. Newman (2007). Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities. *Econometrica 75*(4), 1073–1102.

Lucas Jr, R. E. and B. Moll (2014). Knowledge growth and the allocation of time. *Journal of Political Economy 122*(1).

Merton, R. (1957). Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review 22*(6), 635–659.

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.

Rosenberg, N. (1990). Why do firms do basic research (with their own money)? *Research Policy 19*(2), 165–174.

Scherer, F. (1965). Firm size, market structure, opportunity, and the output of patented inventions. *The American Economic Review 55*(5), 1097–1125.

Stern, S. (2004). Do scientists pay to be scientists? *Management Science 50*(6), 835–853.

Tilton, J. (1971). International diffusion of technology: The case of semiconductors. *The Brookings Institution, Washington, DC*.

Vandenbussche, J., P. Aghion, and C. Meghir (2006). Growth, distance to frontier and composition of human capital. *Journal of economic growth 11*(2), 97–127.