

The Dynamics of Knowledge Diversity and Economic Growth*

Marcus Berliant** and Masahisa Fujita[±]

February 27, 2010

Abstract

How is long run economic growth related to the endogenous diversity of knowledge? We formulate and study a microeconomic model of knowledge creation, through the interactions among a group of heterogeneous R & D workers, embedded in a growth model to address this question. In contrast with the traditional literature, in our model the composition of the research work force in terms of knowledge heterogeneity matters, in addition to its size, in determining the production of new knowledge. Moreover, the heterogeneity of the work force is endogenous. Income to these workers accrues as patent income, whereas transmission of newly created knowledge to all such workers occurs due to public transmission of patent information. Knowledge in common is required for communication, but differential knowledge is useful to bring originality to the endeavor. Whether or not the system reaches the most productive state depends on the strength of the public knowledge transmission technology. Equilibrium paths are found analytically. Long run economic growth is positively related to both the effectiveness of pairwise R & D worker interaction and to the effectiveness of public knowledge transmission. JEL Classification Numbers: D83, O31, D90 Keywords: knowledge creation, knowledge externalities, microfoundations of endogenous growth, knowledge diversity and growth

*We thank David Levine, Kyoji Fukao and an anonymous referee for helpful comments, and Y. Nakamura for creating the figures. The first author is grateful for funding from the Kyoto Institute of Economic Research at Kyoto University, from Washington University in St. Louis, and from the American Philosophical Society. The second author is grateful for Grants Aid for Scientific Research Grant A 18203016 and 21243021 from the Japanese Ministry of Education and Science and for funding from the Weidenbaum Center at Washington University.

**Department of Economics, Washington University, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130-4899 Phone: (1-314) 935-8486, Fax: (1-314) 935-4156, e-mail: berliant@artsci.wustl.edu

[±]Konan University, 8-9-1 Okamoto, Higashinada-ku, Kobe, 658-8501 Japan. Phone and Fax: (81-78) 435-2409, e-mail: fujitam@center.konan-u.ac.jp

1 Introduction

How is economic growth related to the diversity of knowledge? How does knowledge diversity change as an economy grows? Does the knowledge creation activity lead naturally to the formation of teams of creators?¹ Can more effective public knowledge transmission, via the patenting process or the internet, cause the knowledge base to become too homogeneous and slow growth? Given spillovers in the creation of new knowledge, is the equilibrium knowledge production path efficient?

To address these questions, we attempt to provide microfoundations for macro models of endogenous growth, such as Shell (1966), Romer (1986, 1990), Lucas (1988), Jones and Manuelli (1990), Grossman and Helpman (1991) and many papers building on these contributions. In particular, the model proposed below is closely related to the endogenous growth model developed by Romer (1990) in which R & D firms invest resources to develop blueprints for new products. In Romer's model, the productivity of each R & D firm rises in proportion to the stock of general knowledge capital; the latter is assumed to be equal to the cumulative number of products invented in the R & D sector in the past. In addition, all workers in the R & D sector are assumed to be homogeneous. Hence, in Romer's model, when labor is the unique input in the R & D sector, the number of new products developed per unit of time is also proportional to the number of R & D workers at that time.

We maintain the assumption of monopolistic competition in the sector that produces horizontally differentiated consumption goods. A manufacturing sector produces consumption goods for both their workers and the knowledge workers, using a Dixit and Stiglitz (1977) monopolistic competition framework. To produce a consumption commodity under constant returns to manufacturing labor input, a patent must be purchased from the R & D sector.

The main contribution of this paper is to examine the consequences of the introduction of microstructure of the R & D process into a growth model. The microstructure we use is an extension of Berliant and Fujita (2008).² In our model, diversity of researchers in terms of knowledge composition enters into the production of new knowledge. In contrast with Romer's model, where researchers are homogeneous and only their number matters, researchers in

¹For empirical evidence on teams and patents, see for example Guzzo and Dickson (1996), Chatman et al (1998), Trajtenberg et al (2006), and Wuchty et al (2007).

²Other approaches to modelling knowledge creation include Weitzman (1998), Olsson (2000, 2005), Jones (2005) and Ghiglino (2005).

our model are horizontally differentiated by their knowledge, where individual knowledge composition is endogenous and evolves over time. The highest productivity of a partnership between two researchers occurs when knowledge in common, necessary for communication, and exclusive knowledge, bringing originality into the partnership, are in balance.

In contrast with our earlier work, in this paper we introduce public transfer of knowledge through patent information revelation. To be specific, in our model we assume that a certain proportion of new ideas result in patents. In reality, when a patent is granted, generally great detail about the invention is made public. In the US, a wealth of information on patents granted since 1790 is available on the web at <http://patft.uspto.gov/> This information includes a precise description of the invention as well as patents cited by each patent granted, providing evidence of how public knowledge influences the creation of new ideas. Another type of public knowledge, more familiar in academia, are publications in professional journals and their cites. Quantitatively, Walsh and Nagaoka (2009, Figures 18A and 18B) find in surveys of inventors that the scientific literature and patents are the two most important sources of ideas for new inventions in both the US and Japan.

Although this extension makes the analysis more complex, the phenomenon of public knowledge transmission is essential when discussing the modern economy. Again in contrast with our earlier work, that was a stand alone model of R & D, here we embed the model of R & D in a growth model, where public knowledge transmission is essential. The role of the R & D sector in the production of new physical commodities is important in our model. However, for tractability reasons, feedback from the physical goods production sector to the R & D sector is absent, and left to future work.

Manufacturing workers, firms, and consumers in the R & D sector are all farsighted, in the sense that they have rational expectations about prices. The knowledge workers themselves are myopic in their choices concerning R & D partnerships. This simplifying assumption reduces the computational burden for both the knowledge workers and us. Also, for simplicity, we deal exclusively with the case when the agents are symmetric. Our model is analytically tractable, so we do not have to resort to simulations; we find each equilibrium path explicitly.

Our analysis demonstrates that when the speed of public knowledge transmission is not too high, from the largest set of initial conditions the knowledge production system eventually reaches the most productive state. To get there,

the R & D workers are partnering with every other worker for very short intervals in order to avoid building up too much knowledge in common. When the system reaches the most productive state, the research pattern changes dramatically, as smaller research groups form so that partnerships only occur between members of the same group, since this enables workers to maintain the best possible mix of knowledge diversity and knowledge in common. Interactions between groups are weak, as only public knowledge transmission takes place between groups. The size of these research groups is proportional to the importance of the heterogeneity of knowledge in the knowledge production process (e.g. in the biotech industry) and to the transmission strength of the public knowledge spillover (e.g. in the information technology industry). The R & D workers are myopic, whereas there are externalities in partnerships between knowledge workers, so one would not expect efficient outcomes. Here we have added a public knowledge externality; we expect this to be an additional source of market failure. However, when these research groups form to maintain the most productive state, the externalities are internalized endogenously.

The long run economic growth rate is positively related to both the effectiveness of pairwise knowledge worker interaction and, more importantly, to the effectiveness of public knowledge transmission. The latter is not obvious in the context of a model of knowledge diversity, since more effective public knowledge transmission can result in a more homogeneous work force in the R & D sector. Thus, there are both costs and benefits associated with a more effective public knowledge transmission technology. However, *the endogenous adjustment of group size to a better public knowledge transmission technology* implies that better public knowledge transmission improves long run growth. Finally, if we define efficiency constrained by the monopolistic competition environment for consumption goods, for a sufficiently large number of knowledge workers our equilibrium paths are nearly constrained efficient.

In applying our results to real life issues, we must be very careful about interpreting the meaning of the comparative dynamics³ that we derive. The optimal group size increases as the speed of public knowledge transmission increases or the weight given to knowledge in common decreases. In real life, however, once the optimal group size is reached under a fixed set of parameters, group size does not easily adjust to a new optimal size under a new set of

³The term “comparative dynamics” refers to differences in the the equilibrium time paths of endogenous variables under different values of exogenous parameters.

parameters. This is because of the *lock-in effect* of the optimal group size. In particular, the knowledge of a R & D worker in one group will drift apart from the knowledge of R & D workers in other groups. Thus, once the most productive state is achieved under one set of parameters, realigning the R & D workers into larger groups when parameters change will not result in optimal knowledge production, since the R & D workers initially in different groups have differentiated themselves too much from each other. This lock-in effect inherent in an R & D system may partly explain, for example, why the Japanese economy has been suffering from a prolonged recession and slow growth since the early 1990s.

Specifically, the so called IT revolution has significantly increased the speed of public knowledge transmission, whereas new industries displaying rapid growth (e.g. the biotech industry), such as computer software and advanced service industries (including global finance), tend to have a higher weight on knowledge diversity in R & D than traditional manufacturing industries (based mainly on incremental improvements in Japan). Due to the lock-in effect, R & D group size and composition were inherited from past economic circumstances. Our model implies low mobility of Japanese workers and researchers beyond existing institutions, through no fault of their own.⁴ But the Japanese R & D system has not adapted adequately to the new situation. Our analysis implies that research groups in the new industries should be made more diverse and larger. Such a change would generate short term reductions in R & D productivity in exchange for long term gains.

It is important to relate our work to the macro literature on the presence or absence of scale effects in models and the real world. In the endogenous growth models, creation of new ideas is the engine of growth. As discussed in Jones (1999), the early models featured a public goods aspect of idea creation, leading to growth rates that are positively related to population. Unfortunately, this did not match the data, that appears to yield growth rates that are independent of population, or "scale free." Several models implying scale free growth have been proposed, for example Young (1998), Peretto and Smulders (2002), and Dinopoulos and Thompson (1998). We provide an alternative model without scale effects in the growth rate. There are two reasons behind why our model yields a scale free growth rate. First, the long run equilibrium group size for R & D is independent of population. That is, in the long run, direct interaction among R & D workers is only within groups but not between groups, where

⁴Please refer to the conclusion for further discussion of this issue.

the size of these groups is determined by exogenous variables not including population. Second, the learning capacity of R & D workers is limited in proportion to their time and intelligence, so public learning is not a function of population size.

Among the work featuring scale-free growth, the interesting paper of Peretto and Smulders (2002) is the most closely related to our work. Like our work, they consider a model where spillovers within networks/teams of researchers play an important role in technological progress. However, in our model, the additional microstructure for the R & D sector based on the pairwise interaction of R & D workers uncovers the lock-in effect, as we have already detailed. Furthermore, we address efficiency issues here; they are important for policy implications in the presence of externalities.⁵

In sum, we believe that endogenous diversity in individual knowledge composition is essential when developing microfoundations for economic growth models of the modern economy.

Section 2 gives the model and notation, Section 3 analyzes the equilibrium path of dynamics in the knowledge production sector, Section 4 analyzes the equilibrium growth path for the entire economy, whereas Section 5 explores the efficiency properties of the equilibrium path. Section 6 gives our conclusions and suggestions for future knowledge workers. Two appendices provide the proofs of key results.

2 The Model

In this section, we introduce the basic model. There are three types of activity in the economy. There are consumers of physical goods, producers of physical goods, and the R & D sector. The activities in the economy representing physical commodity production and consumption are standard models of product variety with monopolistic competition. The major difference between our model and others is the level of detail in the R & D sector, that generates patents sold to the producers of physical, differentiated products. We shall describe first the consumer side of the economy, namely a market for differentiated products. In the following subsection, we describe the production side of this market. Finally, we describe the R & D sector, the focus of

⁵Moreover, in Peretto and Smulders (2002), the work pattern used by firms for knowledge creation is taken to be exogenous, whereas there is no public knowledge transmission mechanism. Thus, it is not possible to ask questions about either in their framework.

our work.

To begin, there are two types of workers: knowledge workers (\mathbb{K} -workers) engaged in R & D, and manufacturing workers (\mathbb{M} -workers) producing differentiated products. For simplicity, we assume that the type of each worker is exogenously given, so workers cannot change sectors. Let N denote the number of \mathbb{K} -workers, and let L denote the number of \mathbb{M} -workers.

Before getting into the details of the model, it is useful to discuss the rationality assumptions we make regarding the agents. For the producers and the manufacturing workers, we assume that they all have perfect foresight, including knowledge of future prices. When knowledge workers consume, they also have perfect foresight.

The important assumption concerns knowledge workers when they make decisions about knowledge production, in particular which partner to work with at any given time or whether to work alone. In our previous work, we have used a myopic core solution concept. That is, workers in the R & D sector make decisions about their research teams in a cooperative manner but without looking ahead at the long term consequences. Such a concept will be used below. The resulting time-varying pattern of knowledge creation will be found explicitly; the path is both undominated and weakly overtakes any other path. But we also show in Section 5 (under some restrictions) that the myopic core solution is nearly optimal⁶ if a utilitarian social welfare function is used under perfect foresight. Thus, the solution path we propose for a large set of parameters and initial conditions is at the intersection of many solution concepts, and is nearly efficient for the R & D sector (nearly constrained efficient in the entire economy).

In our view, this result is strongest when considering the myopic core solution concept, since in that case we have postulated a model with externalities and R & D workers who are myopic, but attain a nearly constrained efficient outcome in spite of this.

2.1 Consumers

First, we describe consumers' preferences (the time argument is suppressed when no confusion arises). All workers have the same instantaneous utility function given by

⁶Namely, for every $\epsilon > 0$, if N is sufficiently large, then the equilibrium path is within ϵ of the maximal welfare achievable by any feasible path of knowledge creation.

$$u = \left[\int_0^M q(h)^\rho dh \right]^{1/\rho} \quad 0 < \rho < 1 \quad (1)$$

In this expression, M is the total mass of varieties available in the economy at a given time, whereas $q(h)$ represents the consumption of variety $h \in [0, M]$.

If E denotes the expenditure of a consumer at a given time while $p(h)$ is the price of variety h , then the demand function is as follows:

$$q(h) = Ep(h)^{-\sigma} P^{\sigma-1} \quad h \in [0, M] \quad (2)$$

where $\sigma \equiv 1/(1 - \rho)$ represents the elasticity of substitution between any two varieties, and P is the price index of varieties given by

$$P \equiv \left[\int_0^M p(h)^{-(\sigma-1)} dh \right]^{-1/(\sigma-1)} \quad (3)$$

Introducing (3) and (2) into (1) yields the indirect utility function

$$v = E/P$$

We now describe the behavior of an arbitrary consumer i , who is either a \mathbb{K} -worker or an \mathbb{M} -worker. If this consumer chooses an expenditure path, $E_i(t)$ for $t \in [0, \infty)$ such that $E_i(t) \geq 0$, then his indirect utility at time t is given by

$$v_i(t) = E_i(t)/P(t) \quad (4)$$

where $P(t)$ is the price index of the manufactured goods at time t .

The lifetime utility of consumer i at time 0 is then defined by

$$U_i(0) \equiv \int_0^\infty e^{-\gamma t} \ln[v_i(t)] dt \quad (5)$$

where $\gamma > 0$ is the subjective discount rate common to all consumers.

The intertemporal allocation of resources is governed by an interest rate equal to $v(t)$ at time t . We must now specify consumer i 's intertemporal budget constraint, that is, the present value of expenditure equals wealth. Let $y_i(t)$ be the income that this consumer receives at time t . For any \mathbb{M} -worker, their income at time t will be their wage at that time, whereas for any \mathbb{K} -worker, their income at time t will be the value of the patents they create at that time. Then, the present value of income is given by

$$W_i(0) = \int_0^\infty e^{-\bar{v}(t)t} y_i(t) dt \quad (6)$$

where $\bar{\nu}(t) \equiv (1/t) \int_0^t \nu(\tau) d\tau$ is the average interest rate between 0 and t ; in (6), the term $\exp[-\bar{\nu}(t)]$ converts one unit of income at time t to an equivalent unit at time 0. Using the budget flow constraint, Barro and Sala-i-Martin (1995, p. 66) show that the consumer's intertemporal budget constraint may be written as follows:

$$\int_0^\infty E_i(t) e^{-\bar{\nu}(t)t} dt = \omega_i + W_i(0) \quad (7)$$

where ω_i is the value of the consumer's *initial assets*, specified as follows:

$$\omega_i = 0 \text{ for } \mathbb{M}\text{-worker } i$$

and

$$\omega_i = \frac{\Pi(0) \cdot M(0)}{N} \text{ for } \mathbb{K}\text{-worker } i \quad (8)$$

So each \mathbb{K} -worker owns the same number $\frac{M(0)}{N}$ of patents at time 0, where the price of patents at time 0 is $\Pi(0)$.

Then, if $E_i(\cdot)$ stands for an expenditure path that maximizes (5) subject to (7), the first order condition implies that

$$\dot{E}_i(t)/E_i(t) = \nu(t) - \gamma \quad t \geq 0 \quad (9)$$

where $\dot{E}_i(t) \equiv dE_i(t)/dt$. Since (9) must hold for every consumer, it is clear that the following relation must hold

$$\dot{E}(t)/E(t) = \nu(t) - \gamma \quad t \geq 0 \quad (10)$$

where $E(t)$ stands for the total expenditure in the economy at time t .

2.2 Producers

We now turn to the production side of the economy. We normalize the wage rate of manufacturing workers to 1:

$$w^M = 1 \quad t \geq 0 \quad (11)$$

The production of any variety, say h , requires the use of the patent specific to this variety, which has been developed in the R & D sector. Once a firm has acquired the patent at the market price (which corresponds to this firm's fixed cost), it can produce one unit of this variety by using one unit of \mathbb{M} -labor. When the manufacturer of variety h produces $q(h)$ units, the profit is

$$\pi(h) = [p(h) - 1]q(h)$$

which together with the demand function (2) yields the equilibrium price common to all varieties produced:

$$p^* = 1/\rho \quad (12)$$

Then, if M denotes the number of varieties produced at the time in question, substituting (12) into (3) yields

$$P = (1/\rho)(M)^{-1/(\sigma-1)} \quad (13)$$

Furthermore, substituting (12) and (13), we obtain the equilibrium output of any variety produced in the economy:

$$q^* = \rho E/M \quad (14)$$

whereas the equilibrium profit is given by

$$\pi^* = q^*/(\sigma - 1) \quad (15)$$

since $\rho = (\sigma - 1)/\sigma$.

We now study the labor market clearing conditions for the \mathbb{M} -workers. In equilibrium, labor demand is equal to labor supply, so

$$L = Mq^* \quad (16)$$

and, by (14),

$$L = \rho E \quad (17)$$

so that in equilibrium, the total expenditure

$$E^* = L/\rho \quad (18)$$

is independent of time since L is constant. Therefore, we may conclude from (10) that *the equilibrium interest rate is equal to the subjective discount rate over time*

$$\nu^*(t) = \gamma \quad \text{for all } t \geq 0 \quad (19)$$

As a result, using (9), the expenditure of any specific consumer i is also a constant, which is readily obtained from (7) and (19):

$$E_i^* = \gamma[\omega_i + W_i(0)] \quad (20)$$

Substituting (13) into (4) and setting $E_i(t) = E_i^*$ yields

$$v_i(t) = \rho \cdot E_i^* \cdot M(t)^{\frac{1}{\sigma-1}} \quad (21)$$

Using (5) and (21), we obtain the lifetime utility of consumer i as

$$U_i(0) = \frac{E_i^*}{\sigma} \int_0^\infty e^{-\gamma t} \ln(M(t)) dt \quad (22)$$

Finally, using (18) and (22), this sum of \mathbb{M} -workers' and \mathbb{K} -workers' utility is given by

$$W = \frac{L}{\sigma-1} \int_0^\infty e^{-\gamma t} \ln(M(t)) dt \quad (23)$$

2.3 R & D Sector

Production of a new manufactured commodity requires the purchase of a patent. These patents are produced by the R & D sector, consisting of N workers, and they are the only output of this sector. Each new patent embodies a new idea. Not all new ideas result in patents. New ideas are produced by \mathbb{K} -workers using their prior stock of knowledge. The scheme for producing new ideas is described as a knowledge production process.

The basic layout of this sector is similar to Berliant and Fujita (2008). There are, however, two major differences between the R & D sector detailed in our previous paper and the one we use here. First, in our previous work, the value of ideas was given exogenously, whereas in the present paper, the value of an idea is proportional to the price of a patent, an endogenous variable in the model. Second, the previous paper had no public source of knowledge, whereas the present paper has knowledge transmission to all \mathbb{K} -workers through the patent process. Due to these important differences, while avoiding excessive repetition, we present below the details of the R & D process.

At any given time, each \mathbb{K} -worker has a stock of knowledge that has some commonalities with other \mathbb{K} -workers but some knowledge distinct from other workers. Since workers possess knowledge exclusive of others, they may wish to cooperate with each other in the knowledge production process. Heterogeneity of knowledge in a partnership brings more originality, but knowledge in common is important for communication. Thus, \mathbb{K} -worker heterogeneity is an essential feature of the model and of the knowledge production process.

The \mathbb{K} -workers choose to work alone or with a partner, maximizing their myopic payoff, namely the value of patents produced at that time. The solution concept used is the myopic core. If they work alone, new ideas are produced as a function of the total number of ideas known by a \mathbb{K} -worker. If a pair of workers produces new ideas together, their knowledge production is a function of their knowledge in common on the one hand and the knowledge they have that is distinct from their partner on the other. Knowledge that is produced by an agent at a given time becomes part of the stock of knowledge for that agent in the future. In addition, some of these ideas become patented and are sold to the manufacturing sector. The ideas embodied in the patents become public, and will be learned by all the agents in the R & D sector.

The basic unit of knowledge is called an idea.⁷ The number of potential ideas is infinite. In this paper, we will treat ideas symmetrically. In describing the process of knowledge production, that is either accomplished alone or in cooperation with another \mathbb{K} -worker, the sufficient statistics about the state of knowledge of a \mathbb{K} -worker i at a given time can be described as follows. We shall focus on \mathbb{K} -worker i and her potential partner \mathbb{K} -worker j . First, $n_i(t)$ represents the total stock of i 's ideas at time t . Second, $n_{ij}^c(t)$ represents the total stock of ideas that i has in common with \mathbb{K} -worker j at time t . Third, $n_{ij}^d(t)$ represents the stock of ideas that i knows but j doesn't know at time t . Finally, $n_{ji}^d(t)$ represents the stock of ideas that j knows but i doesn't know at time t .

By definition, $n_{ij}^c(t) = n_{ji}^c(t)$.⁸ It also holds by definition that

$$n_i(t) = n_{ij}^c(t) + n_{ij}^d(t) \tag{24}$$

Knowledge is a set of ideas that are possessed by a person at a particular time. However, knowledge is not a static concept. New knowledge can be produced either individually or jointly, and ideas can be shared with others. But all of this activity takes time.

Now we describe the components of the rest of the model. To keep the description as simple as possible, we focus on just two agents, i and j . At each time, each agent faces a decision about whether or not to meet with others. If two agents want to meet at a particular time, a meeting will occur. If an agent decides not to meet with anyone at a given time, then the agent produces separately and also creates new knowledge separately, away from everyone

⁷In principle, all of these time-dependent quantities are positive integers. However, for simplicity we take them to be continuous (in \mathbb{R}_+) throughout the paper.

⁸In general, however, it is not necessary that $n_{ij}^d(t) = n_{ji}^d(t)$.

else. If two persons do decide to meet at a given time, then they collaborate to create new knowledge together.⁹

At each moment of time, there are two mutually exclusive ways to produce new knowledge. The first way is to work alone, away from others. We denote the event that \mathbb{K} -worker i does research alone at time t by $\delta_{ii}(t) = 1$, indicating that i works with herself. Otherwise, $\delta_{ii}(t) = 0$. Alternatively, \mathbb{K} -worker i can choose to work with a partner, say \mathbb{K} -worker j . We denote the event that \mathbb{K} -worker i wishes to work with j at time t by $\delta_{ij}(t) = 1$. Otherwise, $\delta_{ij}(t) = 0$. In equilibrium, this partnership is realized at time t if $\delta_{ij}(t) = \delta_{ji}(t) = 1$.

Consider first the case where \mathbb{K} -worker i works alone. In this case, idea production is simply a function of the stock of i 's ideas at that time. Let $a_{ii}(t)$ be the rate of production of new ideas created by person i in isolation at time t . Then we assume that the creation of new knowledge during isolation is governed by the following equation:

$$a_{ii}(t) = \alpha \cdot n_i(t) \text{ when } \delta_{ii}(t) = 1. \quad (25)$$

If a meeting occurs between i and j at time t ($\delta_{ij}(t) = \delta_{ji}(t) = 1$), then joint knowledge creation occurs, and it is governed by the following dynamics:¹⁰

$$a_{ij}(t) = 2\beta \cdot (n_{ij}^c)^\theta \cdot (n_{ij}^d \cdot n_{ji}^d)^{\frac{1-\theta}{2}} \text{ when } \delta_{ij}(t) = \delta_{ji}(t) = 1 \text{ for } j \neq i \quad (26)$$

where $0 < \theta < 1$, $\beta > 0$. So when two people meet, joint knowledge creation occurs at a rate proportional to the normalized product of their knowledge in common, the differential knowledge of i from j , and the differential knowledge of j from i . The rate of creation of new knowledge is high when the proportions of ideas in common, ideas exclusive to person i , and ideas exclusive to person j are in balance. The parameter θ represents the weight on knowledge in common as opposed to differential knowledge in the production of new ideas. Ideas in common are necessary for communication, whereas ideas exclusive to one person or the other imply more heterogeneity or originality in the collaboration.

⁹Since there is an infinity of potential ideas, the probability that the same idea is duplicated by any \mathbb{K} -worker or \mathbb{K} -workers (even at different points of time) is assumed to be zero.

¹⁰We may generalize equation (26) as follows:

$$a_{ij}(t) = \max \left\{ (\alpha - \varepsilon) n_i(t) + (\alpha - \varepsilon) n_j(t), 2\beta \cdot (n_{ij}^c)^\theta \cdot (n_{ij}^d \cdot n_{ji}^d)^{\frac{1-\theta}{2}} \right\}$$

where $\varepsilon > 0$ represents the costs from the lack of concentration. This generalization, however, does not change the results presented in this paper in any essential way.

Income for the research sector derives from selling patents. But not all ideas are patentable. For every collection of ideas created, we assume that η proportion are patentable as blueprints of new products. Thus, they are sold to the manufacturing sector. The residual ideas, namely $1 - \eta$ proportion of new ideas, becomes tacit knowledge that is only known to the creator or creators of these ideas. They are useful for future creation of yet further ideas.

Let $y_i(t)$ to be the income of \mathbb{K} -worker i at time t , and let $\Pi(t)$ be the price of patents at time t . Then, suppressing t for notational simplicity:

$$y_i = \Pi \cdot \eta \cdot (\delta_{ii} \cdot a_{ii} + \sum_{j \neq i} \delta_{ij} \cdot a_{ij}/2) \quad (27)$$

The formula implies that the revenue from new patents is split evenly if two \mathbb{K} -workers are producing new ideas together.

Concerning the rule used by an agent to choose their best partner, to keep the model tractable in this first analysis, we assume a myopic rule. At each moment of time t , person i would like a meeting with person j when her income while meeting with j is highest among all potential partners, including herself. Maximizing income at a given time amounts to choosing $\{\delta_{ij}\}_{j=1}^N$ so that the right hand side of (27) is highest, meaning that a selection is made only among the most productive partners. Loosely speaking, this interaction could be modeled as a noncooperative game, with player i choosing $\{\delta_{ij}\}_{j=1}^N$ as strategies, and equilibrium implying that for each pair of players i and j , $j \neq i$, $\delta_{ij} = \delta_{ji}$, whereas $\delta_{ij} > 0$ only for those players j that yield maximal payoffs for player i .¹¹

This noncooperative approach is useful for explaining the ideas behind our model, but we employ a cooperative approach for two reasons. First, it gives the same equilibrium path as the noncooperative approach but with less cumbersome notation and structure. Second, as we are attempting to model close interactions within groups, it is plausible that agents will act cooperatively. We assume that at each time, the myopic persons interacting choose a core configuration. That is, we restrict attention to configurations such that at any point in time, no coalition of persons can get together and make themselves better off *in that time period*. In essence, our solution concept at a point in time is the myopic core.

¹¹More formally, out of equilibrium payoffs are defined and a selection or refinement of Nash equilibrium used as in Berliant et al. (2006, pp. 77-78). A refinement of Nash equilibrium is necessary to exclude some trivial equilibria, for example where nobody ever chooses to meet anyone else.

All agents take prices, in this case Π , as given, implying:

$$\max_{\{\delta_{ij}\}_{j=1}^N} (\delta_{ii} \cdot a_{ii} + \sum_{j \neq i} \delta_{ij} \cdot a_{ij}/2) \quad (28)$$

subject to the obvious constraints:

$$\sum_{j=1}^N \delta_{ij} = 1, \delta_{ij} \geq 0 \text{ for } i = 1, \dots, N \quad (29)$$

Since n_i is a stock variable, this is equivalent to

$$\max_{\{\delta_{ij}\}_{j=1}^N} \left(\frac{\delta_{ii} \cdot a_{ii} + \sum_{j \neq i} \delta_{ij} \cdot a_{ij}/2}{n_i} \right) \quad (30)$$

In order to rewrite this problem in a convenient form, we first define the total number of ideas possessed by i and j :

$$n^{ij} = n_{ij}^d + n_{ji}^d + n_{ij}^c \quad (31)$$

and define new variables

$$\begin{aligned} m_{ij}^c &\equiv m_{ji}^c = \frac{n_{ij}^c}{n^{ij}} = \frac{n_{ji}^c}{n^{ij}} \\ m_{ij}^d &= \frac{n_{ij}^d}{n^{ij}}, m_{ji}^d = \frac{n_{ji}^d}{n^{ij}} \end{aligned}$$

By definition, m_{ij}^d represents the proportion of ideas exclusive to person i among all the ideas known by person i or person j . Similarly, m_{ij}^c represents the proportion of ideas known in common by persons i and j among all the ideas known by the pair. From (31), we obtain

$$1 = m_{ij}^d + m_{ji}^d + m_{ij}^c \quad (32)$$

whereas (31) and (24) yield

$$n_i = (1 - m_{ji}^d) \cdot n^{ij} \quad (33)$$

Using these identities and new variables, while recalling the knowledge production function (26) and (25), we obtain (see Technical Appendix a for details)

$$\frac{a_{ij}/2}{n_i} = G(m_{ij}^d, m_{ji}^d) \quad (34)$$

where

$$G(m_{ij}^d, m_{ji}^d) \equiv \frac{\beta (1 - m_{ij}^d - m_{ji}^d)^\theta \cdot (m_{ij}^d \cdot m_{ji}^d)^{\frac{1-\theta}{2}}}{1 - m_{ji}^d} \quad (35)$$

Using (25) and (34), we can rewrite the income function (27) as

$$y_i = \Pi \cdot \eta \cdot n_i \cdot (\delta_{ii} \cdot \alpha + \sum_{j \neq i} \delta_{ij} \cdot G(m_{ij}^d, m_{ji}^d)) \quad (36)$$

and the optimization problem (30) as follows:

$$\max_{\{\delta_{ij}\}_{j=1}^N} (\delta_{ii} \cdot \alpha + \sum_{j \neq i} \delta_{ij} \cdot G(m_{ij}^d, m_{ji}^d)) \quad (37)$$

subject to the obvious constraints (29).

Suppose that for each $i = 1, 2, \dots, N$, $\{\delta_{ij}^*\}_{j=1}^N$ solves the optimization problem immediately above. Furthermore, suppose that it happens to be the case that

$$\delta_{ij}^* = \delta_{ji}^* \text{ for } i, j = 1, 2, \dots, N$$

Then, by construction, $\{\delta_{ij}^*\}_{i,j=1}^N$ must also be the solution to the following social optimization problem:

$$\max \left\{ \sum_{i=1}^N y_i \mid \sum_{j=1}^N \delta_{ij} = 1, \delta_{ij} \geq 0, \delta_{ij} = \delta_{ji} \text{ for } i, j = 1, 2, \dots, N \right\}$$

Thus, $\{\delta_{ij}^*\}_{i,j=1}^N$ is in the myopic core.

We now describe the dynamics of the knowledge system, dropping the time argument. There are two ways to acquire new knowledge for a \mathbb{K} -worker: internal production of new ideas and information from public sources. The first way has the feature that ideas produced alone are attributed to that worker, whereas ideas produced in pairs are attributed to both \mathbb{K} -workers who produce them. In either case, the new ideas are learned by exactly the people who produce them. The second source of knowledge acquisition derives from the new ideas that are patented. The patented ideas become public information. A certain proportion of patented ideas, $\mu(N)$, are learned by all of the \mathbb{K} -workers. In general, $\mu(N)$ will be a decreasing function of N . Limited time and energy determine how many of these new, public ideas can be learned. Due to these limitations, the amount of information a \mathbb{K} -worker can learn from patents at a given time is, roughly, proportional to the number of new ideas she can create in that time. The number of new ideas and thus patents is proportional to the number of \mathbb{K} -workers, so $\mu(N)$ will be inversely proportional to N .¹² Thus, these ideas become knowledge in common for all

¹²In theory, it might be possible to accumulate a stock of ideas patented in past periods to learn in the future. The problem with this is that such information perpetually accumulates, and thus due to time constraints there is never an opportunity to learn the content of older patented ideas.

agents in the research sector.¹³ The net result is an increase in n_{ij}^c for all i and j of $\mu(N) \cdot \eta$ proportion of new ideas created in the economy. The workers in the R & D sector see this flow of new ideas from patents, and account for it when they choose actions at each moment of time. To obtain an expression relating \dot{m}_{ij}^d to m_{ij}^d , we must first examine the knowledge dynamics in terms of the original variables, n_i , n_{ij}^c , and n_{ij}^d .

Let us focus on agent i , as the expressions for the other agents are analogous. Let A be the total number of ideas created at a given moment:

$$A = \sum_{k=1}^N \delta_{kk} \cdot a_{kk} + \left(\sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot a_{kl} \right) / 2 \quad (38)$$

$$= \sum_{k=1}^N \delta_{kk} \cdot \alpha n_k + \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot n_k \cdot G(m_{kl}^d, m_{lk}^d) \quad (39)$$

The dynamics of the knowledge system are based on the assumption that once learned, ideas are not forgotten. Using the argument above, we obtain knowledge system dynamics:

$$\dot{n}_i = \sum_{l=1}^N \delta_{il} \cdot a_{il} + \mu(N) \cdot \eta \cdot \left(A - \sum_{l=1}^N \delta_{il} \cdot a_{il} \right) \quad (40)$$

$$\dot{n}_{ij}^c = \delta_{ij} \cdot a_{ij} + \mu(N) \cdot \eta \cdot \left(A - \delta_{ij} \cdot a_{ij} \right) \text{ for all } j \neq i \quad (41)$$

$$\dot{n}_{ij}^d = \left(1 - \mu(N) \cdot \eta \right) \cdot \sum_{k \neq j} \delta_{ik} \cdot a_{ik} \text{ for all } j \neq i \quad (42)$$

Thus, equation (40) says that the increase in the knowledge of person i is the sum of: the knowledge created in isolation, the knowledge created jointly with someone else, and the transfer of new knowledge from new patents. Equation (41) means that the increase in the knowledge in common for persons i and j equals the new knowledge created jointly by them plus the transfer of knowledge from new patents. Finally, equation (42) means that all the knowledge created by person i either in isolation or joint with persons other than person j becomes a part of the differential knowledge of person i from person j , *except* for patented ideas that are learned by all \mathbb{K} -workers.

¹³It has been suggested that if \mathbb{K} -workers become too homogeneous, they might learn the patented ideas selectively so as not to overlap with the knowledge acquired by other \mathbb{K} -workers in the same fashion. However, this level of coordination, especially when N is large, seems far-fetched. It seems more likely that ideas attractive for whatever reason will be learned by all.

Using (25) and (34), equation (40) can be rewritten as

$$\dot{n}_i = (1 - \mu(N) \cdot \eta) \cdot n_i \cdot (\delta_{ii}\alpha + 2 \sum_{l \neq i}^N \delta_{il} \cdot G(m_{il}^d, m_{li}^d)) + \mu(N) \cdot \eta \cdot A \quad (43)$$

where A is given by (39). Furthermore, using (25), (26), and (34), we have (see Theorem A2 of Technical Appendix a)

$$\begin{aligned} \dot{m}_{ij}^d = & (1 - \mu \cdot \eta)(1 - m_{ij}^d)(1 - m_{ji}^d) \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot 2G(m_{ik}^d, m_{ki}^d) \right\} \\ & - m_{ij}^d [\mu \eta \alpha (1 - m_{ji}^d) \cdot \sum_{k=1}^N \delta_{kk} \cdot \frac{n_k}{n_i} + (1 - \mu \cdot \eta) \cdot \delta_{ij} \cdot (1 - m_{ji}^d) \cdot 2G(m_{ij}^d, m_{ji}^d) \\ & + \mu \cdot \eta \cdot (1 - m_{ji}^d) \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot \frac{n_k}{n_i} \cdot G(m_{kl}^d, m_{lk}^d)] \\ & - m_{ij}^d \cdot (1 - \mu \cdot \eta) \cdot (1 - m_{ij}^d) \cdot \left\{ \delta_{jj} \cdot \alpha + \sum_{k \neq i, j} \delta_{jk} \cdot 2G(m_{jk}^d, m_{kj}^d) \right\} \end{aligned} \quad (44)$$

for $i, j = 1, 2, \dots, N$, $i \neq j$. Thus, using (43) and (44), the knowledge dynamics are described in terms of n_i and m_{ij}^d ($i, j = 1, \dots, N$) only.

3 Knowledge Dynamics

3.1 The Model

Since we are concerned with the macro behavior of the economy and the big picture in terms of growth, we make a number of simplifying assumptions. We impose the assumption that the initial state of knowledge for all \mathbb{K} -workers is pairwise symmetric in terms of heterogeneity. The initial state of knowledge is given by

$$n_{ij}^c(0) = n^c(0) \text{ for all } i \neq j \quad (45)$$

$$n_{ij}^d(0) = n^d(0) \text{ for all } i \neq j \quad (46)$$

implying that

$$n_i(0) = n^c(0) + n^d(0) \equiv n(0) \quad (47)$$

At the initial state, each pair of \mathbb{K} -workers has the same number of ideas, $n^c(0)$, in common. Moreover, for any pair of \mathbb{K} -workers, the number of ideas that one \mathbb{K} -worker knows but the other does not know is the same and equal

to $n^d(0)$. Given that the initial state of knowledge is symmetric among the \mathbb{K} -workers, as seen below, it turns out that the equilibrium configuration at any time also maintains the basic pairwise symmetry among \mathbb{K} -workers.

Suppose that at some given time, all \mathbb{K} -workers are pairwise symmetric to each other. Namely, when

$$m_{ij}^d = m_{ji}^d \text{ for all } i \neq j \quad (48)$$

(37) is simplified as

$$\max_{\{\delta_{ij}\}_{j=1}^N} (\delta_{ii} \cdot \alpha + \sum_{j \neq i} \delta_{ij} \cdot g(m_{ij}^d)) \quad (49)$$

where the function g is defined as

$$g(m) \equiv G(m, m) \equiv \beta \frac{(1 - 2m)^\theta m^{(1-\theta)}}{1 - m} \quad (50)$$

Since $n^{ij} = n^{ji}$ by definition, we can readily see, by using (33), that condition (48) is equivalent to

$$n_i = n_j \text{ for all } i \text{ and } j \quad (51)$$

Furthermore, since $a_{ij} = a_{ji}$ by definition, substituting (48) into (34) yields

$$\frac{a_{ij}/2}{n_i} = \frac{a_{ji}/2}{n_j} = g(m_{ij}^d) \quad (52)$$

Thus, when two \mathbb{K} -workers i and j cooperate in knowledge production and their knowledge states are symmetric, $g(m_{ij}^d)$ represents the creation of new ideas per capita (normalized by the size of individual knowledge input, n_i). In this context, condition (49) means that each \mathbb{K} -worker wishes to engage in knowledge production in a partnership with a person (possibly including herself) leading to the highest \mathbb{K} -productivity.

Figure 1 illustrates the graph of the function $g(m)$ as a bold curve for parameter values $\beta = 1$ and $\theta = 1/3$.

FIGURE 1 GOES HERE

Differentiating $g(m)$ yields

$$g'(m) = g(m) \cdot \frac{(1 - \theta) - (2 - \theta) \cdot m}{(1 - 2m) \cdot m \cdot (1 - m)}$$

implying that

$$g'(m) \begin{cases} > 0 \\ < 0 \end{cases} \text{ as } m \begin{cases} < \frac{1-\theta}{2-\theta} \\ > \frac{1-\theta}{2-\theta} \end{cases} \text{ for } m \in (0, \frac{1}{2}) \quad (53)$$

Thus, $g(m)$ is strictly quasi-concave on $[0, 1/2]$, achieving its maximal value at

$$m^B = \frac{1-\theta}{2-\theta} \quad (54)$$

which we call the ‘‘Bliss Point.’’ It is the point where knowledge productivity is highest for each person. In the remainder of the paper, our main concern is whether or not the dynamics of knowledge interaction will, starting at the initial state given by (45) and (46), lead the system of \mathbb{K} -workers to this bliss point.

When condition (48) holds, using (50) and (51), the dynamics can be written as

$$\begin{aligned} \frac{\dot{m}_{ij}^d}{1-m_{ij}^d} &= (1-\mu \cdot \eta) \cdot (1-m_{ij}^d) \cdot \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i,j} \delta_{ik} \cdot 2g(m_{ik}^d) \right\} \\ &\quad - m_{ij}^d \left\{ \mu \cdot \eta \cdot \alpha \cdot \sum_{k=1}^N \delta_{kk} + (1-\mu \cdot \eta) \cdot \delta_{ij} \cdot 2g(m_{ij}^d) + \mu \cdot \eta \cdot \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot g(m_{kl}^d) \right\} \\ &\quad - m_{ij}^d (1-\mu \cdot \eta) \cdot \left\{ \delta_{jj} \cdot \alpha + \sum_{k \neq i,j} \delta_{jk} \cdot 2g(m_{jk}^d) \right\} \end{aligned} \quad (55)$$

$$\dot{n}_i = (1-\mu(N) \cdot \eta) \cdot n_i \cdot \left(\delta_{ii} \cdot \alpha + 2 \sum_{k \neq i} \delta_{ik} \cdot g(m_{ik}^d) \right) + \mu(N) \cdot \eta \cdot A \quad (56)$$

for $i, j = 1, 2, \dots, N$, where A is given by

$$A = \sum_{k=1}^N \delta_{kk} \cdot \alpha n_k + \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot n_k \cdot g(m_{kl}^d) \quad (57)$$

We observe that the basic rules, (49), (55), and (43), that govern the knowledge dynamics are described in terms of m_{ij}^d and n_i ($i, j = 1, 2, \dots, N$) only. Notice that no market variable is used. This enables us first to solve for the equilibrium path of knowledge dynamics independent of commodity and capital markets.

Next, we discuss the possible equilibrium configurations of partnerships in knowledge creation, noting that the equilibrium configuration can vary with time.¹⁴

¹⁴For details in the case $N = 4$, please refer to Berliant and Fujita (2008).

Although knowledge creation in isolation or in pairs represents the basic forms of knowledge creation, it turns out that the equilibrium path often requires a mixture of these basic forms, namely δ_{ij} takes on fractional values. The reason is that on the equilibrium path, \mathbb{K} -workers wish to form groups where close interaction takes place in pairs within the group but there is no direct interaction between groups. \mathbb{K} -workers in the same group wish to change partners within the group as frequently as possible. The purpose is to balance the proportion of different and common ideas with partners within the same group as best as can be achieved. This suggests a work pattern with rapidly changing partners on the equilibrium path, that is, a work pattern where a worker rotates through fixed partners as fast as possible in order to maximize the instantaneous increase in income. For example, worker 1 chooses \mathbb{K} -workers 2 and 3 as partners, and rotates between the two partners under equilibrium values of δ_{12} and δ_{13} such that $\delta_{12} + \delta_{13} = 1$. Worker 1 might wish to work with workers 2 and 3 for half of each month, but wants to alternate between them so that worker 1 does not have the same partner on consecutive days. As time intervals in this discrete time model become shorter, the limit is a fractional δ_{1j} ($j = 2, 3$) where $\delta_{12} = \delta_{13} = 1/2$. Other \mathbb{K} -workers behave analogously. In order for this type of work pattern to take place, of course, all persons must agree to follow this pattern. In general, we allow $\delta_{ij} \in [0, 1]$, and for all i , $\sum_{j \neq i} \delta_{ij} = 1$. In equilibrium, $\delta_{ij} = \delta_{ji}$ for all $i, j = 1, 2, \dots, N$.

At this point, it is useful to remind the reader that we are using a myopic core concept to determine equilibrium at each point in time. In fact, it is necessary to sharpen that concept in the model with N persons. *When there is more than one vector of strategies that is in the myopic core at a particular time, namely more than one vector of joint strategies implies the same, highest income for all persons, the one with the highest first derivative of income \dot{y}_i is selected.* Furthermore, when the derivative of income is still the same among best options, agent i chooses an option that maximizes the second derivative of income, \ddot{y}_i , and so on. The justification for this assumption is that at each point in time, people are attempting to maximize the flow of income. The formal definition of the myopic core and proof that it is nonempty can be found in Berliant and Fujita (2008, Appendix 0). Although the theorem is general, in the remainder of this paper we shall focus on the symmetric case.

3.2 Equilibrium Path of Knowledge Dynamics

Now we are ready to investigate the actual equilibrium path, depending on the given initial composition of knowledge,

$$m_{ij}^d(0) = m^d(0) = \frac{n^d(0)}{n^c(0) + 2n^d(0)}$$

which is common for all pairs i and j ($i \neq j$). In order to sharpen the results that follow, we introduce a specific form of the parametric function $\mu(N)$, representing the proportion of the public information on new patents that is actually learned by \mathbb{K} -workers as knowledge in common. Assuming that the flow of knowledge that each \mathbb{K} -worker can acquire from public information on new patents is proportional to the flow of new knowledge she can produce, we use the following relation in the analysis below (see Appendix 1.1 for a justification):

$$\mu(N) = \frac{C}{\eta N}$$

or

$$\mu(N) \cdot \eta = \frac{C}{N} \tag{58}$$

where C is a positive constant representing the learning capacity (l -capacity) of each \mathbb{K} -worker.

The reasoning behind this assumption is as follows. Larger populations mean, in general, that more ideas are produced. But the capacity of agents to learn these ideas is limited by their time and ability. Thus, the fraction of public knowledge absorbed is assumed to scale at C/N . This assumption contributes to the result, stated and proved below, that the growth rates in the model are scale free.

In the rest of the paper, we assume that N is sufficiently large, implying:

$$N > C \tag{59}$$

In the remainder of this paper, we also assume that

$$\alpha < g(m^B) \tag{60}$$

so as to avoid the trivial case of all agents always working in isolation.

In Figure 1, let m^J and m^I be defined on the horizontal axis at the left intersection and the right intersection between the $g(m)$ curve and the horizontal line at height α , respectively.

In the following analysis, we shall characterize the equilibrium path of knowledge creation dynamics. The various cases are determined by the initial heterogeneity of the \mathbb{K} -workers. To be precise:

Proposition 1: *The equilibrium path of \mathbb{K} -worker interactions and the sink point of the knowledge creation process depend discontinuously on the initial condition, $m^d(0)$. Assuming that the number of \mathbb{K} -workers N is large, the pattern of interaction between \mathbb{K} -workers and the sink point as a function of the initial condition are as follows.*

(i) For $m^J < m^d(0) \leq m^B$, we define two subcases. Let $\tilde{C} \equiv \frac{2\theta}{1-\theta}$.

(a) $C < \tilde{C}$. The equilibrium path consists of an initial time interval in which each \mathbb{K} -worker is always paired with another but trades partners as rapidly as possible (with $\delta_{ij} = 1/(N-1)$ for all i and for all $j \neq i$). When the bliss point, $m^B = \frac{1-\theta}{2-\theta}$, is attained, the agents split into groups of $\tilde{N}^B = 1 + \frac{1}{\theta - \frac{(1-\theta)C}{2}}$, and they remain at the bliss point.

(b) $C > \tilde{C}$. The equilibrium path has all \mathbb{K} -workers paired with another but trading partners as rapidly as possible (with $\delta_{ij} = 1/(N-1)$ for all i and for all $j \neq i$). This continues forever. The equilibrium path remains to the left of the bliss point, so the bliss point is never attained. The sink point is $\tilde{m}^{d*} = \frac{1}{2 + \frac{C}{2}}$.

(ii) $m^d(0) < m^J < m^B$. Once again, there are two subcases. If C is large, then all \mathbb{K} -workers are in isolation producing new ideas alone forever. The sink point is $\tilde{m}^{d**} = \frac{1}{2+C}$. If C is not large, then the equilibrium path consists of a first phase in which all \mathbb{K} -workers are in isolation producing new ideas. Once the system reaches m^J , the equilibrium path follows that given in case (i).

(iii) $m^B < m^d(0)$ The equilibrium path consists of many phases. First, the N \mathbb{K} -workers are paired arbitrarily and work with their partners for a nonempty interval of time. Second, they switch to new partners and work with their new partners for a nonempty interval of time. Third, each \mathbb{K} -worker pairs alternately with the two partners with whom they worked in the first two phases, but not with a \mathbb{K} -worker with whom they have not worked previously. This process continues, possibly adding more partners.

We wish to alert the reader that the focus of the remainder of the paper, in particular our analysis of economic growth, will be on case (i). Thus, we shall not discuss the other cases in great detail.

3.2.1 Case (i): $m^J < m^d(0) \leq m^B$

We summarize the intuition for the result here, leaving the proof and detailed explanation to Appendix 1.2.

First suppose that the initial state is such that

$$m^J < m^d(0) < m^B$$

The condition $m^d < m^B$ means that the \mathbb{K} -workers have relatively too many ideas in common, and thus they wish to acquire ideas that are different from those of each possible partner as fast as possible. That is, when $m^J < m_{ij}^d = m^d < m^B$ in Figure 1, each \mathbb{K} -worker wishes to move the knowledge composition m_{ij}^d to the right as quickly as possible, thus increasing the \mathbb{K} -productivity $g(m_{ij}^d)$ as fast as possible.

More precisely, when $m^J < m^d(0) = m_{ji}^d(0) < m^B$ for all $i \neq j$, on the equilibrium path, each \mathbb{K} -worker i spends the same amount of time

$$\delta_{ij} = \frac{1}{N-1} \text{ for all } j \neq i \quad (61)$$

with every other \mathbb{K} -worker at the start. This pattern allows each \mathbb{K} -worker to spend the least possible time with each specific partner, thus avoiding the buildup of knowledge in common with any other particular \mathbb{K} -worker. Then, since symmetry holds from the start onward, the same work pattern will continue as long as $m^J < m^d < m^B$. The dynamics of this work pattern are as follows. The creation of new ideas always takes place in pairs. Pairs are cycling rapidly with $\delta_{ij} = 1/(N-1)$ for all $j \neq i$. \mathbb{K} -worker 1, for example, spends $1/(N-1)$ of each period with \mathbb{K} -worker 2, for example, and $(N-2)/(N-1)$ of the time working with other partners.

Under the dynamics (61), $m_{ij}^d = m^d$ for all $i \neq j$ and the sink point of the process is given by¹⁵

$$\tilde{m}^{d*} = \frac{1}{2 + \frac{C}{2}} \quad (62)$$

In the upper half of Figure 2, the \mathbb{K} -productivity curve $g(m)$ is transferred from Figure 1. In the bottom half of Figure 2, the bold curve depicts the sink, \tilde{m}^{d*} , as a function of the l -capacity parameter C .

FIGURE 2 GOES HERE

¹⁵In the remainder of the paper, we take the number of \mathbb{K} -workers to be large, namely $N \rightarrow \infty$.

In the context of Figure 2, we can identify two different possibilities. Suppose that

$$m^B < \tilde{m}^{d*} \quad (63)$$

That is, the sink point of the dynamics is on the right side of the bliss point. In this case, beginning at any point $m^J < m^d(0) < m^B$, the system reaches the bliss point in finite time. In terms of the original parameters, using (54) and (62), condition (63) can be rewritten as

$$C < \tilde{C} \equiv \frac{2\theta}{1-\theta} \quad (64)$$

In Figure 2, C_1 provides an example of this case. The associated sink point is given by m_1^{d*} .

In contrast, suppose that

$$\tilde{m}^{d*} < m^B \quad (65)$$

This occurs exactly when inequality in (64) is reversed. In Figure 2, C_2 represents an example of such a value of C , whereas the associated sink point is given by m_2^{d*} . In this case, starting with any initial point $m^J < m^d(0) < m^B$, the system moves automatically toward $m^{d*} < m^B$, but never reaches the bliss point.

On the downward vertical axis of Figure 2, \tilde{C} gives the value of the parameter C at the boundary of the two cases. Case (63) occurs exactly when the value of the l -capacity C is relatively small, whereas case (65) occurs when C is relatively large. In what follows, under the assumption that N is large, we examine the actual dynamics in each of the two cases.

Case (i-a): $m^J < m^d(0) \leq m^B$ **and** $C < \tilde{C}$ This equilibrium configuration of partnerships at the bliss point m^B can be achieved as follows: When the system reaches m^B , the population splits into smaller groups of equal size,

$$\tilde{N}^B = 1 + \frac{1}{\theta - \frac{(1-\theta) \cdot C}{2}} \quad (66)$$

That is, when all \mathbb{K} -workers reach the bliss point, they stay there by splitting into smaller groups of the same size, \tilde{N}^B , so direct interactions take place only within each group. In this way, each \mathbb{K} -worker maintains the highest \mathbb{K} -productivity while enjoying the knowledge externalities derived from public information on new patents. Figure 3 depicts an example of an equilibrium configuration of \mathbb{K} -worker interactions in which four groups of \mathbb{K} -workers form at the bliss point. The dotted arrows represent indirect interactions through the public revelation of patent information.

FIGURE 3 GOES HERE

The optimal group size for large population \tilde{N}^B increases monotonically with the l -capacity, C ; as C increases, the transmission of public knowledge in common increases, so it is necessary to have a larger group in order to maintain heterogeneity among agents within the group. Recalling that \tilde{C} was defined in (64), the group size becomes infinitely large as C approaches \tilde{C} from the left. Recalling that θ is the weight given to knowledge in common in the \mathbb{K} -production function, as the value of θ increases, \tilde{N}^B decreases, which is not surprising. In Figure 4, for each fixed value of the parameter C , the optimal group size \tilde{N}^B is graphed as a function of θ .

FIGURE 4 GOES HERE

When the system reaches the bliss point, the workers break into groups and the system becomes asymmetric, in the following sense. If \mathbb{K} -worker i belongs to the same group as \mathbb{K} -worker k , then their differential knowledge remains at the bliss point m^B , maintaining the highest \mathbb{K} -productivity $g(m^B)$. If \mathbb{K} -worker j belongs to a different group, then the differential knowledge between i and j diverges, namely it moves away from m^B , thus reducing $g(m_{ij}^d)$. So once the population splits into groups, \mathbb{K} -workers i and j will not want to collaborate again.

To sum up, for partnerships of \mathbb{K} -workers within the same group, their productivity is $g(m^B)$. For potential partnerships of \mathbb{K} -workers in different groups, their potential productivity is $g(\tilde{m}^{d*}) < g(m^B)$. So these potential partnerships are never formed. The implication is that we have endogenous formation of cohesive groups. One interpretation of this phenomenon is that the groups represent research firms, so we have endogenous formation of research firm boundaries.

Case (i-b): $m^J < m^d(0) \leq m^B$ and $C > \tilde{C}$ As explained previously, in this case the dynamics imply that only one large group forms, so each agent works with everyone else an equal amount of time. Heterogeneity m^d changes, approaching the sink point \tilde{m}^{d*} given by (62) to the left of the bliss point, so the bliss point is never reached. In this case

$$m^J < \tilde{m}^{d*} < m^B$$

and one large group is maintained forever, without achieving the highest possible productivity. Intuitively, this is due to the large externality from public knowledge, so it is impossible to attain sufficient heterogeneity.¹⁶

3.2.2 Case (ii): $m^d(0) < m^J < m^B$

Please refer to Appendix 1.3 for more details. Under this set of parameters, $g(m^d(0)) < \alpha$. In other words, at time 0 it is best for everyone to work in isolation rather than in pairs. The sink point of this dynamic is given by

$$\tilde{m}^{d**} = \frac{1}{2 + C}$$

Evidently,

$$\tilde{m}^{d**} < \tilde{m}^{d*}$$

Focusing on this case, there are two possibilities, namely $m^J < \tilde{m}^{d**}$ and $m^J > \tilde{m}^{d**}$.¹⁷ Assuming C is not too large, we concentrate on the first possibility,

$$m^J < \tilde{m}^{d**} \tag{67}$$

The equilibrium path has every \mathbb{K} -worker in isolation to begin, creating new knowledge on their own and moving to the right until they all reach the point m^J . Then one large group forms and all \mathbb{K} -workers create new knowledge working in pairs where each spends equal time with every other. From here, the equilibrium path is exactly the same as in case (i).

3.2.3 Case (iii): $m^B < m^d(0)$

As this case is not the focus of our work, we have relegated it to Appendix 1.4.

4 Growth

Next we assemble the various pieces of our general equilibrium model. Our focus is on case (i-a) of the knowledge dynamics, where the initial state of knowledge heterogeneity is to the left of the bliss point: $m^J < m^d(0) < m^B$.

¹⁶When C is very large, it is possible that $m^{d*} < m^J$, implying that all \mathbb{K} -workers eventually work in isolation. However, this is not an interesting situation, so we neglect it in the discussion that follows.

¹⁷Under the second possibility, $\tilde{m}^{d**} < m^J$, each \mathbb{K} -worker creates knowledge in isolation forever, approaching the sink point \tilde{m}^{d**} .

Proposition 2: Assume that the number of \mathbb{K} -workers N is large and $m^J < m^d(0) < m^B$.¹⁸ Let $\tilde{C} \equiv \frac{2\theta}{1-\theta}$. Then the long run economic growth as a function of the initial condition is given by two subcases.

(a) $C < \tilde{C}$. Let t^B be the time that all \mathbb{K} -workers reach the bliss point m^B . Then

$$\frac{\dot{n}(t)}{n(t)} = g(m^B)(2 + C) = \beta\theta^\theta(1 - \theta)^{1-\theta} \cdot (2 + C) \text{ for } t \geq t^B.$$

Moreover,

$$\lim_{t \rightarrow \infty} \frac{\dot{M}(t)}{M(t)} = \beta\theta^\theta(1 - \theta)^{1-\theta} \cdot (2 + C)$$

and

$$\lim_{t \rightarrow \infty} \frac{\dot{v}_i(t)}{v_i(t)} = \frac{\beta\theta^\theta(1 - \theta)^{1-\theta} \cdot (2 + C)}{\sigma - 1}.$$

(b) $C > \tilde{C}$. Then

$$\lim_{t \rightarrow \infty} \frac{\dot{n}(t)}{n(t)} = \lim_{t \rightarrow \infty} \frac{\dot{M}(t)}{M(t)} = g(\tilde{m}^{d**})(2 + C) = 2\beta\left(\frac{C}{2}\right)^\theta$$

whereas

$$\lim_{t \rightarrow \infty} \frac{\dot{v}_i(t)}{v_i(t)} = \frac{2\beta\left(\frac{C}{2}\right)^\theta}{\sigma - 1}.$$

Therefore the long run economic growth rate is positively related to both β , the parameter reflecting \mathbb{K} -productivity of work in pairs, and C , the speed of public knowledge transmission.

It is evident that, in either case, the long run equilibrium growth rates are scale free in the population of \mathbb{K} -workers, N . This is a consequence of the assumption that knowledge absorption from public sources is inversely proportional to population, and the result that long run optimal group size of \mathbb{K} -workers is independent of population. We examine each case more closely below.

4.1 Case (i-a): $m^J < m^d(0) \leq m^B$ and $C < \tilde{C}$

Please refer to Appendix 1.5 for the formal proof.

It is not surprising that β , the coefficient on the joint knowledge production function, is positively related to the growth of the economy. In contrast, it

¹⁸For the sake of simplicity, the assumption that N is sufficiently large means that the results stated here take N to be set to ∞ .

is surprising that C is positively related to economic growth. On the face of it, when C is higher, agents become relatively homogeneous quicker, since the public transmission of patent knowledge is faster. In theory, it could be the case that the result is lower \mathbb{K} -productivity and thus lower economic growth because the higher homogeneity reduces knowledge productivity. This was our initial conjecture. However, in the model, as indicated by (66), the group size at the bliss point adjusts optimally to the speed of public transmission of knowledge. Group size increases to offset the higher speed of public knowledge transmission and the resulting increase in group homogeneity. The effect of larger groups at the bliss point is to create more heterogeneity within groups, thus maintaining higher economic growth. This is, in essence, a general equilibrium effect that allows the economy to take advantage of a higher speed of public information transmission.¹⁹

4.2 Case (i-b): $m^J < m^d(0) \leq m^B$ and $C > \tilde{C}$

Please refer to Appendix 1.6 for the formal proof.

Again, the asymptotic growth rate of individual knowledge stock, patents, and indirect utility are constants (different from case (i-a)), and depend positively on β and C . The surprising result here is that even though the system does not achieve the bliss point, a higher rate of public knowledge transmission results in higher economic growth. Even though \tilde{m}^{d*} decreases as C increases, and thus the productivity of partnerships $g(\tilde{m}^{d*})$ declines, notice that $g(\tilde{m}^{d*})$ represents the normalized productivity of partnerships. In fact, the total productivity of partnerships is $n \cdot g(\tilde{m}^{d*})$. In the end, the positive effect of increasing n due to public knowledge spillovers more than offsets the negative effect of a decline in $g(\tilde{m}^{d*})$.

5 Efficiency

Next we consider the welfare properties of the equilibrium path. Clearly, it is first necessary to introduce a concept of constrained efficiency that accounts for the nature of the monopolistic competition environment in the market for consumption commodities. There is a market failure associated with this

¹⁹Based on macro equilibrium conditions of the economy, we have derived the relation (20), meaning that the total expenditure per unit of time is a constant independent of time. In Technical Appendix b, using individual budget constraints, we show that the relation (20) indeed holds along the equilibrium path.

feature of the model in itself. However, that is not the focus of our work. Therefore, we employ a notion of constrained efficiency that allows a planner to search for Pareto improvements by using only the choice of the time path of partnerships in the R & D sector, with perfect foresight of the consequences for the other sectors of the model; in particular, the consumption good market features monopolistic competition once the time path in the R & D sector is chosen.

Here we discuss efficiency in the context of an intertemporal utilitarian social welfare function. We consider the following planner's problem, where the planner chooses $\{\delta_{ij}(\cdot)\}_{i,j=1}^N$ in order to maximize the sum of M-workers' and K-workers' utility given by (23).

To be explicit, the planner's problem is given by:

$$\max_{\{\delta_{ij}(\cdot)\}_{i,j=1}^N} W = \frac{L}{\sigma - 1} \int_0^\infty e^{-\gamma t} \cdot \ln(M(t)) dt$$

subject to the dynamics of \dot{M} , \dot{n}_i , and \dot{m}_{ij}^d , given respectively by

$$\begin{aligned} \dot{M} &= \eta \cdot A \\ &= \eta \cdot \sum_{k=1}^N n_k \left(\delta_{kk} \cdot \alpha + \sum_{l \neq k} \delta_{kl} \cdot G(m_{kl}^d, m_{lk}^d) \right), \end{aligned}$$

(43) and (44).

Suppose that the following symmetric initial conditions for case (i-a) are satisfied:

$$\begin{aligned} n_i(0) &= n(0) > 0 \text{ for } i = 1, \dots, N \\ m^J < m_{ij}^d(0) &= m^d(0) < m^B \text{ for } i, j = 1, \dots, N, i \neq j \\ &\text{and } g(m^B) > \alpha \end{aligned}$$

Given the same, symmetric initial conditions, we ask: *How efficient is the myopic core path?*²⁰

We shall divide our discussion into short run efficiency, long run efficiency, and medium run efficiency. Notice first that our equilibrium path is in the myopic core and eventually reaches the bliss point. There are implications

²⁰Our initial attempts to attack the question of efficiency of the equilibrium path were naturally based on optimal control methods using Hamiltonians. However, given that all the dynamics are linear in the control variables δ_{ij} , the optimization problem reduces to a singular control programming problem; we found no existing theorem applicable to our context. Thus, we adopt an alternative strategy.

for both short and long run efficiency. First, regarding short term efficiency, deviation from the myopic core path creates an immediate loss in utility for all knowledge workers. This means that the equilibrium path cannot be dominated by another path over all time periods for all workers. Second, the equilibrium path eventually reaches the bliss point. Since the bliss point has the highest possible productivity for knowledge workers, the implication is that the equilibrium or myopic core path weakly overtakes any other path; by this we mean that there exists an initial time interval after which the equilibrium path is at least as good as any other path.

The intermediate run is much more difficult to handle. Even in the intermediate run, our myopic core path performs well in terms of efficiency when the population of \mathbb{K} -workers N is sufficiently large. The intuition is as follows.

When $m^d(t) < m^B$, then knowledge productivity is higher and m_{ij}^d moves almost as fast to the right as working in isolation if each person works with every other person with equal intensity. The intuition for this result follows from a combination of two reasons. First, productivity is higher when working with others as opposed to working alone on this part of the path. Second, when N is sufficiently large, working with others is very close to working in isolation when the accumulation of differential knowledge is considered, so cooperation with others will be better on net. Once the bliss point is attained, the system reaches the highest productivity possible, and remains there. This intuition indicates that, when $m^d(t) < m^B$, working with a smaller group than the other $N - 1$ \mathbb{K} -workers results in movement to the right that is slower than working with everyone but oneself.

Formally, first we must introduce two additional symmetry restrictions on paths, aside from the first restriction to symmetric initial conditions, to provide analytical tractability. We also introduce notation for welfare along the myopic core path when the number of \mathbb{K} -workers is given. Then we introduce an upper limit on the welfare associated with arbitrary feasible paths subject to the additional symmetry conditions when the number of \mathbb{K} -workers becomes large. Finally we show that when N is sufficiently large, the welfare associated with the myopic core path approaches the upper limit. Complete details of the analysis are found in Technical Appendix c.

We now introduce two additional symmetry conditions. In our analysis of efficiency, we restrict attention to pairwise symmetric paths. If the initial state is pairwise symmetric, we have shown that the equilibrium path, also called the *myopic core* path, is pairwise symmetric. However, in our effi-

ciency analysis, we impose this assumption for tractability and for the sake of comparison. Thus, we restrict our efficiency analysis to paths satisfying two additional symmetry restrictions: First, for each knowledge worker i and for every knowledge worker $j \neq i$,

$$n_i(t) = n_j(t) \equiv n(t) \text{ for all } t$$

As explained in section 3.1, this is equivalent to

$$m_{ij}^d(t) = m_{ji}^d(t) \text{ for all } t$$

Second, at each moment of time, each person interacts with the same number of people with the same intensity. An implication is that at each time, each \mathbb{K} -worker spends a certain fraction of time working in isolation.

Define m_{mc}^d to be the path of the myopic core over time with a given number N of knowledge workers. It satisfies the additional symmetry conditions. We define m_∞^d to be the limit of this path as the number of knowledge workers tends to infinity, called the *limiting myopic core path*. We define m_f^d to be any feasible path satisfying the two symmetry conditions.

For a fixed population of \mathbb{K} -workers N , we define the levels of welfare generated by three different paths. Let $W_\infty(N)$ be the level of welfare generated by the m_∞^d path, but for knowledge worker population size N .²¹ Typically, this path is not feasible, but it is a useful construction. Let $W_f(N)$ be the level of welfare generated by the m_f^d path for the same population size N . As shown in Lemma 4 of Technical Appendix c,

$$W_\infty(N) > W_f(N)$$

implying that *for each fixed population of \mathbb{K} -workers N , the limiting myopic core growth pattern gives an upper bound on the welfare generated by any feasible growth path satisfying the symmetry conditions*. The intuition for this result is as follows. Since \mathbb{K} -productivity is higher as m^d moves from the left toward the bliss point m^B , and since Lemmas 2 and 3 of Technical Appendix c imply that the m_∞^d path reaches m^B faster than any other feasible path, the welfare associated with the m_∞^d path is naturally greater than the welfare associated with any other feasible path.

Now that we have an upper bound on the welfare generated by an arbitrary path with N knowledge workers, we can examine how the myopic core or

²¹By definition, the m_∞^d path itself is independent of N . However, as shown in Technical Appendix c, the welfare generated by m_∞^d under a given value of population N is a function of N since the volume of new ideas created in the economy at each time is dependent on N .

equilibrium path performs relative to these alternative paths. Let $W_{mc}(N)$ be the level of welfare generated by the myopic core path with N knowledge workers.

Proposition 3: *As the number of knowledge workers N tends to infinity, the difference in welfare corresponding to the myopic core path and the welfare corresponding to the limiting myopic core path monotonically converges to 0:*

$$\lim_{N \rightarrow \infty} \{W_{\infty}(N) - W_{mc}(N)\} = 0$$

The proof can be found in Lemma 5 of Technical Appendix c. The intuition for this surprising result is as follows. By definition, $W_{\infty}(N)$ represents the welfare generated by following the limit of the myopic core pairwise knowledge differential paths as the number of \mathbb{K} -workers tends to infinity, but evaluated in terms of knowledge production when the number of knowledge workers is finite. This path is an upper bound on all feasible paths, but is not itself feasible for finite populations of \mathbb{K} -workers. In contrast, $W_{mc}(N)$ is the actual welfare generated by the myopic core path with N knowledge workers. Although $\lim_{N \rightarrow \infty} W_{\infty}(N)$ and $\lim_{N \rightarrow \infty} W_{mc}(N)$ are generally infinite, their difference converges to zero for the following reason. The actual paths of pairwise knowledge differentials for the finite and infinite population cases converge, so the only difference in welfare is due to the difference in how population affects knowledge production along these given paths. The effect is proportional to knowledge worker population in either case, so the logarithmic form of welfare implied by the model means that this difference is actually zero for all N .

We conclude that for large populations of knowledge workers N , our equilibrium or myopic core path performs well with respect to welfare in comparison with arbitrary paths satisfying the symmetry conditions.

6 Conclusions

We have considered a model of knowledge creation and economic growth that is based on individual behavior, allowing knowledge workers to decide whether joint or individual production is best for them at any given time. Our contributions are as follows. Long run economic growth is positively related to both the effectiveness of pairwise knowledge worker interaction and, more importantly, to the effectiveness of public knowledge transmission. The latter is due, in part, to the endogenous adjustment of R & D group size to a better public knowledge transmission technology. Our equilibrium performs well

in efficiency tests, subject to the constraint that the market for consumption goods features monopolistic competition.

It would be interesting to examine the implications of our model for the relationship between knowledge diversity and economic growth in the real world. The role of immigration policy and of the educational systems in various countries would be a topic worthy of further exploration. For example, comparing the US and Japan, the education system in Japan seems to imply more homogeneity whereas immigration policy in the US is less restrictive. Moreover, workers in the US tend to be more mobile both in terms of job and region than in Japan. It would be interesting to explore the relationship between mobility, knowledge diversity and growth.

Institutions play a key role in this discussion. For example, in the US it is uncommon for a new Ph.D. to stay at their graduate school as an assistant professor. Typically, they move to another school, to the private sector, or to government. Such an informal rule benefits the students, as they are exposed to new approaches and ideas, but also benefits society as a whole, since the boundaries between research groups are not as rigid as described above. Thus, the informal rule allows the flexibility to respond to shocks. Similarly, a generous immigration policy for R & D workers allows a flexible response to shocks in the form of new migrants entering groups.

Indeed, Walsh and Nagaoka (2009) provide empirical evidence to support our discussion. They compare surveys of inventors in Japan and the US.²² They find that 25% of inventors in the US moved between institutions during the last 5 years, whereas only 5% of inventors in Japan moved between institutions over the past 5 years. Furthermore, nearly 30% of the inventors in the US are foreign born, whereas this percentage is negligible in Japan.

Many extensions of our work come to mind. Among others, a second manufacturing sector could be introduced, requiring a different type of R & D activity, for example with a different weight θ on homogeneity of knowledge. This addition to the model would imply feedback from the manufacturing sectors to the R & D sector, in the following sense. Each R & D worker would have to choose a sector in which to specialize. This choice is contingent on the time path of patent prices in the two sectors, resulting in the aforementioned feedback.

Location seems to be an important feature of knowledge creation and trans-

²²To be precise, their data come from inventors of patents that are granted in the US, Japan, and the European Patent Office.

fer, so regions and migration are important, along with urban economic concepts more generally; for example, see Duranton and Puga (2001) and Helsley and Strange (2004). A natural extension of our model would have knowledge workers in regions, either allowing those in the different regions to interact, or making migration of knowledge workers between regions feasible, or both.

It would be interesting to put more structure on our concept of knowledge, allowing asymmetry or introducing notions of distance, such as a metric, on the set of ideas²³ or on the space of knowledge. Finally, it would be useful to add vertical differentiation of knowledge, as in Jovanovic and Rob (1989), to our model of horizontally differentiated knowledge.

REFERENCES

Barro, R.J., Sala-i-Martin, X., 1995. *Economic Growth*. McGraw-Hill, New York.

Berliant, M., Fujita, M., 2008. Knowledge creation as a square dance on the Hilbert cube. *International Economic Review* 49, 1251-1295.

Berliant, M., Reed, R., Wang, P., 2006. Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics* 60, 69-95.

Chatman, J.A., Polzer, J.T., Barsade, S.G., Neale, M.A., 1998. Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Administrative Science Quarterly* 43, 749-780.

Dinopoulos, E., Thompson, P., 1998. Schumpeterian growth without scale effects. *Journal of Economic Growth* 3, 313-335.

Dixit, A., Stiglitz, J., 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67, 297-308.

Duranton, G., Puga, D., 2001. Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review* 91, 1454-1477.

Ghiglino, C., 2005. *Balanced Growth with a Network of Ideas*. Queen Mary, University of London, Working Paper 546.

Grossman, G., Helpman, E., 1991. *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, MA.

Guzzo, R.A., Dickson, M.W., 1996. Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology* 47, 307-338.

²³See Berliant et al. (2006).

- Helsley, R.W., Strange, W.C., 2004. Knowledge barter in cities. *Journal of Urban Economics* 56, 327-345.
- Jones, B., 2005. The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder?. NBER Working Paper 11360.
- Jones, C., 1999. Growth: With or without scale effects?. *American Economic Association Papers and Proceedings* 89, 139-144.
- Jones, L., Manuelli, R., 1990. A convex model of equilibrium growth: Theory and policy implications. *Journal of Political Economy* 98, 1008–1038.
- Jovanovic, B., Rob, R., 1989. The growth and diffusion of knowledge. *The Review of Economic Studies* 56, 569-582.
- Lucas, R. E., Jr., 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22, 2–42.
- Olsson, O., 2000. Knowledge as a set in idea space: An epistemological view on growth. *Journal of Economic Growth* 5, 253-275.
- Olsson, O., 2005. Technological opportunity and growth. *Journal of Economic Growth* 10, 35-57.
- Peretto, P., Smulders, S., 2002. Technological distance, growth, and scale effects. *The Economic Journal* 112, 603-624.
- Romer, P., 1986. Increasing returns and long-run growth. *Journal of Political Economy* 94, 1002–1037.
- Romer, P., 1990. Endogenous technological change. *Journal of Political Economy* 98, S71–S102.
- Shell, K., 1966. Toward a theory of inventive activity and capital accumulation. *American Economic Review* 61, 62–68.
- Trajtenberg, M., Shiff, G., Melamed, R., 2006. The “Names Game”: Harnessing Inventors’ Patent Data for Economic Research. NBER Working Paper 12479.
- Walsh, J.P., Nagaoka, S., 2009, Who invents? Evidence from the Japan-US inventor survey, RIETI Discussion Paper Series 09-E-034.
- Weitzman, M.L., 1998. Recombinant growth. *The Quarterly Journal of Economics* 113, 331-360.
- Wuchty, S., Jones, B., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. *Science* 316, 1036-1039.
- Young, A., 1998. Growth without scale effects. *Journal of Political Economy* 106, 41-63.

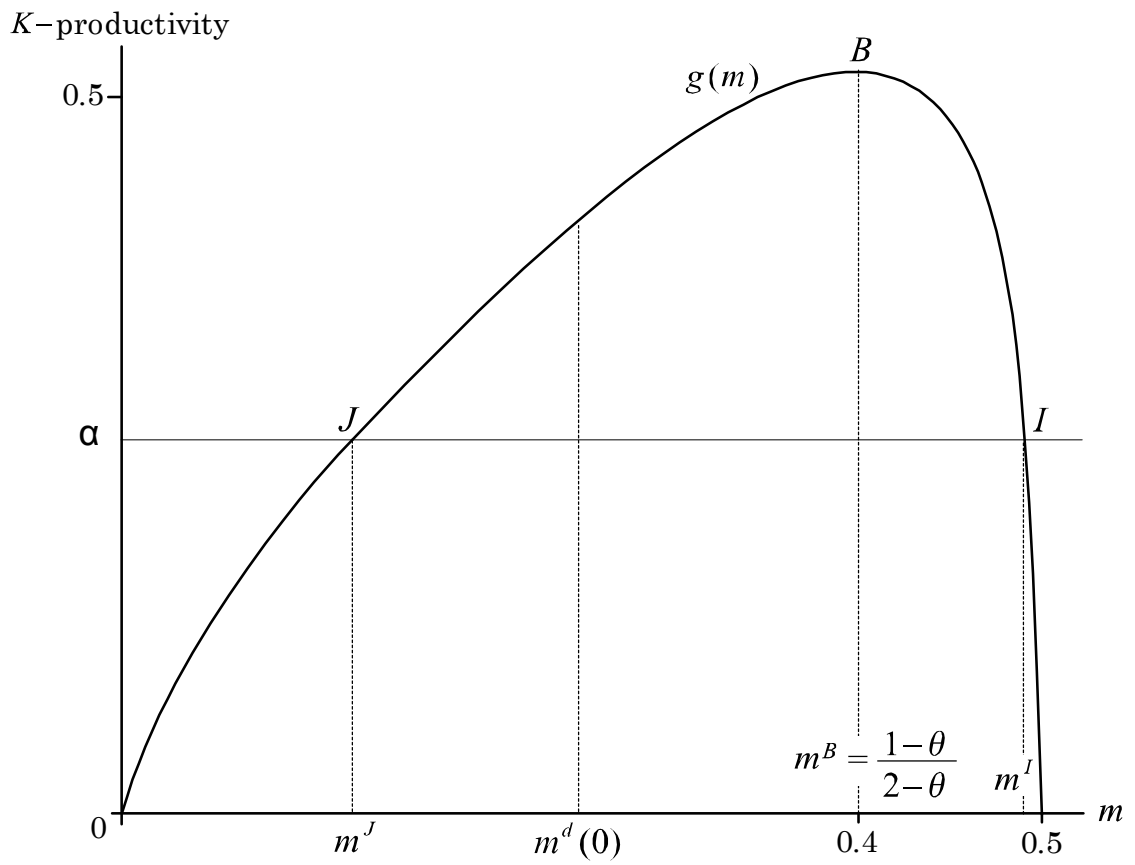


Figure 1: The $g(m)$ curve and the bliss point when $\beta = 1$ and $\theta = 1/3$.

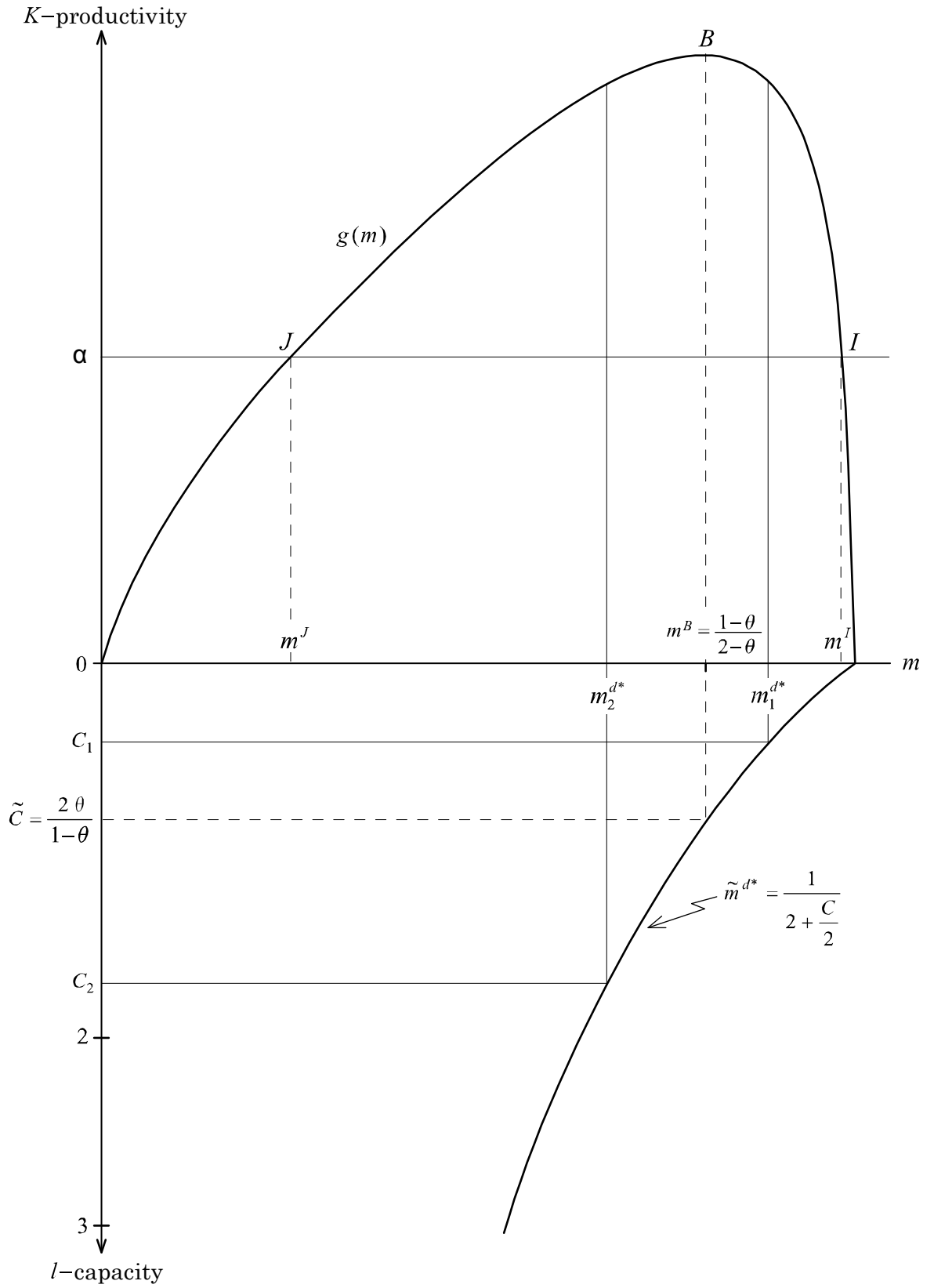


Figure 2: The \mathbb{K} -productivity curve $g(m)$ and the sink curve \tilde{m}^{d*} .

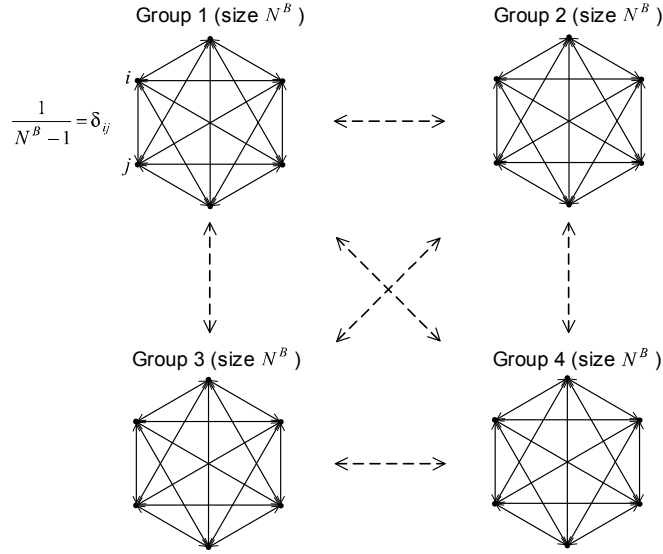


Figure 3: An example of \mathbb{K} -interactions at the bliss point.

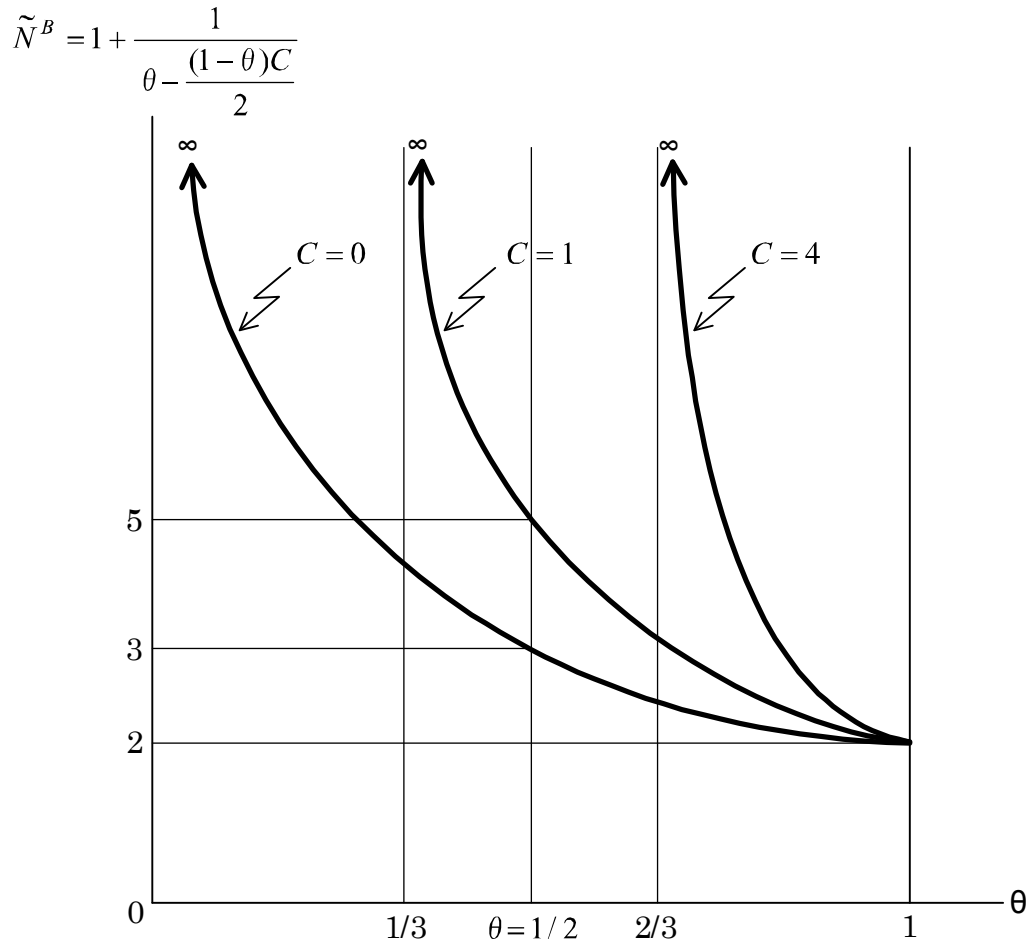


Figure 4: The optimal group size \tilde{N}^B .

The Dynamics of Knowledge Diversity and Economic Growth

Marcus Berliant and Masahisa Fujita

Appendix

Appendix 1: Explanations and Proofs

Appendix 1.1 for section 3.2: Justification of Knowledge Absorption Function

Consider the following statement of the capacity constraint on knowledge absorption from public information on \mathbb{K} -worker i :

$$C \cdot \left(\delta_{ii} \cdot a_{ii} + \sum_{j \neq i} \delta_{ij} \cdot (a_{ij}/2) \right) = \mu \eta \left(A - \sum_{j=1}^N \delta_{ij} \cdot a_{ij} \right) \quad (68)$$

We shall explain the content of this equation piece by piece. On the right hand side of the equation, the term in brackets $A - \sum_{j=1}^N \delta_{ij} \cdot a_{ij}$ represents the new knowledge produced in the economy that does not involve partnerships including \mathbb{K} -worker i . Recall that η gives the rate at which new ideas are patented, whereas μ gives the rate at which publicly revealed ideas can be absorbed by a \mathbb{K} -worker. Therefore the right hand side of the equation represents the public knowledge revealed by patents that is absorbed by \mathbb{K} -worker i . The term in brackets on the left hand side represents new knowledge created by \mathbb{K} -worker i at an instant. In total, the equation means that the new public knowledge that can be absorbed by \mathbb{K} -worker i is proportional to their capacity to produce new ideas. In essence, this is due to the constraint on their time and the productivity of their effort both to absorb new ideas and to produce them.

Equation (68) implies:

$$\mu = \frac{C}{\eta} \cdot \frac{\delta_{ii} \cdot a_{ii} + \sum_{j \neq i} \delta_{ij} \cdot (a_{ij}/2)}{A - \sum_{j=1}^N \delta_{ij} \cdot a_{ij}}$$

Next we consider two special cases, where we assume pairwise symmetry: $n_k = n_i \equiv n$ for all i and k . First, when each agent in the knowledge sector is working alone, namely $\delta_{ii} = 1$ for all i and $\delta_{ij} = 0$ for all $i \neq j$, then

$$A = \sum_{k=1}^N \alpha \cdot n_k = \alpha \cdot n \cdot N$$

and

$$\mu = \frac{C}{\eta} \cdot \frac{\alpha \cdot n}{\alpha \cdot n \cdot (N - 1)} = \frac{C}{\eta} \cdot \frac{1}{N - 1}$$

The second special case is given by $m_{ij}^d = m^d$ for all $i \neq j$ and $g(m^d) > \alpha$. Thus, $\delta_{ii} = 0$ for all i and $a_{ij} = a$ for all $i \neq j$. In this special case, we have

$$\mu = \frac{C}{\eta} \cdot \frac{a/2}{\frac{Na}{2} - a} = \frac{C}{\eta} \cdot \frac{1}{N - 2}$$

Assuming N is sufficiently large, we employ the following specification.

$$\mu(N) \approx \frac{C}{\eta N}$$

6.1 Appendix 1.2 for section 3.2.1: Explanation of Case

(i): $m^J < m^d(0) \leq m^B$

First suppose that the initial state is such that

$$m^J < m^d(0) < m^B$$

Then, since $g(m_{ij}^d(0)) = g(m^d(0)) > \alpha$ for any possible work pairs consisting of i and j , no person wishes to work alone at the start. However, since the value of $g(m_{ij}^d(0))$ is the same for all possible pairs, any feasible pairings are possible equilibrium work configurations at the start. To determine which one of them will actually take place on the equilibrium path, we must consider the first derivative of income for all persons.

In general, consider any time at which all persons have the same composition of knowledge:

$$m_{ij}^d = m^d \text{ for all } i \neq j \quad (69)$$

where

$$g(m^d) > \alpha$$

Focus on person i ; the equations for other persons are analogous. Since person i does not wish to work alone, it follows that

$$\delta_{ii} = 0 \text{ and } \sum_{j \neq i} \delta_{ij} = 1 \quad (70)$$

Substituting (69) and (70) into (27) and using (52) yields

$$y_i = \Pi \cdot \eta \cdot n_i \cdot g(m^d) \quad (71)$$

Likewise, substituting (58), (69) and (70) into (55) and arranging terms gives

$$\dot{m}_{ij}^d = \dot{m}^d = 2(1 - m^d) \cdot g(m^d) \cdot \left\{ \left(1 - \frac{C}{N}\right) \cdot (1 - 2m^d) - \left(1 - \frac{C}{N}\right) \cdot (1 - m^d) \cdot \delta_{ij} - \frac{C}{2} \cdot m^d \right\} \quad (72)$$

for $i \neq j$.

Since the income function (71) is independent of the values of δ_{ij} ($j \neq i$), in order to examine what values of δ_{ij} ($j \neq i$) person i wishes to choose, we must consider the time derivative of y_i . In doing so, however, we cannot use

equation (71) because the original variables have been replaced. Instead, we must go back to the original equation (36). Then, using equations (69) to (72) and setting $\delta_{ij} = \delta_{ji}$ (which must hold for any feasible meeting), we obtain the following (see Berliant and Fujita, 2008, Technical Appendix b for proof):

$$\begin{aligned} \dot{y}_i &= \dot{\Pi} \cdot \eta \cdot n_i \cdot g(m^d) + \Pi \cdot \eta \cdot \dot{n}_i \cdot g(m^d) \\ &\quad + \Pi \cdot \eta \cdot n_i \cdot \sum_{j \neq i} \delta_{ij} \cdot g'(m^d) \cdot \dot{m}_{ij}^d \end{aligned} \quad (73)$$

where

$$\dot{n}_i = g(m^d) \cdot n_i \cdot \left(2 + \frac{N-2}{N} \cdot C\right) \quad (74)$$

and \dot{m}_{ij}^d is given by (72). Substituting (72) into (73) and setting $\sum_{j \neq i} \delta_{ij} = 1$ yields

$$\begin{aligned} \dot{y}_i &= \dot{\Pi} \cdot \eta \cdot n_i \cdot g(m^d) + \Pi \cdot \eta \cdot \dot{n}_i \cdot g(m^d) \\ &\quad + \Pi \cdot \eta \cdot n_i \cdot 2 \cdot (1 - m^d) \cdot g(m^d) \cdot g'(m^d) \cdot \\ &\quad \left\{ \left(1 - \frac{C}{N}\right) \cdot (1 - 2m^d) - \left(1 - \frac{C}{N}\right) \cdot (1 - m^d) \cdot \sum_{j \neq i} \delta_{ij}^2 - \frac{C}{2} \cdot m^d \right\} \end{aligned} \quad (75)$$

All \mathbb{K} -workers take Π and $\dot{\Pi}$ as given, whereas n_i is a state variable. Furthermore, the value of \dot{n}_i given above is independent of the values of δ_{ij} for $j \neq i$. Thus, choosing the values of δ_{ij} for $j \neq i$ is equivalent to choosing the values that maximize the last term in (75).

Now, suppose that

$$m^d < m^B$$

and hence $g'(m^d) > 0$. Then, assuming that $\frac{C}{N} < 1$, in order to maximize the time derivative of the income, person i must solve the following quadratic minimization problem:

$$\min \sum_{j \neq i} \delta_{ij}^2 \quad \text{subject to} \quad \sum_{j \neq i} \delta_{ij} = 1 \quad (76)$$

which yields the solution for person i :

$$\delta_{ij} = \frac{1}{N-1} \quad \text{for all } j \neq i \quad (77)$$

Although we have focused on person i , the vector of optimal strategies is the same for all persons. Thus, all persons agree to a knowledge creation pattern in which each person rotates through all $N-1$ possible partners, sharing time equally.

The intuition behind this result is as follows. The condition $m^d < m^B$ means that the \mathbb{K} -workers have relatively too many ideas in common, and thus they wish to acquire ideas that are different from those of each possible partner as fast as possible. That is, when $m^J < m_{ij}^d = m^d < m^B$ in Figure 1, each \mathbb{K} -worker wishes to move the knowledge composition m_{ij}^d to the right as quickly as possible, thus increasing the \mathbb{K} -productivity $g(m_{ij}^d)$ as fast as possible.

Concerning the general case with $N \geq 4$, when $m^J < m^d(0) = m_{ji}^d(0) < m^B$ for all $i \neq j$, on the equilibrium path, each \mathbb{K} -worker i spends the same amount of time $\delta_{ij} = 1/(N-1)$ for all $j \neq i$ with every other \mathbb{K} -worker at the start. Then, since the symmetric condition (69) holds from the start onward, the same work pattern will continue as long as $m^J < m^d < m^B$. The dynamics of this work pattern are as follows. The creation of new ideas always takes place in pairs. Pairs are cycling rapidly with $\delta_{ij} = 1/(N-1)$ for all $j \neq i$. \mathbb{K} -worker 1, for example, spends $1/(N-1)$ of each period with \mathbb{K} -worker 2, for example, and $(N-2)/(N-1)$ of the time working with other partners. Setting $m_{ij}^d = m^d$ and $\delta_{ij} = 1/(N-1)$ in (72), we obtain

$$\dot{m}^d = 2(1 - m^d) \cdot g(m^d) \cdot \frac{1 - \frac{C}{N}}{N - 1} \left\{ (N - 2) - m^d \left[(2N - 3) + \frac{C}{2} \cdot \frac{N - 1}{1 - \frac{C}{N}} \right] \right\} \quad (78)$$

Setting $\dot{m}^d = 0$ and considering that $m^d < 1$, we obtain the sink point

$$m^{d*} = \frac{N - 2}{(2N - 3) + \frac{C}{2} \cdot \frac{N - 1}{1 - \frac{C}{N}}} \quad (79)$$

As N increases, the value of m^{d*} increases monotonically (provided $N > C$) eventually reaching the limit

$$\tilde{m}^{d*} = \frac{1}{2 + \frac{C}{2}} \quad (80)$$

In the upper half of Figure 2, the \mathbb{K} -productivity curve $g(m)$ is transferred from Figure 1. In the bottom half of Figure 2, the bold curve depicts the limiting sink, \tilde{m}^{d*} , as a function of the l -capacity parameter C . When N is sufficiently large, the actual sink curve, m^{d*} , is close to this limiting curve.

FIGURE 2 GOES HERE

In the context of Figure 2, we can identify two different possibilities. Suppose that

$$m^B < m^{d*} \quad (81)$$

That is, the sink point of the dynamics given in (78) is on the right side of the bliss point. In this case, beginning at any point $m^J < m^d(0) < m^B$, the system reaches the bliss point in finite time. In terms of the original parameters, using (54) and (79), condition (81) can be rewritten as

$$C < \frac{\left\{ \frac{2-\theta}{1-\theta} - \frac{2N-3}{N-2} \right\} \cdot (N-1)}{\frac{2-\theta}{1-\theta} - \frac{2N-3}{N-2} + \frac{N}{2}} \quad (82)$$

Since $m^{d*} \rightarrow \tilde{m}^{d*}$, when N is sufficiently large, condition (82) can be expressed as

$$C < \tilde{C} \equiv \frac{2\theta}{1-\theta} \quad (83)$$

In Figure 2, C_1 provides an example of this case. The associated sink point is given by m_1^{d*} .

In contrast, suppose that

$$m^{d*} < m^B \quad (84)$$

This occurs exactly when the inequality in (82) is reversed. Assuming that N is sufficiently large, it occurs when the inequality in (83) is reversed. In Figure 2, C_2 represents an example of such a value of C , whereas the associated sink point is given by m_2^{d*} . In this case, starting with any initial point $m^J < m^d(0) < m^B$, the system moves automatically toward $m^{d*} < m^B$, but never reaches the bliss point.

On the downward vertical axis of Figure 2, \tilde{C} gives the value of the parameter C at the boundary of the two cases. Case (81) occurs exactly when the value of the l -capacity C is relatively small, whereas case (84) occurs when C is relatively large. In what follows, under the assumption that N is large, we examine the actual dynamics in each of the two cases.

6.1.1 Case (i-a): $m^J < m^d(0) \leq m^B$ and $C < \tilde{C}$

When condition (81) holds, starting with any initial point $m^J < m^d(0) \leq m^B$, the system following the dynamics (72) reaches the bliss point m^B in finite time. Let t^B be the time at which all \mathbb{K} -workers reach the bliss point. At time t^B , we have

$$m_{ij}^d = m^d = m^B \text{ for } i \neq j \quad (85)$$

and $g'(m^d) = g'(m^B) = 0$. Thus, (75) becomes

$$\dot{y}_i = \dot{\Pi} \cdot \eta \cdot n_i \cdot g(m^d) + \Pi \cdot \eta \cdot \dot{n}_i \cdot g(m^d) \quad (86)$$

that is, again, independent of the values of δ_{ij} ($j \neq i$). Thus, we consider the second order condition for income maximization. Replace $g(m^d)$ with

$G(m_{ij}^d, m_{ji}^d)$ in (73) and take the time derivative of the resulting equation. Using (85) and the fact that $g'(m^B) = 0$, by following the logic in Berliant and Fujita (2008, Technical Appendix b) we obtain

$$\begin{aligned} \ddot{y}_i &= \ddot{\Pi} \cdot \eta \cdot n_i \cdot g(m^B) + 2\dot{\Pi} \cdot \eta \cdot \dot{n}_i \cdot g(m^B) + \Pi \cdot \eta \cdot \ddot{n}_i \cdot g(m^B) \\ &\quad + \Pi \cdot \eta \cdot n_i \cdot (1 - m^B)^2 \cdot 4(m^B)^2 g''(m^B) \cdot \\ &\quad \left\{ \sum_{i \neq j} \delta_{ij} \cdot \left[\left(1 - \frac{C}{N}\right) \cdot (1 - 2m^B) - \left(1 - \frac{C}{N}\right) \cdot (1 - m^B) \cdot \delta_{ij} - \frac{C}{2} m^B \right] \right\}^2 \end{aligned} \quad (87)$$

where, using (74),

$$\begin{aligned} \dot{n}_i &= g(m^B) \cdot n_i \cdot \left(2 + \frac{N-2}{N} C\right) \\ \ddot{n}_i &= g(m^B) \cdot \dot{n}_i \cdot \left(2 + \frac{N-2}{N} C\right) \end{aligned}$$

Since the first three terms on the right hand side of (87) are independent of the values of δ_{ij} ($j \neq i$) whereas $g'' < 0$, choosing the values of δ_{ij} ($j \neq i$) to maximize \ddot{y}_i is equivalent to the following optimization problem:

$$\min_{\{\delta_{ij}\}} \left\{ \sum_{i \neq j} \delta_{ij} \cdot \left[\left(1 - \frac{C}{N}\right) \cdot (1 - 2m^B) - \left(1 - \frac{C}{N}\right) \cdot (1 - m^B) \cdot \delta_{ij} - \frac{C}{2} m^B \right] \right\}^2 \quad (88)$$

$$\text{subject to } \sum_{j \neq i} \delta_{ij} = 1$$

This problem can be solved by using the rule that whenever $\delta_{ij} > 0$, the value of the terms inside the square brackets in expression (88) must be zero, or

$$\delta_{ij} > 0 \implies \delta_{ij} = \frac{\left(1 - \frac{C}{N}\right) \cdot (1 - 2m^B) - \frac{C}{2} m^B}{\left(1 - \frac{C}{N}\right) \cdot (1 - m^B)} \equiv \delta^B \quad (89)$$

whereas the number of partners for \mathbb{K} -worker i must be chosen to satisfy the constraint $\sum_{j \neq i} \delta_{ij} = 1$. This applies to all \mathbb{K} -workers.

This equilibrium configuration of partnerships at the bliss point m^B can be achieved as follows: When the system reaches m^B , the population splits into smaller groups of equal size,²⁴

$$N^B \equiv 1 + \frac{1}{\delta^B} \quad (90)$$

²⁴The configuration of workers necessary to maintain the bliss point is not unique. Each \mathbb{K} -worker must have $N^B - 1$ links to other \mathbb{K} -workers, communicating with each for an equal share of time. For example, when $N^B = 4$, groups of 4 may form, where each worker within a group communicates equally with every other worker in that group. However,

so each person works with $N^B - 1$ other persons in their group for the same proportion of time, δ^B . Recalling (72), rule (89) is equivalent to

$$\delta_{ij} > 0 \implies \dot{m}_{ij}^d = 0 \text{ at } m_{ij}^d = m^B$$

That is, when all \mathbb{K} -workers reach the bliss point, they stay there by splitting into smaller groups of the same size, N^B , so direct interactions take place only within each group. In this way, each \mathbb{K} -worker maintains the highest \mathbb{K} -productivity while enjoying the knowledge externalities derived from public information on new patents. Figure 3 depicts an example of an equilibrium configuration of \mathbb{K} -worker interactions in which four groups of \mathbb{K} -workers form at the bliss point. The dotted arrows represent indirect interactions through the public revelation of patent information.

FIGURE 3 GOES HERE

Substituting (54) into (89), using (90) and arranging terms, the optimal group size N^B is given by

$$N^B = 1 + \frac{1}{\theta - \frac{(1-\theta) \cdot C}{2} \cdot \frac{N}{N-C}} \quad (91)$$

As N becomes large, the optimal group size approaches

$$\tilde{N}^B = 1 + \frac{1}{\theta - \frac{(1-\theta) \cdot C}{2}} \quad (92)$$

The optimal group size for large population \tilde{N}^B (as well as the optimal group size for finite population N^B) increases monotonically with the l -capacity, C ; as C increases, the transmission of public knowledge in common increases, so it is necessary to have a larger group in order to maintain heterogeneity among agents within the group. Recalling that \tilde{C} was defined in (83), the group size becomes infinitely large as C approaches \tilde{C} from the left. Recalling that θ is the weight given to knowledge in common in the \mathbb{K} -production function, as

with $N^B = 4$ it is also possible to have, say, groups of six forming. With such groups, each \mathbb{K} -worker has communication links to only three other \mathbb{K} -workers within their group. So not all possible links within a group are actually active. If groups at the bliss point are larger, then their communication structure must become more sparse to maintain the bliss point. The minimal size of groups that coalesce at the bliss point is clearly N^B . Nevertheless, all of the calculations apply independent of the size of groups that form at the bliss point. The same remarks apply to the various cases detailed below, except when \mathbb{K} -workers are in isolation.

the value of θ increases, \tilde{N}^B decreases, which is not surprising. In Figure 4, for each fixed value of the parameter C , the optimal group size \tilde{N}^B is graphed as a function of θ .

FIGURE 4 GOES HERE

Substituting $\delta^B = 1/(N^B - 1)$ for δ_{ij} in equation (72), then by construction, m^B is the sink point of the dynamics

$$\dot{m}_{ij}^d = 2(1 - m^d) \cdot g(m^d) \cdot \left\{ \left(1 - \frac{C}{N}\right) \cdot (1 - 2m^d) - \left(1 - \frac{C}{N}\right) \cdot (1 - m^d) \cdot \frac{1}{N^B - 1} - \frac{C}{2} \cdot m^d \right\} \quad (93)$$

for $i \neq j$ but i and j in the same group. Thus, starting with any initial point $m_{ij}^d(0) = m^d(0) \in (0, 1/2)$, if each person participates in a group of N^B persons, and if they maintain the same group structure where each person works with each of the $N^B - 1$ other people in their group for the same proportion of time δ^B , then the system monotonically approaches the bliss point m^B . However, when $m_{ij}^d(0) = m^d(0) < m^B$, if all N persons form a single group while setting $\delta_{ij} = 1/(N - 1)$, the system can reach the bliss point m^B fastest.²⁵

When the system reaches the bliss point, the workers break into groups and the system becomes asymmetric, in the following sense. If \mathbb{K} -worker i belongs to the same group as \mathbb{K} -worker k , then their differential knowledge remains at the bliss point m^B , maintaining the highest \mathbb{K} -productivity $g(m^B)$. If \mathbb{K} -worker j belongs to a different group, then the differential knowledge between i and j diverges, namely it moves away from m^B , thus reducing $g(m_{ij}^d)$. So once the population splits into groups, \mathbb{K} -workers i and j will not want to collaborate again.

Formally, setting $\delta_{ij} = 0$ in equation (72), the dynamics of differential knowledge for \mathbb{K} -workers i and j in different groups is given by

$$\dot{m}_{ij}^d = \dot{m}^d = 2(1 - m^d) \cdot g(m^d) \cdot \left\{ \left(1 - \frac{C}{N}\right) \cdot (1 - 2m^d) - \frac{C}{2} \cdot m^d \right\} \quad (94)$$

that yields a sink point

$$m^{d*} = \frac{1}{2 + \frac{C}{2(1 - \frac{C}{N})}}$$

As $N \rightarrow \infty$, the sink point becomes

$$m^{d*} = \frac{1}{2 + \frac{C}{2}} = \tilde{m}^{d*} \quad (95)$$

²⁵With a starting point $m_{ij}^d(0) = m^d(0) < m^B$, if the population forms groups of size less than N but larger than N^B , then the system will still reach the bliss point, but at a slower speed than if the group size were N .

Notice that this is the same as expression (80). As the number of \mathbb{K} -workers becomes large, the difference between pairs of workers who interact at intensity $1/(N-1)$ and pairs of workers in different groups who don't interact is close to zero, so they tend to the same sink point.

To sum up, for partnerships of \mathbb{K} -workers within the same group, their productivity is $g(m^B)$. For potential partnerships of \mathbb{K} -workers in different groups, their potential productivity is $g(\tilde{m}^{d*}) < g(m^B)$. So these potential partnerships are never formed.

The implication is that we have endogenous formation of cohesive groups. One interpretation of this phenomenon is that the groups represent research firms, so we have endogenous formation of research firm boundaries.

6.1.2 Case (i-b): $m^J < m^d(0) \leq m^B$ and $C > \tilde{C}$

As explained previously, in this case the dynamics imply that only one large group forms, so each agent works with everyone else an equal amount of time. Heterogeneity m^d changes, approaching the sink point m^{d*} given by (79) to the left of the bliss point, so the bliss point is never reached. In this case

$$m^J < m^{d*} < m^B$$

and one large group is maintained forever, without achieving the highest possible productivity. Intuitively, this is due to the large externality from public knowledge, so it is impossible to attain sufficient heterogeneity.²⁶

6.2 Appendix 1.3 for section 3.2.2: Explanation of Case

(ii): $m^d(0) < m^J < m^B$

Under this set of parameters, $g(m^d(0)) < \alpha$. In other words, at time 0 it is best for everyone to work in isolation rather than in pairs. Substituting $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$ into (55), and using (58), we obtain dynamics for work in isolation:

$$\dot{m}_{ij}^d = \dot{m}^d = (1 - m^d) \cdot \alpha \cdot \left\{ \left(1 - \frac{C}{N}\right) \cdot (1 - 2m^d) - C \cdot m^d \right\} \quad (96)$$

that yields the sink point

$$m^{d**} = \frac{1}{2 + \frac{C}{1 - \frac{C}{N}}}$$

²⁶When C is very large, it is possible that $m^{d*} < m^J$, implying that all \mathbb{K} -workers eventually work in isolation. However, this is not an interesting situation, so we neglect it in the discussion that follows.

As $N \rightarrow \infty$, the sink point approaches

$$\tilde{m}^{d^{**}} = \frac{1}{2 + C}$$

Evidently, $\tilde{m}^{d^{**}} < \tilde{m}^{d^*}$. When N is sufficiently large, it follows that

$$m^{d^{**}} < m^{d^*}$$

Focusing on this case, there are two possibilities, namely $m^J < m^{d^{**}}$ and $m^J > m^{d^{**}}$.²⁷ Assuming C is not too large, we concentrate on the first possibility,

$$m^J < m^{d^{**}} \tag{97}$$

The equilibrium path has every \mathbb{K} -worker in isolation to begin, creating new knowledge on their own and moving to the right until they all reach the point m^J . Then one large group forms and all \mathbb{K} -workers create new knowledge working in pairs where each spends equal time with every other. From here, the equilibrium path is exactly the same as in case (i).

Appendix 1.4 for section 3.2.3: Explanation of Case (iii):

$$m^B < m^d(0)$$

Let us consider the dynamics of the system when it begins to the right of m^B . First we consider the situation where $m^B < m^d(0) < m^I$, where m^I was introduced in Figure 1. In other words, the initial state reflects a higher degree of heterogeneity than the bliss point, but $g(m^d(0)) > \alpha$. Since the initial state reflects a higher degree of heterogeneity than the bliss point, the \mathbb{K} -workers want to increase the knowledge they have in common as fast as possible, leading to fidelity and *pairwise knowledge creation*.

To be precise, since $m_{ij}^d(0) = m^d(0)$ for all $i \neq j$ and $g(m^d(0)) > \alpha$, the situation at time 0 is the same as that in Case (i) except that we now have $m^d(0) > m^B$. Hence, focusing on person i as before, the time derivative of income y_i at time 0 is given by (75). However, since $g'(m^d) = g'(m^d(0)) < 0$ at time 0, in order to maximize the right hand side of equation (75), person i now must solve now the following quadratic maximization problem:

$$\max \sum_{j \neq i} \delta_{ij}^2 \quad \text{subject to} \quad \sum_{j \neq i} \delta_{ij} = 1 \tag{98}$$

²⁷Under the second possibility, $m^{d^{**}} < m^J$, each \mathbb{K} -worker creates knowledge in isolation forever, approaching the sink point $m^{d^{**}}$.

Thus, person i wishes to choose any partner, say k , and set $\delta_{ik} = 1$, whereas $\delta_{ij} = 0$ for all $j \neq k$. The situation is the same for all \mathbb{K} -workers. Hence, without loss of generality, we can assume that N persons agree at time 0 to form the following combination of partnerships:

$$P_1 \equiv \{\{1, 2\}, \{3, 4\}, \{5, 6\}, \dots, \{N-1, N\}\} \quad (99)$$

and initiate a pairwise knowledge creation work pattern such that²⁸

$$\delta_{ij} = \delta_{ji} = 1 \text{ for } \{i, j\} \in P_1, \delta_{ij} = \delta_{ji} = 0 \text{ for } \{i, j\} \notin P_1 \quad (100)$$

Similar to Berliant and Fujita (2008, case (ii)), the equilibrium path can be described as follows. The equilibrium path consists of several phases. First, in order to increase income and \mathbb{K} -productivity as fast as possible, the \mathbb{K} -workers want to develop knowledge in common with their partner as fast as possible. Therefore, the N persons are paired arbitrarily and work with their partners for a nonempty interval of time. This implies fast movement to the left, because there is both shared knowledge creation and public knowledge transfer. If potential partners are not actually meeting, their differential knowledge will converge to the sink point of the process where no persons meet, given by (95) and illustrated by m_1^{d*} in Figure 2. This process moves to the left beyond the bliss point because \mathbb{K} -workers cannot switch to any new partner that will allow them to maintain the bliss point. The actual partners move quickly to the left of the bliss point and their \mathbb{K} -productivity decreases rapidly. When their productivity matches that of a potential partner with whom they have not worked, they switch to new partners and work with their new partners for a nonempty interval of time. Once again, the two actual partners increase their knowledge in common quickly, past the bliss point, and their productivity decreases rapidly, while the differential knowledge with their potential partners moves slowly toward m^{d*} , until the productivity of their current partnership and their previous partnership are the same. Next, each person works alternately with the two partners with whom they worked in the first two phases, but not with a person with whom they have not worked previously. This process continues, but the productivity of each \mathbb{K} -worker oscillates between $g(m_1^{d*})$ and $g(m^B)$. The equilibrium path in this case crosses the bliss point, but this is not a sink of the process, due to the myopic behavior of the \mathbb{K} -workers.

²⁸Here we adopt the convention that $\{i, j\} \in P_1$ means either $\{i, j\} \in P_1$ or $\{j, i\} \in P_1$, whereas $\{i, j\} \notin P_1$ means neither $\{i, j\} \in P_1$ nor $\{j, i\} \in P_1$.

Appendix 1.5 for section 4.1: Growth in Case (i-a):

Recall from equation (77) that the initial pattern of knowledge creation has each \mathbb{K} -worker interacting with every other \mathbb{K} -worker for an equal share of time, so the dynamics are symmetric and given by (78). The associated sink point is given by (79). Summarizing, the assumption of case (i-a) implies that the sink point, where each \mathbb{K} -worker is interacting with every other \mathbb{K} -worker with the same intensity, is to the right of the bliss point:

$$m^J < m^d(0) < m^B < m^{d*}$$

Let t^B be the unique finite time such that the dynamics reach the bliss point, so that when i and j belong to the same group:

$$m_{ij}^d(t) = m^d(t) = m^B \text{ for } t \geq t^B.$$

Due to the symmetry of the path in case (i-a), for all t when $\delta_{ij}(t) > 0$,

$$\begin{aligned} g(m_{ij}^d(t)) &= g(m^d(t)) \\ n_i(t) &= n(t) \\ a_{ij}(t) &= a(t) = n(t) \cdot 2g(m^d(t)) \end{aligned}$$

In particular

$$\begin{aligned} g(m^d(t)) &= g(m^B) \text{ for } t \geq t^B \\ a(t) &= n(t) \cdot 2g(m^B) \text{ for } t \geq t^B \end{aligned}$$

Setting $m^d = m^B$ in equation (74), we have

$$\dot{n}(t) = n(t) \cdot \phi^B \text{ for } t \geq t^B$$

where

$$\phi^B \equiv g(m^B)(2 + C)$$

so

$$\frac{\dot{n}(t)}{n(t)} = \phi^B \text{ for } t \geq t^B \tag{101}$$

Thus, once the system reaches the bliss point, the size of each \mathbb{K} -worker's knowledge expands at an exponential rate of ϕ^B , and we have:

$$n(t) = n(t^B) \cdot e^{\phi^B(t-t^B)} \text{ for } t \geq t^B$$

Recall that the number of varieties of manufactured goods at time t , that is equal to the number of patents present at time t , is given by $M(t)$. Since $A(t)$ is the total number of ideas created at time t , whereas the the proportion of new ideas that are patented is given by η , the rate of increase in patents at time t is given by

$$\dot{M}(t) = \eta \cdot A(t) \quad (102)$$

In the present context of case (i-a), using equation (38),

$$A(t) = \frac{N \cdot a(t)}{2} = N \cdot n(t) \cdot g(m^d(t))$$

and hence

$$\dot{M}(t) = \eta \cdot N \cdot n(t) \cdot g(m^d(t))$$

In particular,

$$\dot{M}(t) = \eta \cdot N \cdot g(m^B) \cdot n(t^B) \cdot e^{\phi^B(t-t^B)} \text{ for } t \geq t^B \quad (103)$$

With this in hand, we can proceed to the calculation of the asymptotic rate of growth of patents. First, for $t \geq t^B$:

$$M(t) = M(t^B) + \int_{t^B}^t \dot{M}(\tau) d\tau$$

Using (103),

$$\int_{t^B}^t \dot{M}(\tau) d\tau = \eta \cdot N \cdot g(m^B) \cdot n(t^B) \cdot \frac{e^{\phi^B(t-t^B)} - 1}{\phi^B}$$

Hence

$$\frac{\dot{M}(t)}{M(t)} = \frac{\eta \cdot N \cdot g(m^B) \cdot n(t^B) \cdot e^{\phi^B(t-t^B)}}{M(t^B) + \eta \cdot N \cdot g(m^B) \cdot n(t^B) \cdot \frac{e^{\phi^B(t-t^B)} - 1}{\phi^B}}$$

implying

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\dot{M}(t)}{M(t)} &= \phi^B \equiv g(m^B)(2 + C) \\ &= \beta\theta^\theta(1 - \theta)^{1-\theta} \cdot (2 + C) \end{aligned} \quad (104)$$

Notice that the asymptotic growth rate of M is the same as the asymptotic growth rate of n .

Next we calculate the rate of growth of indirect utility of consumers. Using (4) and (20), for any consumer i , we have:

$$v_i(t) = E_i \cdot \rho M(t)^{1/(\sigma-1)}$$

Since E_i is constant, this leads to

$$\frac{\dot{v}_i(t)}{v_i(t)} = \frac{1}{\sigma - 1} \cdot \frac{\dot{M}(t)}{M(t)}$$

Thus, using (104)

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\dot{v}_i(t)}{v_i(t)} &= \frac{\phi^B}{\sigma - 1} = \frac{g(m^B)(2 + C)}{\sigma - 1} \\ &= \frac{\beta\theta^\theta(1 - \theta)^{1-\theta} \cdot (2 + C)}{\sigma - 1} \end{aligned} \quad (105)$$

Therefore the growth rate of indirect utility approaches a constant.

In summary, the growth rate of the individual stock of knowledge (101), the growth rate of patents (104) and the growth rate of indirect utility (105) approach constants as t tends to infinity. These constants are positively related to both C and β .

Appendix 1.6 for section 4.2: Growth in Case (i-b)

The assumption that applies for this case implies that $\tilde{m}^{d*} < m^B$. This case is very similar to the previous one. The only change in the calculations is that $g(m^B)$ is replaced with $g(\tilde{m}^{d*})$, where \tilde{m}^{d*} is given by equation (80). The system tends to \tilde{m}^{d*} as $t \rightarrow \infty$. Defining

$$\begin{aligned} \phi^* &\equiv g(\tilde{m}^{d*}) \cdot (2 + C) \\ &= 2\beta\left(\frac{C}{2}\right)^\theta \end{aligned}$$

Analogous calculations yield

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\dot{n}(t)}{n(t)} &= \phi^* \\ \lim_{t \rightarrow \infty} \frac{\dot{M}(t)}{M(t)} &= \phi^* \equiv 2\beta\left(\frac{C}{2}\right)^\theta \\ \lim_{t \rightarrow \infty} \frac{\dot{v}_i(t)}{v_i(t)} &= \frac{\phi^*}{\sigma - 1} = \frac{2\beta\left(\frac{C}{2}\right)^\theta}{\sigma - 1} \end{aligned}$$

Appendix 2: Technical Appendix

Appendix a

Theorem A1: The following identity holds for $i \neq j$:

$$\frac{a_{ij}/2}{n_i} = G(m_{ij}^d, m_{ji}^d)$$

where G is defined in (35).

Proof: Using (26) and (33),

$$\begin{aligned}
& \frac{a_{ij}/2}{n_i} \\
&= \frac{n^{ij}}{n_i} \cdot \frac{a_{ij}/2}{n^{ij}} \\
&= \frac{1}{1 - m_{ji}^d} \cdot \beta(m_{ij}^c)^\theta \cdot (m_{ij}^d \cdot m_{ji}^d)^{\frac{1-\theta}{2}} \\
&= \frac{\beta(1 - m_{ij}^d - m_{ji}^d)^\theta \cdot (m_{ij}^d \cdot m_{ji}^d)^{\frac{1-\theta}{2}}}{1 - m_{ji}^d} \\
&= G(m_{ij}^d, m_{ji}^d)
\end{aligned}$$

which leads to (34).

Theorem A2: Knowledge dynamics evolve according to the system:

$$\begin{aligned}
\dot{m}_{ij}^d &= (1 - \mu \cdot \eta)(1 - m_{ij}^d)(1 - m_{ji}^d) \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot 2G(m_{ik}^d, m_{ki}^d) \right\} \\
&\quad - m_{ij}^d [\mu \eta \alpha (1 - m_{ji}^d) \cdot \sum_{k=1}^N \delta_{kk} \cdot \frac{n_k}{n_i} + (1 - \mu \cdot \eta) \cdot \delta_{ij} \cdot (1 - m_{ji}^d) \cdot 2G(m_{ij}^d, m_{ji}^d) \\
&\quad + \mu \cdot \eta \cdot (1 - m_{ji}^d) \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot \frac{n_k}{n_i} \cdot G(m_{kl}^d, m_{lk}^d)] \\
&\quad - m_{ij}^d \cdot (1 - \mu \cdot \eta) \cdot (1 - m_{ij}^d) \cdot \left\{ \delta_{jj} \cdot \alpha + \sum_{k \neq i, j} \delta_{jk} \cdot 2G(m_{jk}^d, m_{kj}^d) \right\}
\end{aligned}$$

for $i, j = 1, 2, \dots, N, i \neq j$

Proof: By definition,

$$\begin{aligned}
\dot{m}_{ij}^d &= \frac{d(n_{ij}^d/n^{ij})}{dt} \\
&= \frac{\dot{n}_{ij}^d}{n^{ij}} - \frac{n_{ij}^d}{n^{ij}} \cdot \frac{\dot{n}^{ij}}{n^{ij}} \\
&= \frac{\dot{n}_{ij}^d}{n^{ij}} - m_{ij}^d \cdot \frac{\dot{n}^{ij}}{n^{ij}} \\
&= \frac{\dot{n}_{ij}^d}{n^{ij}} - m_{ij}^d \cdot \left(\frac{\dot{n}_{ij}^c}{n^{ij}} + \frac{\dot{n}_{ij}^d}{n^{ij}} + \frac{\dot{n}_{ji}^d}{n^{ij}} \right) \\
&= (1 - m_{ij}^d) \cdot \frac{\dot{n}_{ij}^d}{n^{ij}} - m_{ij}^d \cdot \left(\frac{\dot{n}_{ij}^c}{n^{ij}} + \frac{\dot{n}_{ji}^d}{n^{ij}} \right)
\end{aligned}$$

Setting $\mu = \mu(N)$, and using (42) and (33), we have

$$\begin{aligned}
\frac{\dot{n}_{ij}^d}{n^{ij}} &= \frac{(1 - \mu \cdot \eta) \cdot \sum_{k \neq j} \delta_{ik} \cdot a_{ik}}{n^{ij}} \\
&= (1 - \mu \cdot \eta) \cdot \left[\frac{\delta_{ii} \cdot \alpha \cdot n_i}{n^{ij}} + \sum_{k \neq i, j} \delta_{ik} \cdot \frac{a_{ik}}{n^{ij}} \right] \\
&= (1 - \mu \cdot \eta) \cdot \left[\frac{\delta_{ii} \cdot \alpha \cdot n_i}{n^{ij}} + \sum_{k \neq i, j} \delta_{ik} \cdot \frac{n_i}{n^{ij}} \cdot \frac{n^{ik}}{n_i} \cdot \frac{a_{ik}}{n^{ik}} \right] \\
&= (1 - \mu \cdot \eta) \cdot \frac{n_i}{n^{ij}} \cdot \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot \frac{n^{ik}}{n_i} \cdot \frac{a_{ik}}{n^{ik}} \right\} \\
&= (1 - \mu \cdot \eta) \cdot (1 - m_{ji}^d) \cdot \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot \frac{1}{1 - m_{ki}^d} \cdot 2\beta (1 - m_{ik}^d - m_{ki}^d)^\theta \cdot (m_{ik}^d \cdot m_{ki}^d)^{\frac{1-\theta}{2}} \right\} \\
&= (1 - \mu \cdot \eta) \cdot (1 - m_{ji}^d) \cdot \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot 2G(m_{ik}^d, m_{ki}^d) \right\}
\end{aligned}$$

Similarly,

$$\frac{\dot{n}_{ji}^d}{n^{ij}} = (1 - \mu \cdot \eta) \cdot (1 - m_{ij}^d) \cdot \left\{ \delta_{jj} \cdot \alpha + \sum_{k \neq i, j} \delta_{jk} \cdot 2G(m_{jk}^d, m_{kj}^d) \right\}$$

while using (41) yields

$$\begin{aligned}
\frac{\dot{n}_{ij}^c}{n^{ij}} &= \frac{(1 - \mu \cdot \eta) \delta_{ij} \cdot a_{ij} + \mu \cdot \eta A}{n^{ij}} \\
&= \frac{(1 - \mu \cdot \eta) \delta_{ij} \cdot a_{ij} + \mu \cdot \eta \cdot \sum_{k=1}^N \delta_{kk} \cdot a_{kk} + \mu \cdot \eta \cdot \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot (a_{kl}/2)}{n^{ij}} \\
&= \frac{\mu \cdot \eta \cdot \sum_{k=1}^N \delta_{kk} \cdot a_{kk}}{n^{ij}} + \frac{(1 - \mu \cdot \eta) \cdot \delta_{ij} \cdot a_{ij}}{n^{ij}} + \frac{\mu \cdot \eta \cdot \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot (a_{kl}/2)}{n^{ij}}
\end{aligned}$$

Using equations (25), (26), (34), and (33), we have:

$$\begin{aligned}
\frac{\dot{n}_{ij}^c}{n^{ij}} &= \mu \cdot \eta \cdot \alpha \cdot (1 - m_{ji}^d) \cdot \sum_{k=1}^N \delta_{kk} \cdot \frac{n_k}{n_i} + (1 - \mu \cdot \eta) \cdot \delta_{ij} \cdot 2\beta \cdot (1 - m_{ij}^d - m_{ji}^d)^\theta \cdot (m_{ij}^d \cdot m_{ji}^d)^{\frac{1-\theta}{2}} \\
&\quad + \mu \cdot \eta \cdot \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot \frac{n^{kl}}{n^{ij}} \cdot \beta \cdot (1 - m_{kl}^d - m_{lk}^d)^\theta \cdot (m_{kl}^d \cdot m_{lk}^d)^{\frac{1-\theta}{2}} \\
&= \mu \cdot \eta \cdot \alpha \cdot (1 - m_{ji}^d) \cdot \sum_{k=1}^N \delta_{kk} \cdot \frac{n_k}{n_i} + (1 - \mu \cdot \eta) \cdot \delta_{ij} \cdot (1 - m_{ji}^d) \cdot 2G(m_{ij}^d, m_{ji}^d) \\
&\quad + \mu \cdot \eta \cdot \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot \frac{n_k}{n_i} \cdot (1 - m_{ji}^d) \cdot G(m_{kl}^d, m_{lk}^d)
\end{aligned}$$

Thus,

$$\begin{aligned}
\dot{m}_{ij}^d &= (1 - \mu \cdot \eta)(1 - m_{ij}^d)(1 - m_{ji}^d) \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot 2G(m_{ik}^d, m_{ki}^d) \right\} \\
&\quad - m_{ij}^d [\mu \eta \alpha (1 - m_{ji}^d) \cdot \sum_{k=1}^N \delta_{kk} \cdot \frac{n_k}{n_i} + (1 - \mu \cdot \eta) \cdot \delta_{ij} \cdot (1 - m_{ji}^d) \cdot 2G(m_{ij}^d, m_{ji}^d)] \\
&\quad + \mu \cdot \eta \cdot (1 - m_{ji}^d) \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot \frac{n_k}{n_i} \cdot G(m_{kl}^d, m_{lk}^d) \\
&\quad - m_{ij}^d \cdot (1 - \mu \cdot \eta) \cdot (1 - m_{ij}^d) \cdot \left\{ \delta_{jj} \cdot \alpha + \sum_{k \neq i, j} \delta_{jk} \cdot 2G(m_{jk}^d, m_{kj}^d) \right\}
\end{aligned}$$

Appendix b

Here we confirm that when the expenditure of any specific consumer i is constant over time and given by (20), the total equilibrium expenditure in the economy is indeed given by equation (18).

For case (i-a), along the equilibrium path, the present value of income at time 0 for \mathbb{K} -worker i is given by

$$\begin{aligned}
W_i(0) &= \int_0^\infty e^{-\gamma t} y_i(t) dt \\
&= \int_0^\infty e^{-\gamma t} \cdot \left[\frac{\eta \cdot a(t)}{2} \cdot \Pi(t) \right] dt
\end{aligned}$$

Furthermore, as explained in case (i-a) of the Growth section,

$$\dot{M} = \eta \cdot A(t) = \frac{\eta \cdot N \cdot a(t)}{2}$$

Thus,

$$W_i(0) = \frac{1}{N} \int_0^\infty e^{-\gamma t} \cdot \dot{M}(t) \cdot \Pi(t) dt \tag{106}$$

where the patent price $\Pi(t)$ is obtained as follows, by using (15) and (16):

$$\begin{aligned}
\Pi(t) &= \int_t^\infty e^{-\gamma(\tau-t)} \cdot \pi^*(\tau) d\tau \\
&= \int_t^\infty e^{-\gamma(\tau-t)} \cdot \frac{q^*(\tau)}{\sigma - 1} d\tau \\
&= \frac{L}{\sigma - 1} \cdot \int_t^\infty e^{-\gamma(\tau-t)} \cdot \frac{1}{M(\tau)} d\tau
\end{aligned} \tag{107}$$

that yields

$$\dot{\Pi}(t) = \gamma \Pi(t) - \frac{L}{\sigma - 1} \cdot \frac{1}{M(t)} \tag{108}$$

Next, integrating (106) by parts and using (108), we obtain:

$$\begin{aligned}
W_i(0) &= \frac{1}{N} \int_0^\infty \dot{M}(t) \cdot (e^{-\gamma t} \cdot \Pi(t)) dt \\
&= \frac{1}{N} \left\{ [e^{-\gamma t} \cdot M(t) \cdot \Pi(t)] \Big|_0^\infty - \int_0^\infty M(t) \cdot \frac{d(e^{-\gamma t} \cdot \Pi(t))}{dt} dt \right\} \\
&= \frac{1}{N} \left\{ [e^{-\gamma t} \cdot M(t) \cdot \Pi(t)] \Big|_0^\infty - \int_0^\infty M(t) \cdot (-\gamma e^{-\gamma t} \cdot \Pi(t) + e^{-\gamma t} \cdot \dot{\Pi}(t)) dt \right\} \\
&= \frac{1}{N} \left\{ [e^{-\gamma t} \cdot M(t) \cdot \Pi(t)] \Big|_0^\infty + \gamma \int_0^\infty M(t) \cdot e^{-\gamma t} \cdot \Pi(t) dt - \int_0^\infty e^{-\gamma t} \cdot M(t) \cdot \dot{\Pi}(t) dt \right\} \\
&= \frac{1}{N} \left\{ [e^{-\gamma t} \cdot M(t) \cdot \Pi(t)] \Big|_0^\infty + \gamma \int_0^\infty e^{-\gamma t} \cdot M(t) \cdot \Pi(t) dt \right. \\
&\quad \left. - \int_0^\infty e^{-\gamma t} \cdot (M(t) \cdot \gamma \Pi(t) - \frac{L}{\sigma - 1}) dt \right\} \\
&= \frac{1}{N} \left\{ [e^{-\gamma t} \cdot M(t) \cdot \Pi(t)] \Big|_0^\infty + \int_0^\infty e^{-\gamma t} \cdot \frac{L}{\sigma - 1} dt \right\}
\end{aligned}$$

that leads to

$$W_i(0) = \frac{1}{N} \left\{ \lim_{t \rightarrow \infty} e^{-\gamma t} \cdot M(t) \cdot \Pi(t) - M(0) \cdot \Pi(0) + \frac{L}{\sigma - 1} \cdot \frac{1}{\gamma} \right\}$$

Thus, using (20) and (8), for any specific \mathbb{K} -worker i , expenditure is:

$$\begin{aligned}
E_i &= \gamma \left(\frac{\Pi(0) \cdot M(0)}{N} + W_i(0) \right) \\
&= \frac{1}{N} \left\{ \gamma \cdot \lim_{t \rightarrow \infty} e^{-\gamma t} \cdot M(t) \cdot \Pi(t) + \frac{L}{\sigma - 1} \right\}
\end{aligned}$$

In order to evaluate the first term in this expression, observe that by (107),

$$\begin{aligned}
M(t) \cdot \Pi(t) &= \frac{L}{\sigma - 1} \cdot \int_t^\infty e^{-\gamma(\tau-t)} \cdot \frac{M(t)}{M(\tau)} d\tau \\
&< \frac{L}{\sigma - 1} \cdot \int_t^\infty e^{-\gamma(\tau-t)} d\tau = \frac{L}{\sigma - 1} \cdot \frac{1}{\gamma}
\end{aligned}$$

The second line follows since for $\tau > t$, $\frac{M(t)}{M(\tau)} < 1$. So

$$\lim_{t \rightarrow \infty} e^{-\gamma t} \cdot M(t) \cdot \Pi(t) = 0 \tag{109}$$

and

$$E_i = \frac{1}{N} \cdot \frac{L}{\sigma - 1}$$

Therefore, the total expenditure of all \mathbb{K} -workers together is:

$$N \cdot E_i = \frac{L}{\sigma - 1} \tag{110}$$

For any specific \mathbb{M} -worker i , we have $y_i(t) = w^M \equiv 1$ for every time t . Thus,

$$W_i(0) = \int_0^\infty e^{-\gamma t} \cdot y_i(t) dt = \frac{1}{\gamma}$$

Noting that $\omega_i = 0$ by assumption for any \mathbb{M} -worker, equation (20) yields:

$$E_i = 1$$

So the total expenditure of all \mathbb{M} -workers together is:

$$L \cdot E_i = L \tag{111}$$

Summing (110) and (111) yields the total expenditure of consumers in the economy:

$$E^* = \frac{L}{\sigma - 1} + L = \frac{\sigma L}{\sigma - 1} = \frac{L}{\rho} \tag{112}$$

Therefore relation (18) is verified for the equilibrium path.

Appendix c

Here we discuss efficiency in the context of an intertemporal utilitarian social welfare function. We consider the following planner's problem, where the planner chooses $\{\delta_{ij}(\cdot)\}_{i,j=1}^N$ in order to maximize the sum of \mathbb{M} -workers' and \mathbb{K} -workers' utility given by (23).

Using (102), (43) and (44), the planner's problem is given by:

Choose piecewise continuous $\{\delta_{ij}(t)\}$ for $i, j = 1, \dots, N$ and $t \geq 0$ so as to

$$\max W = \frac{L}{\sigma - 1} \int_0^\infty e^{-\gamma t} \cdot \ln(M(t)) dt \tag{113}$$

subject to

$$\begin{aligned} \dot{M} &= \eta \cdot A \\ &= \eta \cdot \sum_{k=1}^N n_k \left(\delta_{kk} \cdot \alpha + \sum_{l \neq k} \delta_{kl} \cdot G(m_{kl}^d, m_{lk}^d) \right) \\ \dot{n}_i &= (1 - \mu\eta) \cdot n_i \cdot \left\{ \delta_{ii} \cdot \alpha + 2 \sum_{j \neq i} \delta_{ij} \cdot G(m_{ij}^d, m_{ji}^d) \right\} \\ &\quad + \mu\eta \cdot \sum_{k=1}^N n_k \left(\delta_{kk} \cdot \alpha + \sum_{l \neq k} \delta_{kl} \cdot G(m_{kl}^d, m_{lk}^d) \right) \end{aligned}$$

and

$$\begin{aligned}
\dot{m}_{ij}^d &= (1 - \mu \cdot \eta)(1 - m_{ij}^d)(1 - m_{ji}^d) \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot 2G(m_{ik}^d, m_{ki}^d) \right\} \\
&\quad - m_{ij}^d [\mu \eta \alpha (1 - m_{ji}^d) \cdot \sum_{k=1}^N \delta_{kk} \cdot \frac{n_k}{n_i} + (1 - \mu \cdot \eta) \cdot \delta_{ij} \cdot (1 - m_{ji}^d) \cdot 2G(m_{ij}^d, m_{ji}^d) \\
&\quad + \mu \cdot \eta \cdot (1 - m_{ji}^d) \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot \frac{n_k}{n_i} \cdot G(m_{kl}^d, m_{lk}^d)] \\
&\quad - m_{ij}^d \cdot (1 - \mu \cdot \eta) \cdot (1 - m_{ij}^d) \cdot \left\{ \delta_{jj} \cdot \alpha + \sum_{k \neq i, j} \delta_{jk} \cdot 2G(m_{jk}^d, m_{kj}^d) \right\}
\end{aligned}$$

given $n_i(0) > 0$ and $m_{ij}^d(0) > 0$, for $i, j = 1, \dots, N$. The equality in the objective function follows from (22) and (112). We must also account for the obvious constraints:

$$\begin{aligned}
\sum_{j=1}^N \delta_{ij} &= 1 \text{ for each } i = 1, \dots, N \\
\delta_{ij} &= \delta_{ji} \text{ for each } i, j = 1, \dots, N \\
\delta_{ij} &\geq 0 \text{ for each } i, j = 1, \dots, N
\end{aligned}$$

We assume that the discount rate is sufficiently large, $\gamma > g(m^B)$, in order to ensure that the objective is finite.

We assume symmetric initial conditions,

$$n_i(0) = n_j(0) = n(0) > 0 \text{ for } i, j = 1, \dots, N. \quad (114)$$

that implies

$$m_{ij}^d(0) = m_{ji}^d(0) \equiv m^d(0) > 0 \text{ for } i, j = 1, \dots, N. \quad (115)$$

Given our welfare function in equation (113), choosing $\{\delta_{ij}\}$ to maximize total income at a given time t , namely choosing the myopic core path, yields a growth path that is not dominated in the very short run. In the long run, our myopic core path reaches the bliss point in finite time; the bliss point is the maximal productivity for any path. Thus, after a certain interval of time, our myopic core path weakly overtakes any other path, in the sense that after this initial interval of time, the payoffs from the myopic core path are at least as high as those from any other path. Next, we focus on the intermediate run, the time intervals not covered by the short and long run cases.

To study the intermediate run, our analysis proceeds as follows. First, in Lemma 1, we shall compute the rate of increase in differential knowledge for pairs of \mathbb{K} -workers, given the initial conditions for case (i-a), (118), for our limiting myopic core path as the number of \mathbb{K} -workers tends to infinity. Lemmas 2 and 3 show that this rate of increase dominates that of any alternative path from time 0 until the bliss point is attained, provided that the alternative path satisfies two further symmetry conditions. We conclude from the first three lemmas that for any alternative feasible path satisfying the additional symmetry conditions, if the number of \mathbb{K} -workers is sufficiently large, our limiting myopic core path will dominate it in terms of the rate of increase of differential knowledge. This is depicted in Figure 5. Lemma 4 shows that an improvement in the rate of increase of differential knowledge has a positive impact on the paths of patent and idea production. Thus, the limiting myopic core path provides an upper bound on welfare achievable by any alternative feasible path satisfying the symmetry conditions. Finally, Lemma 5 shows that the difference between the welfare level generated by our myopic core path and our limiting myopic core path tends to zero as the number of knowledge workers tends to infinity, thus demonstrating that our myopic core path is asymptotically efficient as the number of knowledge workers tends to infinity.

When $m^d(t) < m^B$, then knowledge productivity is higher and m_{ij}^d moves almost as fast to the right as working in isolation if each person works with every other person with equal intensity. The intuition for this result follows from a combination of two reasons. First, productivity is higher when working with others as opposed to working alone on this part of the path. Second, when N is sufficiently large, working with others is very close to working in isolation when the accumulation of differential knowledge is considered, so cooperation with others will be better on net. Once the bliss point is attained, the system reaches the highest productivity possible, and remains there. This intuition indicates that, when $m^d(t) < m^B$, working with a smaller group than the other $N - 1$ \mathbb{K} -workers results in movement to the right that is slower than working with everyone but oneself.

Here we introduce symmetry of the admissible paths, a second symmetry restriction after the restriction to symmetric initial conditions. In our analysis of efficiency, we restrict attention to pairwise symmetric paths. If the initial state is pairwise symmetric, then the equilibrium is pairwise symmetric. However, in our efficiency analysis, we impose this assumption for tractability.

Later, we shall impose a third symmetry restriction for some of the analysis.

For tractability and the sake of comparison with the equilibrium path, we restrict our efficiency analysis to paths satisfying the symmetry restriction: For each knowledge worker i and for every knowledge worker $j \neq i$,

$$n_i(t) = n_j(t) \equiv n(t) \text{ for all } t \quad (116)$$

As explained in section 3.1, this is equivalent to

$$m_{ij}^d(t) = m_{ji}^d(t) \text{ for all } t \quad (117)$$

Furthermore, we focus on case (i-a), namely when the initial heterogeneity is to the left of the bliss point and where the public knowledge externality is not too strong:

$$m^J < m^d(0) \leq m^B \text{ and } C < \tilde{C}. \quad (118)$$

Using these restrictions and (58), we can restate the optimization problem in a simpler fashion: Maximize

$$W = \frac{L}{\sigma - 1} \int_0^\infty e^{-\gamma t} \cdot \ln(M) dt \quad (119)$$

subject to

$$\begin{aligned} \dot{M} &= \eta \cdot A \\ &= \eta \cdot \sum_{k=1}^N n_i \left\{ \delta_{kk} \cdot \alpha + \sum_{l \neq k} \delta_{kl} \cdot g(m_{kl}^d) \right\} \end{aligned} \quad (120)$$

$$\begin{aligned} \dot{n}_i &= \left(1 - \frac{C}{N}\right) \cdot n_i \cdot \left\{ \delta_{ii} \cdot \alpha + 2 \sum_{j \neq i} \delta_{ij} \cdot g(m_{ij}^d) \right\} \\ &+ \frac{C}{N} \cdot \sum_{k=1}^N n_k \left\{ \delta_{kk} \cdot \alpha + \sum_{l \neq k} \delta_{kl} \cdot g(m_{kl}^d) \right\} \end{aligned} \quad (121)$$

and

$$\begin{aligned} \dot{m}_{ij}^d &= \left(1 - \frac{C}{N}\right) (1 - m_{ij}^d) (1 - m_{ij}^d) \left\{ \delta_{ii} \cdot \alpha + \sum_{k \neq i, j} \delta_{ik} \cdot 2g(m_{ik}^d) \right\} \\ &- m_{ij}^d \left[\frac{C}{N} \cdot \alpha (1 - m_{ij}^d) \cdot \sum_{k=1}^N \delta_{kk} + \left(1 - \frac{C}{N}\right) \cdot \delta_{ij} \cdot (1 - m_{ij}^d) \cdot 2g(m_{ij}^d) \right. \\ &+ \left. \frac{C}{N} \cdot (1 - m_{ij}^d) \sum_{k=1}^N \sum_{l \neq k} \delta_{kl} \cdot g(m_{kl}^d) \right] \\ &- m_{ij}^d \cdot \left(1 - \frac{C}{N}\right) \cdot (1 - m_{ij}^d) \cdot \left\{ \delta_{jj} \cdot \alpha + \sum_{k \neq i, j} \delta_{jk} \cdot 2g(m_{jk}^d) \right\} \end{aligned} \quad (122)$$

subject to initial conditions (114) and (115) as well as symmetry conditions (116) and (117) under the obvious restrictions on the control variables.

Using (27) and (38), we obtain

$$\sum_{i=1}^N y_i = \Pi \eta A$$

On the equilibrium path, each \mathbb{K} -worker takes the patent price Π as given. Hence

$$\max \sum_{i=1}^N y_i \Leftrightarrow \max \dot{M} = \eta A$$

Lemma 1: *For case (i-a), on the myopic core path, for each fixed m^d , the value of \dot{m}^d given by equation (78) is increasing in the number of \mathbb{K} -workers, N . Moreover, along the myopic core path,*

$$\begin{aligned} \lim_{N \rightarrow \infty} \dot{m}^d &\equiv \dot{m}_\infty^d = 2(1 - m^d) \cdot g(m^d) \cdot \left\{1 - m^d \left[2 + \frac{C}{2}\right]\right\} \\ &= (1 - m^d) \cdot g(m^d) \cdot \{2(1 - 2m^d) - Cm^d\} \end{aligned} \quad (123)$$

Proof: Inspecting equation (78), for fixed m^d the expression on the right hand side is increasing in N . Taking the limit of equation (78) as N tends to infinity yields the second result. ■

Solving the differential equation (123) using initial conditions (115), we define the t_∞^B to be the time when the limiting myopic core path reaches the bliss point, namely

$$m_\infty^d(t_\infty^B) = m^B$$

When the limiting myopic core path reaches m^B , we assume that the \mathbb{K} -workers split into groups of optimal size \tilde{N}^B given by (92), and stay at the bliss point m^B . Hence we set

$$m_\infty^d(t) = m^B \text{ for } t \geq t_\infty^B$$

The top curve in Figure 5 depicts the limiting myopic core path.

As promised above, we now introduce a third symmetry condition. At each moment of time, each person interacts with the same number of people with the same intensity. In other words,

For each \mathbb{K} -worker i , there is a subset of \mathbb{K} -workers \mathbb{N}_i^a , $i \notin \mathbb{N}_i^a$ and $1 \geq \delta_I \geq 0$

$$(124)$$

such that $|\mathbb{N}_i^a| = N^a \geq 1$ for all i ,

and for all $t < t_s^B$, $j \in \mathbb{N}_i^a \implies \delta_{ij}(t) = \delta_a \equiv \frac{1 - \delta_I}{N^a}$

where t_s^B will be defined shortly. An implication is that at each time, each \mathbb{K} -worker spends δ_I fraction of time working in isolation.²⁹

For the myopic core path, notice that $N^a = N - 1$, whereas $\delta_I = 0$. So for admissible alternative paths, \mathbb{K} -workers can form smaller subgroups or work in isolation.

Next, in preparation for Lemma 2, we perform some preliminary calculations of path dynamics for the active and shadow partners. With a focus on \mathbb{K} -worker i and potential partner j , we calculate the dynamics of m_{ij}^d when i and j are active partners, namely $j \in \mathbb{N}_i^a$. Since all active pairs are symmetric, we have

$$\begin{aligned} m_{ij}^d(t) &= m_{ji}^d(t) \equiv m_a^d(t) \text{ for each } j \in \mathbb{N}_i^a, \\ &\text{for all } i = 1, 2, \dots, N \text{ and for all } t \\ m_{ij}^d(t) &= m_{ji}^d(t) \equiv m_s^d(t) \text{ for each } j \notin \mathbb{N}_i^a, \\ &\text{for all } i = 1, 2, \dots, N \text{ and for all } t \end{aligned} \tag{125}$$

In (122), using (125), setting $\delta_{kk} = \delta_I$ for all $k = 1, \dots, N$, we set

$$\begin{aligned} \delta_{ij}(t) &= \frac{1 - \delta_I}{N^a} \text{ when } j \in \mathbb{N}_i^a \\ \delta_{ik}(t) &= \frac{1 - \delta_I}{N^a} \text{ when } k \in \mathbb{N}_i^a \\ \delta_{kl}(t) &= \frac{1 - \delta_I}{N^a} \text{ when } l \in \mathbb{N}_k^a \end{aligned}$$

Then for active partner $j \in \mathbb{N}_i^a$, setting $m_{ij}^d = m_a^d$, we obtain

$$\begin{aligned} \frac{\dot{m}_a^d}{1 - m_a^d} &= \left(1 - \frac{C}{N}\right)(1 - m_a^d) \left\{ \delta_I \cdot \alpha + 2g(m_a^d)(1 - \delta_I - \delta_a) \right\} \\ &\quad - m_a^d \left[C \cdot \alpha \cdot \delta_I + \left(1 - \frac{C}{N}\right) \cdot 2g(m_a^d)\delta_a \right. \\ &\quad \left. + C \cdot g(m_a^d) \cdot (1 - \delta_I) \right] \\ &\quad - m_a^d \cdot \left(1 - \frac{C}{N}\right) \cdot \left\{ \delta_I \cdot \alpha + 2g(m_a^d)(1 - \delta_I - \delta_a) \right\} \\ &= \left\{ 2\left(1 - \frac{C}{N}\right)(1 - 2m_a^d) - Cm_a^d \right\} \cdot g(m_a^d) \cdot (1 - \delta_I) \\ &\quad + \left\{ \left(1 - \frac{C}{N}\right)(1 - 2m_a^d) - Cm_a^d \right\} \alpha \cdot \delta_I - \left(1 - \frac{C}{N}\right) \cdot (1 - m_a^d) \cdot 2g(m_a^d) \cdot \delta_a \end{aligned} \tag{126}$$

²⁹Even when we allow each interaction set $\mathbb{N}_i^a(t)$ to vary with time, we get essentially the same results below. But such a generalization would complicate the notation and the arguments. In fact, we permit (but do not require) a change in δ_I and $\mathbb{N}_i^a(t)$ at time t_s^B in our analysis.

whereas for shadow partner $j \notin N_i^a$, setting $m_{ij}^d = m_s^d$

$$\begin{aligned}
\frac{\dot{m}_s^d}{1 - m_s^d} &= \left(1 - \frac{C}{N}\right)(1 - m_s^d) \left\{ \delta_I \cdot \alpha + 2g(m_a^d)(1 - \delta_I) \right\} \\
&\quad - m_s^d [C \cdot \alpha \cdot \delta_I + C \cdot g(m_a^d)(1 - \delta_I)] \\
&\quad - m_s^d \cdot \left(1 - \frac{C}{N}\right) \cdot \left\{ \delta_I \cdot \alpha + 2g(m_a^d)(1 - \delta_I) \right\} \\
&= \left\{ 2\left(1 - \frac{C}{N}\right)(1 - 2m_s^d) - C m_s^d \right\} \cdot g(m_a^d) \cdot (1 - \delta_I) \\
&\quad + \left\{ \left(1 - \frac{C}{N}\right)(1 - 2m_s^d) - C m_s^d \right\} \alpha \cdot \delta_I
\end{aligned} \tag{127}$$

Solving this system of differential equations (126) and (127) using initial conditions (115), we obtain the active path, m_a^d , and the shadow path, m_s^d . Then we define the t_s^B to be the time when the shadow path reaches the bliss point, namely³⁰

$$m_s^d(t_s^B) = m^B$$

Lemma 2: *Under the additional symmetry condition (124), $m_s^d(t) > m_a^d(t)$ for all $0 < t < t_s^B$.*

Proof: From the initial condition (115), $m_s^d(0) = m_a^d(0) = m^d(0)$. Thus, using (126) and (127)

$$\dot{m}_s^d(0) - \dot{m}_a^d(0) = \left(1 - \frac{C}{N}\right) \cdot (1 - m_a^d) \cdot 2g(m_a^d) \cdot \delta_a$$

that is positive by assumption. Thus, at least for small t , $m_s^d(t) > m_a^d(t)$. Next, we show by contradiction that this relationship holds for all $t > 0$. Suppose to the contrary, at time t that $m_s^d(t) \leq m_a^d(t)$. Then by continuity of the active and shadow paths, there is some minimal time $0 < t' < t$ such that $m_s^d(t') = m_a^d(t')$. Then using (126) and (127), we have that at time t'

$$\dot{m}_s^d(t') - \dot{m}_a^d(t') = \left(1 - \frac{C}{N}\right) \cdot (1 - m_a^d) \cdot 2g(m_a^d) \cdot \delta_a$$

that is positive again by assumption. This results in a contradiction, because $m_s^d(t') = m_a^d(t')$ and $\dot{m}_s^d(t') > \dot{m}_a^d(t')$ mean that there is a $0 < \hat{t} < t'$ where $m_s^d(\hat{t}) = m_a^d(\hat{t})$. ■

³⁰If the bliss point is not attained by the shadow path in finite time, then set $t_s^B = \infty$.

Next we show that the path representing shadow partners, m_s^d , is dominated by the limiting myopic path m_∞^d , as illustrated in Figure 5.

Lemma 3: For all $0 < t < t_s^B$,

$$\dot{m}_\infty^d(t') > \dot{m}_s^d(t) \text{ whenever } m_\infty^d(t') = m_s^d(t) \quad (128)$$

implying that

$$m_\infty^d(t) > m_s^d(t) \text{ for all } 0 < t < t_s^B$$

Proof: When $m_\infty^d(t') = m_s^d(t) \equiv m^d$, using (123) and (127)

$$\begin{aligned} \frac{\dot{m}_\infty^d(t') - \dot{m}_s^d(t)}{1 - m^d} &= \{2(1 - 2m^d) - Cm^d\} \cdot g(m^d) \\ &\quad - \{2(1 - \frac{C}{N})(1 - 2m^d) - Cm^d\} \cdot g(m_a^d(t)) \cdot (1 - \delta_I) \\ &\quad - \{(1 - \frac{C}{N})(1 - 2m^d) - Cm^d\} \alpha \cdot \delta_I \end{aligned} \quad (129)$$

By Lemma 2, $m^B > m^d(t) > m_a^d(t)$ for all $0 < t < t_s^B$. When $\delta_I = 0$,

$$\begin{aligned} \frac{\dot{m}_\infty^d(t') - \dot{m}_s^d(t)}{1 - m^d} &= \{2(1 - 2m^d) - Cm^d\} \cdot g(m^d) \\ &\quad - \{2(1 - \frac{C}{N})(1 - 2m^d) - Cm^d\} \cdot g(m_a^d(t)) \end{aligned}$$

Since $g(m^d(t)) > g(m_a^d(t))$, we can readily conclude that $\dot{m}_\infty^d(t') > \dot{m}_s^d(t)$. When $\delta_I = 1$,

$$\begin{aligned} \frac{\dot{m}_\infty^d(t') - \dot{m}_s^d(t)}{1 - m^d} &= \{2(1 - 2m^d) - Cm^d\} \cdot g(m^d) \\ &\quad - \{(1 - \frac{C}{N})(1 - 2m^d) - Cm^d\} \alpha \end{aligned}$$

By assumption, $g(m^d) > \alpha$, and hence we can conclude that $\dot{m}_\infty^d(t') > \dot{m}_s^d(t)$. By the linearity of (129), we can conclude that (128) holds for all $0 < t < t_s^B$. Given the same initial conditions for the two differential equations, the result is proved. ■

Combining Lemmas 2 and 3, we have the situation depicted in Figure 5. For any $0 < t < t_s^B$, the limiting myopic core path m_∞^d strictly dominates the shadow path m_s^d , which in turn strictly dominates the active path m_a^d .

Fixing δ_I and \mathbb{N}_i^a , we have obtained the growth path (m_a^d, m_s^d) for $0 < t < t_s^B$. In fact, they act in concert, generating a growth pattern for both

the active and shadow partners. Beyond t_s^B , we consider two alternative symmetric growth patterns. At time t_s^B , we allow the \mathbb{K} -workers to choose any new δ_I' and $\mathbb{N}_i^{a'}$ (possibly the same as δ_I and \mathbb{N}_i^a) satisfying (124). We solve the differential equations (126) and (127) as before, with initial conditions

$$\begin{aligned} j \in \mathbb{N}_i^a &\Rightarrow m_{ij}^d(t_s^B) = m_a^d(t_s^B) \\ j \notin \mathbb{N}_i^a &\Rightarrow m_{ij}^d(t_s^B) = m_s^d(t_s^B) = m^B \end{aligned}$$

Analogous to previous notation, we denote the associated growth path by (m_a^{df}, m_s^{df}) for $t \geq t_s^B$. We denote by $m_f^d = (m_a^{df}, m_s^{df})$ the growth pattern that follows (m_a^d, m_s^d) for $0 < t < t_s^B$ and (m_a^{df}, m_s^{df}) for $t \geq t_s^B$, that is by definition feasible. By varying δ_I' and $\mathbb{N}_i^{a'}$, we can obtain many growth patterns. Next we establish an upper bound on all such growth patterns. Define this upper bound, called \bar{m}^d , by

$$\begin{aligned} \bar{m}_{ij}^d(t) &= m_a^d \text{ if } j \in \mathbb{N}_i^a, \bar{m}_{ij}^d(t) = m_s^d \text{ if } j \notin \mathbb{N}_i^a \text{ for } 0 < t < t_s^B \\ \bar{m}_{ij}^d(t) &= \bar{m}^d(t) = m^B \text{ for all } i, j = 1, 2, \dots, N \text{ and for all } t \geq t_s^B \end{aligned}$$

This path might not be feasible after time t_s^B , but it nevertheless establishes an upper bound on feasible paths.

Now we have three alternative growth patterns: m_∞^d , \bar{m}^d , and an arbitrary feasible path, m_f^d . By comparing the welfare generated by each of the three growth patterns, we can readily conclude as follows.

Lemma 4: *The limiting myopic core growth pattern gives an upper bound on the patent and idea paths generated by any feasible growth pattern. More precisely,*

$$\begin{aligned} \dot{M}(t) \Big|_{m_\infty^d} &> \dot{M}(t) \Big|_{\bar{m}^d} \geq \dot{M}(t) \Big|_{m_f^d} \text{ for } 0 < t < t_s^B \\ \dot{M}(t) \Big|_{m_\infty^d} &= \dot{M}(t) \Big|_{\bar{m}^d} \geq \dot{M}(t) \Big|_{m_f^d} \text{ for } t > t_s^B \\ \dot{n}_i(t) \Big|_{m_\infty^d} &> \dot{n}_i(t) \Big|_{\bar{m}^d} \geq \dot{n}_i(t) \Big|_{m_f^d} \text{ for all } i = 1, 2, \dots, N, \text{ for all } 0 < t < t_s^B \\ \dot{n}_i(t) \Big|_{m_\infty^d} &> \dot{n}_i(t) \Big|_{\bar{m}^d} \geq \dot{n}_i(t) \Big|_{m_f^d} \text{ for all } i = 1, 2, \dots, N, \text{ for all } t > t_s^B \end{aligned} \tag{130}$$

Proof: In equation (120), we focus on the terms in parentheses, evaluating

them before and after time t_s^B : for all $k = 1, 2, \dots, N$, for all $0 < t < t_s^B$,

$$\begin{aligned} \delta_{kk}(t) \cdot \alpha + \sum_{l \neq k} \delta_{kl}(t) \cdot g(m_{kl}^d(t)) \big|_{m_\infty^d} &= g(m_\infty^d(t)) \\ &> g(m_a^d(t)) = \delta_{kk}(t) \cdot \alpha + \sum_{l \neq k} \delta_{kl}(t) \cdot g(m_{kl}^d(t)) \big|_{\bar{m}^d} \\ &= \delta_{kk}(t) \cdot \alpha + \sum_{l \neq k} \delta_{kl}(t) \cdot g(m_{kl}^d(t)) \big|_{m_f^d} \end{aligned}$$

For all $k = 1, 2, \dots, N$, for all $t \geq t_s^B$,

$$\begin{aligned} \delta_{kk}(t) \cdot \alpha + \sum_{l \neq k} \delta_{kl}(t) \cdot g(m_{kl}^d(t)) \big|_{m_\infty^d} &= g(m^B) \\ &= \delta_{kk}(t) \cdot \alpha + \sum_{l \neq k} \delta_{kl}(t) \cdot g(m_{kl}^d(t)) \big|_{\bar{m}^d} \\ &\geq \delta_{kk}(t) \cdot \alpha + \sum_{l \neq k} \delta_{kl}(t) \cdot g(m_{kl}^d(t)) \big|_{m_f^d} \end{aligned}$$

For the terms in both parentheses in equation (121), we obtain an analogous result. Thus, by the nature of the differential equations (120) and (121), the four lines of the relationship (130) follow. ■

Next we outline our strategy for the remainder of the analysis. First, for a fixed population of \mathbb{K} -workers N , we define the levels of welfare generated by three different paths. Let $W_\infty(N)$ be the level of welfare generated by the m_∞^d path, but for population size N . Let $\bar{W}(N)$ be the level of welfare generated by the \bar{m}^d path for the same population size N . Finally, let $W_f(N)$ be the level of welfare generated by the m_f^d path for the same population size N . By Lemma 4, it is clear that

$$W_\infty(N) > \bar{W}(N) > W_f(N)$$

Therefore we can conclude that *for each fixed population of \mathbb{K} -workers N , the limiting myopic core growth pattern gives an upper bound on the welfare generated by any feasible growth pattern.*

Let $W_{mc}(N)$ be the level of welfare generated by the myopic core path with N \mathbb{K} -workers. It remains for us to show that:

$$\lim_{N \rightarrow \infty} \{W_\infty(N) - W_{mc}(N)\} = 0$$

To accomplish this task, we must define $W_{mc}(N)$ and $W_\infty(N)$ formally. After that, the result follows almost immediately.

First, recalling (78) and applying (119) to (122) to the case of the myopic core path, $W_{mc}(N)$ is defined as follows:

$$W_{mc}(N) \equiv \frac{L}{\rho} \cdot \int_0^\infty e^{-\gamma t} \ln[M_{mc}(t)] dt$$

where

$$\dot{M}_{mc} = \eta \cdot n_{mc} \cdot g(m_{mc}^d) \cdot N \quad (131)$$

$$\dot{n}_{mc} = n_{mc} \cdot g(m_{mc}^d) \cdot n \cdot \left(2 + \frac{N-2}{N}C\right) \quad (132)$$

$$\dot{m}_{mc}^d = \begin{cases} 2(1-m_{mc}^d) \cdot g(m_{mc}^d) \cdot \frac{1-\frac{C}{N}}{N-1} \left\{ (N-2)-m_{mc}^d \left[(2N-3) + \frac{C}{2} \cdot \frac{N-1}{1-\frac{C}{N}} \right] \right\} & \text{for } t < t_{mc}^B \\ 0 & \text{for } t \geq t_{mc}^B \end{cases} \quad (133)$$

given the initial conditions

$$m_{mc}^d(0) = m^d(0) \quad (134)$$

$$n_{mc}(0) = n(0) \quad (135)$$

$$M_{mc}(0) = z(0) \cdot N \quad (136)$$

Here t_{mc}^B is the time when the m_{mc}^d reaches m^B , whereas $z(0)$ is the initial number of patents owned by each \mathbb{K} -worker.

Let $m_{mc}^d(t, N)$ be the solution to the differential equation (133) subject to the initial condition (134) for the given N . Then, using (135), we can solve (132) for $n_{mc}(t, N)$, that satisfies the following relationship:

$$n_{mc}(t, N) = n(0) + \left\{ \int_0^t g(m_{mc}^d(\tau, N)) \cdot n_{mc}(\tau, N) d\tau \right\} \cdot \left(2 + \frac{N-2}{N}C\right)$$

Finally, solving (131) with initial condition (136), we obtain

$$M_{mc}(t, N) = f_{mc}(t, N) \cdot N$$

where

$$f_{mc}(t, N) = z(0) + \eta \cdot \int_0^t n_{mc}(\tau, N) \cdot g(m_{mc}^d(\tau, N)) d\tau$$

This path yields the associated level of welfare:

$$W_{mc}(N) = \frac{L}{\rho} \cdot \int_0^\infty e^{-\gamma t} \ln[f_{mc}(t, N) \cdot N] d\tau$$

Similarly, we define the welfare generated by the m_∞^d path as follows:

$$W_\infty(N) \equiv \frac{L}{\rho} \cdot \int_0^\infty e^{-\gamma t} \ln[M_\infty(t)] dt$$

where

$$\dot{M}_\infty = \eta \cdot n_\infty \cdot g(m_\infty^d) \cdot N \quad (137)$$

$$\dot{n}_\infty = n_\infty \cdot g(m_\infty^d) \cdot (2 + C) \quad (138)$$

$$\dot{m}_\infty^d = \begin{cases} (1-m_\infty^d) \cdot g(m_\infty^d) \cdot \{2(1-2m_\infty^d) - Cm_\infty^d\} & \text{for } t < t_{mc}^B \\ 0 & \text{for } t \geq t_{mc}^B \end{cases} \quad (139)$$

given the initial conditions

$$m_\infty^d(0) = m^d(0) \quad (140)$$

$$n_\infty(0) = n(0) \quad (141)$$

$$M_\infty(0) = z(0) \cdot N \quad (142)$$

Let $m_\infty^d(t)$ be the solution to the differential equation (139) subject to the initial condition (140). Then, using (141), we can solve (138) for $n_\infty(t)$. It satisfies the following relationship:

$$n_\infty(t) = n(0) + \left\{ \int_0^t g(m_\infty^d(\tau)) \cdot n_\infty(\tau) d\tau \right\} \cdot (2 + C)$$

that is independent of N . Finally, solving (137) with initial condition (140), we obtain

$$M_\infty(t, N) = f_\infty(t) \cdot N$$

where

$$f_\infty(t) = z(0) + \eta \cdot \int_0^t n_\infty(\tau) \cdot g(m_\infty^d(\tau)) d\tau$$

Notice that $f_\infty(t)$ is also independent of N . This path yields the associated level of welfare:

$$W_\infty(N) = \frac{L}{\rho} \cdot \int_0^\infty e^{-\gamma t} \ln[f_\infty(t) \cdot N] dt$$

We can readily see that by construction,

$$\lim_{N \rightarrow \infty} f_{mc}(t, N) = f_\infty(t) \text{ uniformly in } t$$

Therefore, taking the difference of the welfare levels,

$$\begin{aligned}
\lim_{N \rightarrow \infty} \{W_\infty(N) - W_{mc}(N)\} &= \lim_{N \rightarrow \infty} \left\{ \frac{L}{\rho} \cdot \int_0^\infty e^{-\gamma t} \ln[f_\infty(t) \cdot N] d\tau \right. \\
&\quad \left. - \frac{L}{\rho} \cdot \int_0^\infty e^{-\gamma t} \ln[f_{mc}(t, N) \cdot N] d\tau \right\} \\
&= \frac{L}{\rho} \lim_{N \rightarrow \infty} \left\{ \int_0^\infty e^{-\gamma t} \ln[f_\infty(t)] d\tau + \int_0^\infty e^{-\gamma t} \ln[N] d\tau \right. \\
&\quad \left. - \int_0^\infty e^{-\gamma t} \ln[f_{mc}(t, N)] d\tau - \int_0^\infty e^{-\gamma t} \ln[N] d\tau \right\} \\
&= \frac{L}{\rho} \lim_{N \rightarrow \infty} \left\{ \int_0^\infty e^{-\gamma t} (\ln[f_\infty(t)] - \ln[f_{mc}(t, N)]) d\tau \right\} \\
&= 0
\end{aligned}$$

Finally, we can conclude our analysis as follows:

Lemma 5: *As the number of knowledge workers N tends to infinity, the difference in welfare corresponding to the myopic core path and the welfare corresponding to the limiting myopic core path monotonically converges to 0.*

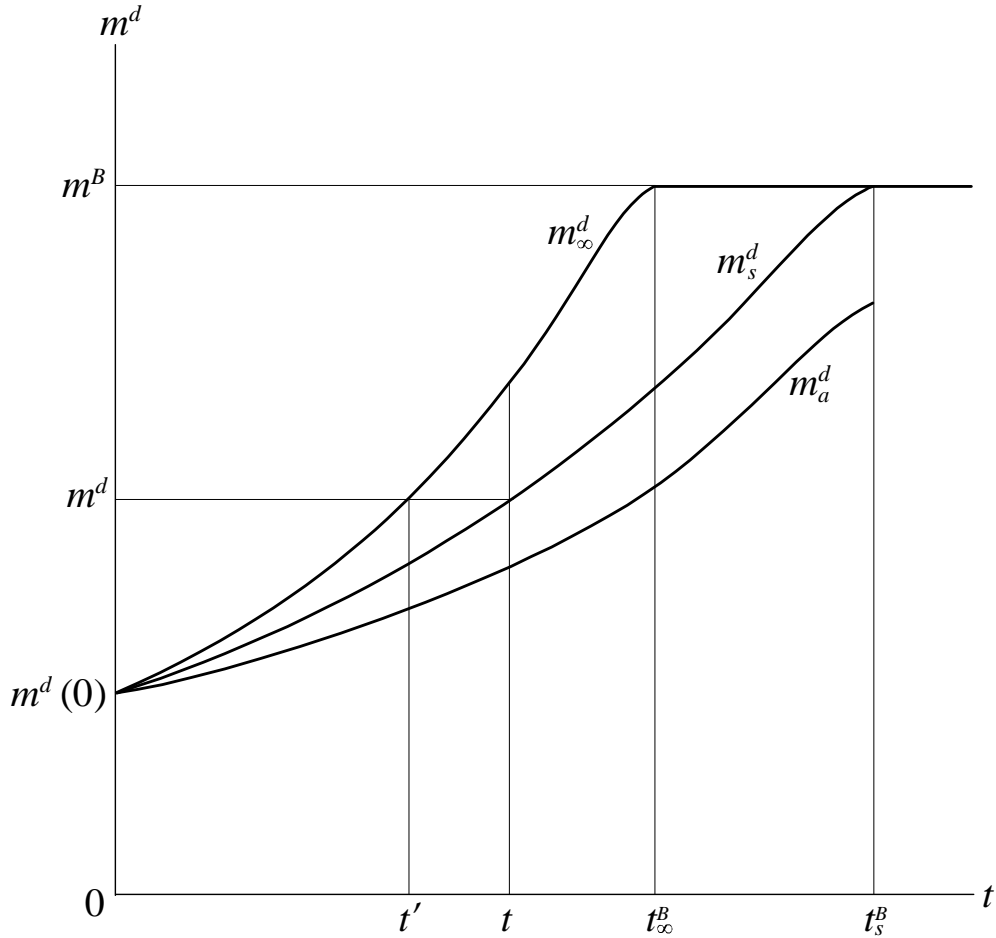


Figure 5: The path m_∞^d dominates the paths m_s^d and m_a^d .