Knowledge Creation through Multimodal Communication*

Marcus Berliant† and Masahisa Fujita‡

April 21, 2024

Abstract

Knowledge creation either in isolation or joint with another person, using either face to face or internet contact and incorporating internet search ability is analyzed. Both a conceptual phase and a technical phase of research are analyzed, allowing workers to choose endogenously their mode of communication. In addition to formal knowledge, tacit knowledge plays an essential role in the knowledge production process, as it is not internalized. Lead time for face to face communication plays a key role. The sink point is inefficient. Our framework is applied to pandemic restrictions on face to face communication.

JEL codes: D83, L86 Keywords: Knowledge creation; Tacit knowledge; Multimodal communication; Pandemic restrictions

*We thank the Whiteley Center in Friday Harbor at the University of Washington for their kind hospitality during our work on this paper. The second author is grateful for Grants Aid for Scientific Research Grant 23K01348 from the Japanese Ministry of Education and Science. We are grateful to two anonymous referees, Toshitaka Gokan, Nobuaki Hamaguchi, David Nagy, Kentaro Nakajima, and Tony Smith for helpful comments. The authors retain responsibility for any errors.

†Department of Economics, Washington University, MSC 1208-228-308, 1 Brookings Drive, St. Louis, MO 63130-4899 USA. Fax: (1-314) 935-4156, e-mail: berliant@wustl.edu

‡Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-01 Japan. Phone · Fax: (81-75) 753-7120, e-mail: fujita@kier.kyoto-u.ac.jp

1 Introduction

1.1 Motivation

How does the knowledge creation process function when there are multiple ways for people to communicate, for example face to face or using the internet? Relative to the one communication channel case, what different patterns of joint research among knowledge workers emerge, and how is the productivity of research work affected? Under what conditions are the conceptual and technical phases of knowledge production best accomplished through each of the communication channels?

When analyzing the dynamics of knowledge creation, as we do here, the concept of *tacit knowledge* among the people creating new knowledge arises organically. As tacit knowledge is the part of knowledge that is not embodied in the final product, readers will recognize it as what is learned by authors in the academic research setting that is not embedded in a published paper, including this one. What are the efficiency consequences of tacit knowledge? What kind of innovation policy is appropriate in the presence of tacit knowledge? Can artificial intelligence have tacit knowledge? To address these questions, we must first build a model that incorporates tacit knowledge. We are not aware of any formal models of either multimodal communication or tacit knowledge in the prior literature.

1.2 Preview of the results

To address our motivating questions, as depicted in Figure 1, a person, say i, with knowledge K_i can develop new ideas in isolation while interacting with *the Server*. Or person i may create new ideas jointly with another person, say j, by interacting through *the Net* or by working *F2F (Face to Face)*. Using this extension of our earlier models, that are discussed below, we can examine the impact of rapidly developing ICT (including AI) on knowledge creation activities.

Figure 1

This paper also aims to provide a theoretical framework for the study of specific recent issues such as the impact of the Covid-19 pandemic on knowledge creation, or the effect of urban structure on the productivity of knowledge workers as well as on the pattern of knowledge work in large cities.

Our main findings are as follows. In contrast with our previous work, where tacit knowledge was absent, the incorporation of tacit knowledge explicitly into the framework changes the results. In particular, knowledge is generated not only by patent output but also through search and independent thinking in preparation for joint or independent research. This represents tacit knowledge retained by a specific worker for future work. It slows the accumulation of knowledge in common. Our conclusions are threefold. First, the dynamics of the system with two workers and tacit knowledge are, in the end, very different from the dynamics without tacit knowledge. Second, the steady state will not, in general, be the state with highest productivity. The net effect is that achieving and maintaining the highest productivity profile of knowledge in common and differential knowledge requires more heterogeneity or larger research groups than we found in our previous work. Third, the effect of tacit knowledge on knowledge productivity is not internalized by the knowledge workers. We shall provide more discussion of these points in the conclusions.

Applying this framework to pandemic restrictions, we show, for example, how the productivity of knowledge workers with longer commutes to work is affected less than those with shorter commutes when pandemic restrictions on face to face work are implemented. This application requires the introduction of multimodal communication to our model.

1.3 Related literature

Through a sequence of four related papers, Berliant and Fujita (2008, 2009, 2011, 2012), we have developed a model of knowledge creation based on the interactions among a group of heterogeneous people. These papers form the basis for the analysis here. This paper is at a more micro level than the earlier work, necessitated by addressing questions concerning multiple modes of communication and tacit knowledge during the knowledge creation process. In the earlier work, we considered knowledge creation to be opening up boxes, each containing one idea, either alone or with another person. Opening up a box and learning the content, in other words creating a new idea, takes time, considered to be an opportunity cost. In contrast, here we model, to an extent, the process of opening a box.[1] Analogous to producing a car, we model the construction of the various components of an idea, including both

[1] In ongoing work, we tackle the completely microfounded process of knowledge creation using artificial intelligence.

the conceptual and technical aspects, and then their assembly into a final product. For each phase, agents optimize over the type of communication medium used for joint work. The construction of components generates tacit knowledge.

In Berliant and Fujita (2008), we develop the basic model and analyze dynamic interactions among a group of knowledge workers under the assumption of symmetry of the state of knowledge. Berliant and Fujita (2009) investigates further the case of two knowledge workers, relaxing the symmetry assumption and allowing knowledge transfer in addition to knowledge creation, whereas Berliant and Fujita (2011) embeds the basic model in a growth framework to analyze macroeconomic dynamics and the efficiency properties of equilibrium. Finally, Berliant and Fujita (2012) constructs a two region version of the model to examine the emergence of cultures in the knowledge production community. All of these papers feature only one channel of communication, whereas here we consider both face to face as well as electronic communication. None of these papers touch on tacit knowledge.

Part of our framework here is based on the insightful empirical paper by Lin et al (2022). They break down the knowledge creation process into *conceptual* and *technical* phases. The early conceptual phase involves tacit knowledge deployment, whereas the later technical phase involves explicit knowledge. Using a large data set, they find that face to face communication is more effective in the conceptual phase, whereas remote teams can be effective in the technical phase.[2] Our model allows agents to select endogenously the mode of communication they employ in each phase.

From a wider perspective, in related work, Aghion et al. (2017) address the interesting ways in which artificial intelligence can impact economic growth at the macro level using a Cobb-Douglas production technology. In contrast, the work here addresses how internet communication and search affects knowledge creation at a very micro level. Krugman (1991) points out that face to face contact can promote knowledge externalities between agents. Atkin et al (2022) verify this empirically. Ceci et al (2020) conduct a case study in the aerospace industry, finding that both face to face and electronic communication are important to knowledge creation in a company, but they are used differentially by employees depending on both the relationship between communicators and the activity. Panahi et al. (2012) hypothesize that social media are a channel for tacit knowledge sharing.

[2] Finke et al (1992) present a similar abstract framework called Geneplore.

The interesting work of Bloom et al (2020) looks at idea creation from an endogenous growth perspective. They find that research productivity in various sectors as well as the aggregate economy has been declining, since ideas are becoming harder to find. The phenomenon has been offset by an increasing number of researchers and research effort.[3] Tacit knowledge could play a big role here, as it is not part of the empirical productivity accounting or theoretical models.

1.4 Outline of the paper

Our analysis proceeds as follows. In section 2, we develop our model of two knowledge workers using multiple modes of communication and generating tacit knowledge in addition to patents, and analyze the steady state. In section 3, we examine the comparison static effect of lead time cost on the optimal mode of communication. Section 4 considers knowledge growth under symmetry, whereas section 5 introduces tacit knowledge and analyzes dynamics in the two person case. Section 6 presents our conclusions and suggestions for future research. The three appendices provide extensions and technical details, including an application of our model to pandemic restrictions in Appendix A.

2 The model with two persons in the stationary state

In this section, we consider two persons/researchers, i and j, and extend the model of Berliant and Fujita (2008, 2009) by incorporating multiple modes of interaction. Wherever possible, we use the same notation as in Berliant and Fujita (2008, 2009).

Following on our earlier work, we model knowledge creation as a process of opening up boxes containing ideas. The labels on the boxes, that describe their contents, are known to all, but it takes time to understand the *contents* of the boxes. An example of such a box of knowledge is the creation of this paper by its authors. The title is its label. Another example is a new recipe for curry rice. So the discovery of a new idea is analogous to opening a box that contains its details.

[3]They also note the public good nature of ideas, as we do in Berliant and Fujita (2011), that can overcome decreasing returns and imply exponential growth in per capita income.

There is a countable infinity of boxes, each with its own unique label. They are numbered $k = 1, 2, ...$ so that the state of a knowledge worker at time $t \in [0, \infty)$ is given by a (countably infinite) vector of zeroes and ones, where a zero represents an unopened box to *that person*, and a one represents an opened box, in other words an idea that has been discovered by that person. At any finite time, the number of ones in the vector is finite. Moreover, ideas that have been discovered by others but not by a researcher are assigned zero to that researcher.[4] We model potential states of knowledge, represented by this vector, as infinite, since the potential for new discoveries is unbounded. The knowledge state of each person at a given time is thus represented by a vertex of the Hilbert cube.

The rate at which the boxes can be opened depends on the stock already opened by a particular person, either alone or with someone else. When working in isolation, the total stock of knowledge or boxes already opened by that person affects the rate at which new boxes are opened. When working jointly, both the total number of boxes opened previously and their profile matter. Whether they were opened together, and thus become mutual knowledge, or independently, and thus become exclusive knowledge, determines the relative heterogeneity of agents and the productivity of joint work. Notice that the heterogeneity of knowledge workers is endogenous to the model.

As there is an infinite number of boxes or potential ideas, we assume that the probability that knowledge workers who are not working together open the same box is zero.

In what follows, in contrast with our earlier work, we allow more channels of communication between knowledge workers, namely face to face and internet communication. Moreover, we decompose knowledge creation into more elementary units or phases, to be described formally in this section. These phases involve internet search and thinking on one's own, both when creating knowledge alone and when preparing to work with someone else. When working with someone else, there will also be time spent communicating either face to face or over the internet. We allow each person to optimize over the allocation of time or frequency to these various activities. For example, when creating a new recipe for curry rice, a chef might search the internet for recipes good and bad (respecting reviews on line), and then might engage in trial and error. The allocation of time is chosen by the chef. Much of what the chef

[4] In other work, for example Berliant and Fujita (2011), we have examined public knowledge transmission.

learns becomes tacit knowledge *beyond* the final recipe; this final recipe might be secret and used by his restaurant, or sold. The rate at which new recipes are created will depend on the chef's experience with this type of knowledge work, but also on *the tacit knowledge accumulated through the creation process.*

To elaborate our model, let us consider a specific time, $t \in [0, \infty)$, and let the following variables represent the state of each person's knowledge at time t (whenever clear, dropping t for simplicity):

n_k: the size of person k's knowledge (or number of ideas known by person k at time t); $k = i, j$

$n_{ij}^c \equiv n_{ji}^c$: the size of knowledge that i and j both know, or the *common knowledge* for i and j

n_{ij}^d: the size of knowledge known by i but not known by j, or the *differential knowledge* of i from j,

n_{ji}^d: the size of knowledge known by j but not known by i, or the *differential knowledge* of j from i.

By definition,

$$n_i = n_{ij}^c + n_{ij}^d, \; n_j = n_{ij}^c + n_{ji}^d.$$

Let

$$n^{ij} \equiv n_{ij}^c + n_{ij}^d + n_{ji}^d = n_i + n_j - n_{ij}^c$$

be the size of total knowledge that is known either by i or j.

Next, we define the proportion of each type of knowledge in the total size of knowledge n^{ij}:

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

implying that

$$m_{ij}^c + m_{ij}^d + m_{ji}^d = 1, \tag{1}$$

and hence

$$n_i = n^{ij} \cdot (1 - m_{ji}^d), \; n_j = n^{ij} \cdot (1 - m_{ij}^d), \tag{2}$$

or

$$\frac{n_i}{n^{ij}} = 1 - m_{ji}^d, \; \frac{n_j}{n^{ij}} = 1 - m_{ij}^d.$$

Using this notation, we describe next the two alternative ways of creating knowledge (at time t).

In what follows, consistent with the notation introduced above, lower case letters such as i and j represent persons, whereas upper case letters represent activities. Examples of the latter include I, representing *isolated* or *independent* activity, and J, representing *joint* activity or activity for the purpose of *joint* knowledge creation. Likewise, S represents *search* activity, to be explained next.

What we mean by *search* is to search the web by oneself to prepare for knowledge creation activity that will occur either in isolation or jointly with another. Examples are using Google or ChatGPT. The activity is a form of directed search, in contrast with undirected search. That is, the search terms or questions guide the use of the web in an important and nonrandom way, and are progressively refined over the time used for search. Just a few years ago, when beginning a project, an economics researcher would search the Econlit database, for example, using search terms that define the new project. The purpose would be to find related papers at the frontier of knowledge and to see similar work in terms of assumptions, implications, models, and empirics. For example, aside from the references we knew about from previous joint work, to compose this paper we searched on key phrases such as "R&D during the Covid 19 pandemic." Refinement of search terms, as well as digesting material, takes time and effort. Nowadays, with Google Scholar and ChatGPT, the effectiveness of search has improved dramatically over primitive times. We will parameterize this effectiveness of directed search in our model.

Given two persons i and j, as noted previously, a person, say i, can develop new knowledge in alternatively ways:

(i) knowledge creation in isolation

(ii) joint knowledge creation of i and j working together

Next, we describe in detail each type of knowledge creation, starting with the simpler one.

(i) Knowledge creation in isolation

Figure 2 depicts the *activity tree* for knowledge creation by person i in isolation.

Figure 2

The top node $\square$ in Figure 2 represents the final output A^I_{ii}, which is produced by appropriately combining the outputs of the two basic activities in the second tier: a_i^{IT} (the intermediate output of *Independent Thinking* by i), and a_i^{IS} (the intermediate output of *Independent Search* by i). The output, or the number of new intermediate ideas created per unit of time by each basic activity is governed by the following equations:

$$a_i^{IT} = \alpha_{IT} \cdot n_i, \quad (\alpha_{IT} > 0) \tag{3}$$

$$a_i^{IS} = \alpha_{IS} \cdot n_i, \quad (\alpha_{IS} > 0) \tag{4}$$

where the positive constant α_{IT} represents the effectiveness of thinking alone, whereas the positive constant α_{IS} represents the effectiveness of search. In both basic activities, productivity depends on what a person already knows.[5]

The final output of knowledge creation activity by person i in isolation per unit of time is governed by the following equation:

$$A_{ii}^{I} = \alpha_I \cdot \left[\omega_{IT} \cdot a_i^{IT}\right]^{\rho_{IT}} \cdot \left[\omega_{IS} \cdot a_i^{IS}\right]^{\rho_{IS}} \quad (\alpha_I > 0) \tag{5}$$

where A_{ii}^{I} is the number of ideas produced.[6] On the right hand side, $\rho_{IS}, \rho_{IT} \geq 0$ are fixed parameters that weight the search and thought activities; we assume that $\rho_{IS} + \rho_{IT} = 1$. Knowledge creation activity in a time period is divided into two parts: search in isolation with frequency $\omega_{IS} \geq 0$ and thinking in isolation with frequency $\omega_{IT} \geq 0$, where $\omega_{IS} + \omega_{IT} = 1$. The positive constant α_I represents the overall productivity of knowledge creation in isolation. Notice that given the Cobb-Douglas functional form, individual thinking and individual search are not perfect substitutes.

Knowledge worker i optimizes the choice of frequency over the two activities:

$$\max \left\{ A_{ii}^{I} \middle| \, \omega_{IT} + \omega_{IS} = 1, \omega_{IT} \geq 0, \omega_{IS} \geq 0 \right\}$$

yielding the optimal choices of frequencies ω_{IT}^{*} and ω_{IS}^{*}:

$$\omega_{IT}^{*} = \rho_{IT}, \omega_{IS}^{*} = \rho_{IS}.$$

Thus, the optimized value of knowledge output per unit of time for person i in isolation is:

$$\begin{aligned} A_{ii}^{I*} &= \alpha_I \cdot \left[\rho_{IT} \cdot a_i^{IT}\right]^{\rho_{IT}} \cdot \left[\rho_{IS} \cdot a_i^{IS}\right]^{\rho_{IS}} && (6) \\ &= \Phi_I \cdot n_i && (7) \end{aligned}$$

[5] "What you can learn from Wikipedia depends on what you already know."

[6] We note here, at the first introduction of a knowledge production function, that we use extensively the Cobb-Douglas functional form thoughout this paper. Although use of this particular function has drawbacks, for instance in generality of the results, its big advantage is that the functional form makes calculations vastly simpler than they would be otherwise, and permits analytical tractability.

where

$$\Phi_I \equiv \alpha_I \cdot (\alpha_{IT} \cdot \rho_{IT})^{\rho_{IT}} \cdot (\alpha_{IST} \cdot \rho_{IS})^{\rho_{IS}} .$$

Likewise, for person j, we have

$$A_{jj}^{I*} = \Phi_I \cdot n_j \tag{8}$$

where Φ_I represents the same function of parameters as above. We may note that the search productivity parameter α_{IS} depends on several factors. First, it depends on the search technology at the time the work is done. For example, Google (1998) preceded Google Scholar (2004) and more recently ChatGPT. Second, it depends on the knowledge stock at the time, for example in Wikipedia. Third, it depends on the learning capacity of person i. Similarly for the parameter ρ_{IS}. In the long run, parameters might change, but here, for simplicity, we take the parameters as fixed.

(ii) Joint knowledge creation of i and j through multimodal communication

Next, we turn to the main focus of this section, that is, the modeling of joint knowledge creation by two persons in communication through the Net and F2F. Our framework is motivated by the empirical work of Lin et al (2023). We divide the knowledge production process into *conceptual* and *technical* phases, allowing researchers to choose the best mode of communication, F2F or Net, for each phase. Lin et al (2023) find that remote collaboration is less effective in the conceptual phase of research, particularly where tacit knowledge is useful, whereas remote collaboration is used in the later stage of technical tasks. These will be outcomes of our model.

Figure 3 represents the *activity tree* for joint knowledge creation by two persons, i and j. In comparing figures 2 and 3, we can see that the process of joint knowledge production is much more complex than the case of knowledge creation in isolation.

Figure 3

The top ■ in Figure 3 represents the final output A_{ij}^J, which is obtained by appropriately combining the outputs of four main branches of activities. The left branch corresponds to the joint task of *conceptual development* for new patents (or papers). Each of the two branches in the middle of the second tier corresponds to the task of further development of intermediate ideas for the purpose of joint knowledge creation *independently* by each person. The right branch corresponds to the task of joint *technical development* of new patents (or papers).

The middle two branches of independent development are included to reflect the fact that when two persons are developing a new patent or paper, they spend time not only working jointly F2F or through the net, but each person also spends a significant amount of time working independently *for the purpose of joint development* of the final output. It also needs to be noted that over the time of actual development of a new patent or paper, the four branches of activities depicted in Figure 3 may be repeated sequentially many times.

Next, we describe each branch of activity in detail.

$(ii-1)$ *Joint conceptual development*

As shown in the third tier of the left branch in Figure 3, Conceptual development is based on four basic activities: a_{ij}^{CTF} (*Conceptual* development by *Thinking* together *F2F*), a_{ij}^{CTN} (*Conceptual* development by *Thinking* together through the *Net*), a_i^{CSI} (*Conceptual* development through *Search Independently* by person i), and a_j^{CSI} (*Conceptual* development through *Search Independently* by person j).

Each basic activity is governed by the following equations:

$$a_{ij}^{CTF} = \alpha_{CF} \cdot (n_{ij}^c)^{1-\theta_C} \cdot (n_{ij}^d \cdot n_{ji}^d)^{\frac{\theta_C}{2}}, \ (\alpha_{CF} > 0, \ 0 < \theta_C < 1) \quad (9)$$

$$a_{ij}^{CTN} = \alpha_{CN} \cdot (n_{ij}^c)^{1-\theta_C} \cdot (n_{ij}^d \cdot n_{ji}^d)^{\frac{\theta_C}{2}}, \ (\alpha_{CN} > 0) \quad (10)$$

$$a_i^{CSI} = \alpha_{CS} \cdot n_i, \qquad a_j^{CSI} = \alpha_{CS} \cdot n_j, (\alpha_{CS} > 0) \quad (11)$$

Equation (9) states that when two persons are thinking jointly for conceptual development F2F, the output a_{ij}^{CTF} of conceptual intermediate ideas is generated at a rate proportional to the normalized product of their knowledge in common, n_{ij}^c, the differential knowledge of i from j, n_{ij}^d, and the differential knowledge of j from i, n_{ji}^d. The rate of creation of new intermediate ideas is high when the proportions of knowledge in common, knowledge exclusive to person i, and knowledge exclusive to person j are in balance. The parameter θ_C represents the weight on differential knowledge as opposed to knowledge in common in the production of new ideas in the conceptual development phase. Knowledge in common is necessary for communication between the two persons, whereas knowledge exclusive to one person or the other implies more heterogeneity or originality in the collaboration. The positive constant α_{CF} represents the effectiveness of conceptual development by thinking jointly F2F.

Equation (10) gives the rate of creation of conceptual intermediate ideas thinking jointly through the Net. This equation is same as equation (9) except that effectiveness parameter α_{CF} is replaced with α_{CN}. Although the two equations are of the same form, the difference between the two parameters, α_{CF} and α_{CN}, plays the essential role in the analysis later.

Each of the two equations in (11) states that the rate of creation of conceptual intermediate ideas through search by each person is proportional to the size of their knowledge, where the effectiveness parameter α_{CS} is common for the two persons.

The outputs of the four basic activities at the bottom of the left branch are combined in two steps, yielding the output a_{ij}^{C} of conceptual development. First, intermediate conceptual ideas, a_{ij}^{CTF} and a_{ij}^{CTN}, generated through *thinking jointly* through F2F and Net are combined as follows:

$$a_{ij}^{CT} = \lambda_{CF} \cdot a_{ij}^{CTF} + \lambda_{CN} \cdot a_{ij}^{CTN}, \tag{12}$$

where λ_{CF} and λ_{CN} denote the frequency of basic activities a_{ij}^{CTF} and a_{ij}^{CTN}, respectively. The two persons can jointly choose λ_{CF} and λ_{CN} freely, subject to the following constraint:

$$(1 + \varepsilon_F) \cdot \lambda_{CF} + \lambda_{CN} = 1, \quad \lambda_{CF} \geq 0, \; \lambda_{CN} \geq 0 \tag{13}$$

where $\varepsilon_F > 0$ represents the *lead time* of joint thinking for conceptual development F2F. In practice, ε_F reflects the time cost of preparing for a F2F meeting, such as commuting time to the common CBD office (or common university), or travel time between two cities or two countries where each of the two persons reside separately.

Alternatively, we may generalize the constraint (13) as follows:

$$(1 + \varepsilon_F) \cdot \lambda_{CF} + (1 + \varepsilon_N) \cdot \lambda_{CN} = 1, \quad \lambda_{CF} \geq 0, \; \lambda_{CN} \geq 0,$$

where ε_N represents the lead time of joint thinking for conceptual development through the net. In practice, ε_N is much smaller than ε_F. Furthermore, to reflect the nature of *net-technology*, the value of ε_N is essentially independent of "geographical distance" between the pair of communicators. Hence, given that we focus in this paper on the obstacle to communication caused by "geographical distance" between the pair of communicators, and also for simplicity of notation, we set $\varepsilon_N = 0$ and use the time constraint given by (13).[7]

[7] Actually, if we denote the original parameter ε_F as ε_F^o, and define the parameter ε_F by

Based on the constraint (13), the two persons jointly choose the optimal combination of $\{\lambda_{CF}, \lambda_{CN}\}$ by solving the this problem:

$$\max\left\{ a_{ij}^{CT} \middle| (1+\varepsilon_F)\cdot\lambda_{CF}+\lambda_{CN}=1,\quad \lambda_{CF}\geq 0,\ \lambda_{CN}\geq 0\right\}$$

which yields the following solution:

$$\text{when } \frac{\alpha_{CF}}{\alpha_{CN}}\cdot\frac{1}{1+\varepsilon_F} > 1,\ \lambda_{CF}^*=\frac{1}{1+\varepsilon_F},\ \lambda_{CN}^*=0 \tag{14}$$

$$\text{when } \frac{\alpha_{CF}}{\alpha_{CN}}\cdot\frac{1}{1+\varepsilon_F} \leq 1,\ \lambda_{CF}^*=0,\ \lambda_{CN}^*=1 \tag{15}$$

Using (14) and (15), the optimized value of the output of joint conceptual thinking over a unit of time is given by

$$a_{ij}^{CT*}=\lambda_{CF}^*\cdot a_{ij}^{CTF}+\lambda_{CN}^*\cdot a_{ij}^{CTN}. \tag{16}$$

In the second step, the output of joint thinking, a_{ij}^{CT*}, and the output of independent search by each person, $a_i^{CSI}$and a_j^{CSI}, in (11) are combined to yield the final output a_{ij}^C of the *conceptual development phase* as follows:

$$a_{ij}^C=\alpha_{ij}^C\cdot[\omega_{CT}\cdot a_{ij}^{CT*}]^{\rho_{CT}}\cdot[\omega_{CS}\cdot(a_i^{CSI}+a_j^{CSI})]^{\rho_{CS}} \tag{17}$$

where the positive constant α_{ij}^C represents the overall effectiveness of conceptual development activities. Each of ρ_{CT} and ρ_{CS} represents a given positive weight on each type of basic activity where $\rho_{CT}+\rho_{CS}=1$. Each of $\omega_{CT}\geq 0$ and $\omega_{CS}\geq 0$ denotes the frequency of the corresponding activity over a unit of time, where $\omega_{CT}+\omega_{CS}=1$. The two persons choose jointly the optimal combination of $\{\omega_{CT}, \omega_{CS}\}$ by solving the next problem:

$$\max\{ a_{ij}^C \big| \omega_{CT}+\omega_{CS}=1,\ \omega_{CT}\geq 0,\ \omega_{CS}\geq 0\}$$

yielding the solution:

$$\omega_{CT}^*=\rho_{CT},\ \omega_{CS}^*=\rho_{CS}. \tag{18}$$

Substituting (16) and (18) into (17), the optimized value of the final output of conceptual development phase is obtained as follows:

$$\begin{aligned} a_{ij}^{C*} &= \alpha_{ij}^C\cdot[\rho_{CT}\cdot(\lambda_{CF}^*\cdot a_{ij}^{CTF}+\lambda_{CN}^*\cdot a_{ij}^{CTN})]^{\rho_{CT}}\cdot[\rho_{CS}\cdot(a_i^{CSI}+a_j^{CSI})]^{\rho_{CS}} && (19)\\ &= \alpha_{ij}^C\cdot[\rho_{CT}\cdot(\lambda_{CF}^*\cdot\alpha_{CF}+\lambda_{CN}^*\cdot\alpha_{CN})\cdot(n_{ij}^c)^{1-\theta_C}\cdot(n_{ij}^d\cdot n_{ji}^d)^{\frac{\theta_C}{2}}]^{\rho_{CT}} \\ &\quad \times[\rho_{CS}\cdot\alpha_{CS}\cdot(n_i+n_j)]^{\rho_{CS}}, && (20)\end{aligned}$$

the relation:

$$\frac{1+\varepsilon_N}{1+\varepsilon_F^o}=\frac{1}{1+\varepsilon_F},\ \text{that is, } \varepsilon_F\equiv\frac{\varepsilon_F^o-\varepsilon_N}{1+\varepsilon_N},$$

then we can see that there is no loss of generality in setting $\epsilon_N=0$.

where λ^*_{CF} and λ^*_{CN} are defined through (14) and (15).

(ii − 2) Independent development by each person for the purpose of joint creation

As noted before, when two persons are jointly creating a new patent or paper, each person spends a significant amount of time working independently to develop intermediate ideas for the final joint output. The modeling of this phase of independent development by each person is rather similar to *Knowledge creation in Isolation* explained in (i) using Figure 2. Here, focusing on the middle left branch of *Independent development by person i* in Figure 3, the output of each basic activity at the bottom of this branch is given by the following equations:

$$a_i^{ITJ} = \alpha_{ITJ} \cdot n_i \quad (\alpha_{ITJ} > 0) \tag{21}$$

$$a_i^{ISJ} = \alpha_{ISJ} \cdot n_i \quad (\alpha_{ISJ} > 0) \tag{22}$$

where a_i^{ITJ} represents the number of new basic ideas created per unit time through *Independent Thinking* by person i for *Joint* creation purposes, and a_i^{ISJ} represents that basic ideas created through *Independent Search* for *Joint* purposes by i. These outputs of basic activities are combined as follows for developing intermediate ideas:

$$a_{ii}^I = \alpha_{IJ} \cdot \left[\omega_{ITJ} \cdot a_i^{ITJ}\right]^{\rho_{ITJ}} \cdot \left[\omega_{ISJ} \cdot a_i^{ISJ}\right]^{\rho_{ISJ}} \tag{23}$$

where a_{ii}^I is the number of new intermediate ideas developed per unit time. Each multiplier, ρ_{ITJ} and ρ_{ISJ}, represents the fixed weight on the output of the corresponding basic activity, where $\rho_{ITJ} + \rho_{ISJ} = 1$. The frequency of each basic activity, ω_{ITJ} and ω_{ISJ}, is chosen optimally by person i as follows:

$$\max \left\{ a_{ii}^I \middle| \, \omega_{ITJ} + \omega_{ISJ} = 1, \; \omega_{ITJ} \geq 0, \omega_{ISJ} \geq 0 \right\},$$

yielding the optimal choice of frequencies ω^*_{ITJ} and ω^*_{ISJ}:

$$\omega^*_{ITJ} = \rho_{ITJ}, \; \omega^*_{ISJ} = \rho_{ISJ}. \tag{24}$$

Thus, the optimized value of intermediate output per unit of time for person i in this phase of Independent development is obtained by substituting (21), (22) and (24) into (23) as follows:

$$\begin{aligned} a_{ii}^{I*} &= \alpha_{IJ} \cdot \left[\rho_{ITJ} \cdot a_i^{ITJ}\right]^{\rho_{ITJ}} \cdot \left[\rho_{ISJ} \cdot a_i^{ISJ}\right]^{\rho_{ISJ}} && (25) \\ &= \Phi_{IJ} \cdot n_i && (26) \end{aligned}$$

where

$$\Phi_{IJ} \equiv \alpha_{IJ} \cdot (\rho_{ITJ} \cdot \alpha_{ITJ})^{\rho_{ITJ}} \cdot [\rho_{ISJ} \cdot \alpha_{ISJ}]^{\rho_{ISJ}}\,. \tag{27}$$

Likewise, for Independent development by person j in the middle right branch of Figure 3, we have

$$a_{jj}^{I*} = \Phi_{IJ} \cdot n_j \tag{28}$$

where Φ_{IJ} represents the same function of parameters as above.

(ii − 3) Joint technical development

As shown in the third tier of the right branch in Figure 3, *technical development* is based on four basic activities: $a_{ij}^{\tau TF}$ (*$\mathcal{T}$echnical* development by *Thinking* together *F2F*), $a_{ij}^{\tau TN}$ (*$\mathcal{T}$echnical* development by *Thinking* together through the *Net*), $a_i^{\tau SI}$ (*$\mathcal{T}$echnical* development through *Search Independently* by person i, and $a_j^{\tau SI}$ (*$\mathcal{T}$echnical* development through *Search Independently* by person j).

These basic activities in the right branch of Technical development are in parallel to those in the left branch of Conceptual development. Furthermore, the form of each basic activity in Technical development is essentially the same as the corresponding one in Conceptual development, where each upper case letter C (indicating Conceptual development) is replaced with τ (indicating $\mathcal{T}$echnical development). Hence, the explanation of each specification below shall be short.

The basic activities in Technical Development are governed by the following equations:

$$\begin{aligned}
a_{ij}^{\tau TF} &= \alpha_{\tau F} \cdot (n_{ij}^c)^{1-\theta_\tau} \cdot (n_{ij}^d \cdot n_{ji}^d)^{\frac{\theta_\tau}{2}},\ (\alpha_{\tau F} > 0,\ 0 < \theta_\tau < 1) && (29)\\
a_{ij}^{\tau TN} &= \alpha_{\tau N} \cdot (n_{ij}^c)^{1-\theta_\tau} \cdot (n_{ij}^d \cdot n_{ji}^d)^{\frac{\theta_\tau}{2}},\ (\alpha_{\tau N} > 0) && (30)\\
a_i^{\tau SI} &= \alpha_{\tau S} \cdot n_i,\ a_j^{\tau SI} = \alpha_{\tau S} \cdot n_j,\ (\alpha_{\tau S} > 0) && (31)
\end{aligned}$$

Although the three equations above are of the same form as, respectively, (9), (10) and (11), differences in parameters $\alpha_{\tau F}$ and $\alpha_{\tau N}$ from parameters α_{CF} and α_{CN} in (9) and (10) play an essential role in the choice of communication mode later. In (29) and (30), the parameter θ_τ represents the weight on differential knowledge as opposed to knowledge in common in joint Technical development. This parallels the parameter θ_C.

To get the final output of the Technical development phase, first the intermediate concept ideas, $a_{ij}^{\tau TF}$ and $a_{ij}^{\tau TN}$, are combined as follows:

$$a_{ij}^{\tau T} = \lambda_{\tau F} \cdot a_{ij}^{\tau TF} + \lambda_{\tau N} \cdot a_{ij}^{\tau TN} \tag{32}$$

The two persons can jointly choose the frequencies $\lambda_{\tau F}$ and $\lambda_{\tau N}$ freely, subject to the following constraint:

$$(1+\varepsilon_F)\cdot\lambda_{\tau F}+\lambda_{\tau N}=1,\ \lambda_{\tau F}\geq 0,\ \lambda_{\tau N}\geq 0 \tag{33}$$

where ε_F is the same lead time parameter for F2F meetings introduced in (13). The optimal combination of $\{\lambda_{\tau F},\lambda_{\tau N}\}$ is obtained by solving the following problem:

$$\max\left\{a_{ij}^{\tau T}\,\middle|\,(1+\varepsilon_F)\cdot\lambda_{\tau F}+\lambda_{\tau N}=1,\ \lambda_{\tau F}\geq 0,\ \lambda_{\tau N}\geq 0\right\}$$

which yields the solution:

$$\text{when } \frac{\alpha_{\tau F}}{\alpha_{\tau N}}\cdot\frac{1}{1+\varepsilon_F} > 1,\quad \lambda_{\tau F}^*=\frac{1}{1+\varepsilon_F},\ \lambda_{\tau N}^*=0 \tag{34}$$

$$\text{when } \frac{\alpha_{\tau F}}{\alpha_{\tau N}}\cdot\frac{1}{1+\varepsilon_F} \leq 1,\quad \lambda_{\tau F}^*=0,\ \lambda_{\tau N}^*=1 \tag{35}$$

Using (34) and (35), the optimized value of joint technical thinking over a unit of time is given by

$$a_{ij}^{\tau T*}=\lambda_{\tau F}^*\cdot a_{ij}^{\tau TF}+\lambda_{\tau N}^*\cdot a_{ij}^{\tau TN}. \tag{36}$$

Next, the output of joint thinking, $a_{ij}^{\tau T*}$, and the output of independent search by each person, $a_i^{\tau SI}$ and $a_j^{\tau SI}$, in (31) are combined to yield the final output of the *technical development phase* as follows:

$$a_{ij}^{\tau}=\alpha_{ij}^{\tau}\cdot[\omega_{\tau T}\cdot a_{ij}^{\tau T*}]^{\rho_{\tau T}}\cdot[\omega_{\tau S}\cdot(a_i^{\tau SI}+a_j^{\tau SI})]^{\rho_{\tau S}} \tag{37}$$

The two persons jointly choose the optimal combination of frequencies, $\{\omega_{\tau T},\omega_{\tau S}\}$, by solving the following problem:

$$\max\{a_{ij}^{\tau}\,|\,\omega_{\tau T}+\omega_{\tau S}=1,\ \omega_{\tau T}\geq 0,\ \omega_{\tau S}\geq 0\}$$

yielding the solution:

$$\omega_{\tau T}^*=\rho_{\tau T},\ \omega_{\tau S}^*=\rho_{\tau S}. \tag{38}$$

Substituting (36) and (38) into (37), we obtain the optimized value of the final output of the technical development phase as follows:

$$a_{ij}^{\tau*} = \alpha_{ij}^{\tau}\cdot[\rho_{\tau T}\cdot(\lambda_{\tau F}^*\cdot a_{ij}^{\tau TF}+\lambda_{\tau N}^*\cdot a_{ij}^{\tau TN})]^{\rho_{\tau T}}\cdot[\rho_{\tau S}\cdot(a_i^{\tau SI}+a_j^{\tau SI})]^{\rho_{\tau S}} \tag{39}$$

$$= \alpha_{ij}^{\tau}\cdot[\rho_{\tau T}\cdot(\lambda_{\tau F}^*\cdot\alpha_{\tau F}+\lambda_{\tau N}^*\cdot\alpha_{\tau N})\cdot(n_{ij}^c)^{1-\theta_\tau}\cdot(n_{ij}^d\cdot n_{ji}^d)^{\frac{\theta_\tau}{2}}]^{\rho_{\tau T}} \times[\rho_{\tau S}\cdot\alpha_{\tau S}\cdot(n_i+n_j)]^{\rho_{\tau S}}, \tag{40}$$

where $\lambda^*_{\tau F}$ and $\lambda^*_{\tau N}$ are defined through (34) and (35).

(ii − 4) The final output of joint work

As indicated in Figure 3, the final output A^J_{ij} of joint work for knowledge creation is generated by combining the outputs of the four branches of activities, (20), (25), (39) and (40), as follows:

$$A^J_{ij} = [\mu_C \cdot a^{C*}_{ij}]^{\rho_C} \cdot [\mu_I \cdot (a^{I*}_{ii} + a^{I*}_{jj})]^{\rho_I} \cdot [\mu_\tau \cdot a^{\tau *}_{ij}]^{\rho_\tau} \tag{41}$$

where each of μ_C, μ_I and μ_N represents a given positive weight on each task of branch-activity such that

$$\mu_C + \mu_I + \mu_\tau = 1,$$

whereas each of μ_C, μ_I and μ_τ denotes the frequency of each branch-activity over a unit of time. The two persons can jointly choose the optimal combination of $\{\mu_C, \mu_I, \mu_\tau\}$ by solving the following problem:

$$\max\{ A^J_{ij} \big| \, \mu_C + \mu_I + \mu_\tau = 1, \ \mu_C \geq 0, \ \mu_I \geq 0, \ \mu_\tau \geq 0\},$$

which yields the following solution:

$$\mu^*_C = \rho_C, \ \mu^*_I = \rho_I, \ \mu^*_\tau = \rho_\tau. \tag{42}$$

Substituting (42) into (41), the maximized value of A^J_{ij} is given by

$$A^{J*}_{ij} = [\rho_C \cdot a^{C*}_{ij}]^{\rho_C} \cdot [\rho_I \cdot (a^{I*}_{ii} + a^{I*}_{jj})]^{\rho_I} \cdot [\rho_\tau \cdot a^{\tau *}_{ij}]^{\rho_\tau} \tag{43}$$

To be more specific, substitution of (20), (26), (28) and (40) into (43) leads to

$$\begin{aligned} A^{J*}_{ij} &= \Phi_J \cdot [n_i + n_j]^{\rho_I + \rho_{CT} \cdot \rho_C + \rho_{\tau T} \cdot \rho_\tau} \\ &\times \left[(n^c_{ij})^{1-\theta_C} \cdot (n^d_{ij} \cdot n^d_{ji})^{\frac{\theta_C}{2}} \right]^{\rho_{CT} \cdot \rho_C} \cdot \left[(n^c_{ij})^{1-\theta_\tau} \cdot (n^d_{ij} \cdot n^d_{ji})^{\frac{\theta_\tau}{2}} \right]^{\rho_{\tau T} \cdot \rho_\tau} \\ &\times [\lambda^*_{CF} \cdot \alpha_{CF} + \lambda^*_{CN} \cdot \alpha_{CN}]^{\rho_{CT} \cdot \rho_C} \cdot [\lambda^*_{\tau F} \cdot \alpha_{\tau F} + \lambda^*_{\tau N} \cdot \alpha_{\tau N}]^{\rho_{\tau T} \cdot \rho_\tau} \end{aligned} \tag{44}$$

where λ^*_{CF}, λ^*_{CN}, $\lambda^*_{\tau F}$ and $\lambda^*_{\tau N}$ are defined respectively through (14), (15), (34) and (35), whereas Φ_J represents a function of parameters given by

$$\begin{aligned} \Phi_J &= (\rho_C \cdot \alpha^C_{ij})^{\rho_C} \cdot {\rho_I}^{\rho_I} \cdot \alpha^{\rho_I}_{IJ} \cdot (\rho_{ITJ} \cdot \alpha_{ITJ})^{\rho_{ITJ} \cdot \rho_I} \cdot (\rho_{ISJ} \cdot \alpha_{ISJ})^{\rho_{ISJ} \cdot \rho_I} \\ &\times (\rho_\tau \cdot \alpha^\tau_{ij})^{\rho_\tau} \cdot (\rho_{CT})^{\rho_{CT} \cdot \rho_C} \cdot (\rho_{CS} \cdot \alpha_{CS})^{\rho_{CS} \cdot \rho_C} \cdot (\rho_{\tau S} \cdot \alpha_{\tau S})^{\rho_{\tau S} \cdot \rho_\tau} \end{aligned} \tag{45}$$

Finally, recalling the definitions in the first part of this section, we can rewrite (44) in terms of m^c_{ij}, m^d_{ij} and m^d_{ji} as follows:

$$\begin{aligned} A^{J*}_{ij} &= n^{ij} \cdot \Phi_J \cdot \left[(1 - m^d_{ji}) + (1 - m^d_{ij}) \right]^{\rho_I + \rho_{CT} \cdot \rho_C + \rho_{\tau T} \cdot \rho_\tau} \\ &\times \left[(m^c_{ij})^{1-\theta_C} \cdot (m^d_{ij} \cdot m^d_{ji})^{\frac{\theta_C}{2}} \right]^{\rho_{CT} \cdot \rho_C} \cdot \left[(m^c_{ij})^{1-\theta_\tau} \cdot (m^d_{ij} \cdot m^d_{ji})^{\frac{\theta_\tau}{2}} \right]^{\rho_{\tau T} \cdot \rho_\tau} \\ &\times [\lambda^*_{CF} \cdot \alpha_{CF} + \lambda^*_{CN} \cdot \alpha_{CN}]^{\rho_{CT} \cdot \rho_C} \cdot [\lambda^*_{\tau F} \cdot \alpha_{\tau F} + \lambda^*_{\tau N} \cdot \alpha_{\tau N}]^{\rho_{\tau T} \cdot \rho_\tau} \end{aligned} \tag{46}$$

where Φ_J is a function of parameters given by (45).

Our model of the knowledge creation process is analogous to the manufacturing of an automobile. The elementary components are first synthesized from raw materials. Then the elementary components are assembled into complex parts. Lastly, the final product is created from the parts. This is what the customers see as sold in the market.

3 Effect of F2F lead time on the mode of communication and joint knowledge productivity

In this section, using the model developed in Section 2 (ii), we examine how the communication mode and joint knowledge productivity are affected by a change in the value of ε_F. Recall that the parameter ε_F represents the lead time necessary for realizing actual F2F communication by a pair of knowledge workers. In the context of intra-city communication, ε_F reflects the commuting time to the common workplace such as a CBD office, research institute or university. In the context of inter-city or inter-country communication, ε_F reflects the travel time between two cities or countries for realizing research work in F2F mode. Here, travel by each researcher between two cities or countries is assumed to take place with equal frequency, so cost is symmetric.

Recall parameter α_{CF} introduced in equation (9), which represents the effectiveness of *conceptual* development by thinking jointly *F2F*, and parameter α_{CN} in (10), representing the effectiveness of *conceptual* development by thinking jointly through the *Net*. Recall also parameters $\alpha_{\tau F}$ and $\alpha_{\tau N}$ introduced respectively in (29) and (30), representing respectively the effectiveness of *technical* development by thinking jointly *F2F* and through the *Net*. In the following analysis, we assume the following relations hold among these four parameters:

Assumption 1. The following relations hold among communication parameters $\{\alpha_{CF}, \alpha_{CN}, \alpha_{\tau F}, \alpha_{\tau N}\}$:

$$\frac{\alpha_{CF}}{\alpha_{CN}} > 1, \ \frac{\alpha_{\tau F}}{\alpha_{\tau N}} > 1, \ \text{and} \ \frac{\alpha_{CF}}{\alpha_{CN}} > \frac{\alpha_{\tau F}}{\alpha_{\tau N}}. \tag{47}$$

That is, in both the phase of conceptual development and the phase of technical development, thinking jointly F2F is more effective than thinking jointly through the Net. Furthermore, the relative effectiveness of F2F relative to the Net is greater in conceptual development than in technical development.

This assumption is based on two ideas. First, tacit knowledge is best developed face to face. Nonaka and Takeuchi (1995) provide case studies, including the invention of the first automatic home bakery machine by Matsushita and the development of the City car model by Honda. Both involved the exploitation of tacit knowledge. Second, empirical evidence developed by Lin et al (2022) is consistent with this assumption.

Based on Assumption 1, Figure 4 explains the endogenous formation of three ranges of communication mode, (a), (b) and (c), depending on the parameter value of ε_F when $\alpha_{CF}/\alpha_{CN} = 2$ and $\alpha_{\tau F}/\alpha_{\tau N} = 1.5$.

Figure 4

In this figure, the horizontal axis displays values of the parameter ε_F, whereas the vertical axis gives the productivity of F2F interaction relative to Net interaction. The upper curve shows how the relative effectiveness of F2F relative to the Net in *conceptual* development changes with parameter ε_F; the lower curve shows the same for relative effectiveness in *technical* development. Let ε_F^C (respectively ε_F^τ) represent the value of ε_F where the upper curve (respectively the lower curve) meets the horizontal line with height 1. That is, ε_F^C and ε_F^τ are values of ε_F such that

$$\frac{\alpha_{CF}}{\alpha_{CN}} \cdot \frac{1}{1+\varepsilon_F^C} = 1, \text{ i.e., } \varepsilon_F^C = \frac{\alpha_{CF}}{\alpha_{CN}} - 1, \quad (48)$$

$$\frac{\alpha_{\tau F}}{\alpha_{\tau N}} \cdot \frac{1}{1+\varepsilon_F^\tau} = 1, \text{ i.e., } \varepsilon_F^\tau = \frac{\alpha_{\tau F}}{\alpha_{\tau N}} - 1. \quad (49)$$

Next, let (a), (b) and (c) represent respectively the three ranges of parameter value ε_F as shown in Figure 4. Then, recalling (14), (15), (34) and (35), and using Figure 4, we can readily obtain the following result:

Proposition 1. Under Assumption 1, let ε_F^C and ε_F^τ be the values of parameter ε_F defined by (48) and (49). Then, the combination of optimal frequencies, $\{\lambda_{CF}^*, \lambda_{CN}^*, \lambda_{\tau F}^*, \lambda_{\tau N}^*\}$, associated with the four basic joint-thinking activities, $\{a_{ij}^{CTF}, a_{ij}^{CTN}, a_{ij}^{\tau TF}, a_{ij}^{\tau TN}\}$, forms the following three ranges in terms of the parameter value of ε_F:

$$\text{(a) for } \varepsilon_F \in [0, \varepsilon_F^\tau)\,;\ \lambda_{CF}^* = \frac{1}{1+\varepsilon_F},\ \lambda_{CN}^* = 0,\ \lambda_{\tau F}^* = \frac{1}{1+\varepsilon_F},\ \lambda_{\tau N}^* = 0 \tag{50}$$

$$\text{(b) for } \varepsilon_F \in \left[\varepsilon_F^\tau, \varepsilon_F^C\right)\,;\ \lambda_{CF}^* = \frac{1}{1+\varepsilon_F},\ \lambda_{CN}^* = 0,\ \lambda_{\tau F}^* = 0,\ \lambda_{\tau N}^* = 1 \tag{51}$$

$$\text{(c) for } \varepsilon_F \in \left[\varepsilon_F^C, \infty\right)\,;\ \lambda_{CF}^* = 0,\ \lambda_{CN}^* = 1,\ \lambda_{\tau F}^* = 0,\ \lambda_{\tau N}^* = 1. \tag{52}$$

That is, within range (a) of Figure 4, in both the phase of conceptual development and the phase of technical development, joint work is conducted solely F2F, reflecting the high effectiveness of F2F communication and the small obstacle of F2F lead time. Within range (b), joint thinking for conceptual development is conducted F2F whereas joint thinking for technical development is conducted through the Net. This reflects the fact that the F2F lead time is a relatively small obstacle for conceptual joint work whereas it is relatively a heavy burden for technical joint work. In contrast, within range (c), both the phase of conceptual development and the phase of technical development are conducted solely through the net, reflecting the big obstacle of F2F lead time.

(Dear Marcus: Following each proposition or group of propositions, could you provide systematic discussions of relevant empirical feats / papers. In this first draft, I state only theoretical results.

Dear Masa: There isn't much directly related. There is only one paper that even divides work into conceptual and technical phases, and we cite it heavily. The other papers don't distinguish the phases. Most of the empirical literature is concerned with pandemic restrictions, the topic of the next subsection.)

Next, we examine how the productivity of joint work is affected by the parameter value ε_F. For this purpose, let us define

$$A_{ij}^{J*}(\varepsilon_F) \equiv \text{the value of } A_{ij}^{J*} \text{ at } \varepsilon_F \in [0, \infty)\,,$$

where A_{ij}^{J*} is the optimized output of joint work given by equation (44). The values of frequencies, λ_{CF}^*, λ_{CN}^*, $\lambda_{\tau F}^*$ and $\lambda_{\tau N}^*$, are respectively given by (50), (51) and (52), depending on the value of ε_F. Next, let

$$h(\varepsilon_F) \equiv \frac{A_{ij}^{J*}(\varepsilon_F)}{A_{ij}^{J*}(0)} = \frac{[\lambda_{CF}^* \cdot \alpha_{CF} + \lambda_{CN}^* \cdot \alpha_{CN}]^{\rho_{CT} \cdot \rho_C} \cdot [\lambda_{\tau F}^* \cdot \alpha_{\tau F} + \lambda_{\tau N}^* \cdot \alpha_{\tau N}]^{\rho_{\tau T} \cdot \rho_\tau}}{[\alpha_{CF}]^{\rho_{CT} \cdot \rho_C} \cdot [\alpha_{\tau F}]^{\rho_{\tau T} \cdot \rho_\tau}} \tag{53}$$

be the *relative productivity of joint work at* ε_F in comparison with the productivity of joint work at $\varepsilon_F = 0$, where the right side of equation (53) has been obtained from (50) since $\lambda^*_{CF} = 1$, $\lambda^*_{CN} = 0$, $\lambda^*_{\tau F} = 1$ and $\lambda^*_{\tau N} = 0$ at $\varepsilon_F = 0$.

Then, based on Proposition 1, we can readily obtain the relative productivity function, $h(\varepsilon_F)$, for each range of ε_F as follows:
(a) for $0 \leq \varepsilon_F < \varepsilon_F^\tau$,

$$h_a(\varepsilon_F) \equiv h(\varepsilon_F \,|0 \leq \varepsilon_F < \varepsilon_F^\tau) = (1+\varepsilon_F)^{-(\rho_{CT}\cdot\rho_C + \rho_{\tau T}\cdot\rho_\tau)}, \tag{54}$$

(b) for $\varepsilon_F^\tau \leq \varepsilon_F < \varepsilon_F^C$,

$$h_b(\varepsilon_F) \equiv h(\varepsilon_F \left|\varepsilon_F^\tau \leq \varepsilon_F < \varepsilon_F^C\right.) = (1+\varepsilon_F)^{-\rho_{CT}\cdot\rho_C} \cdot [\alpha_{\tau F}/\alpha_{\tau N}]^{-\rho_{\tau T}\cdot\rho_\tau}, \tag{55}$$

(c) for $\varepsilon_F^C \leq \varepsilon_F < \infty$,

$$h_c(\varepsilon_F) \equiv h(\varepsilon_F \left|\varepsilon_F^C \leq \varepsilon_F < \infty\right.) = [\alpha_{CF}/\alpha_{CN}]^{-\rho_{CT}\cdot\rho_C} \cdot [\alpha_{\tau F}/\alpha_{\tau N}]^{-\rho_{\tau T}\cdot\rho_\tau}, \tag{56}$$

Denoting by " $'$ " the derivative of a function with respect to parameter ε_F, it follows that:
(a) for $0 \leq \varepsilon_F < \varepsilon_F^\tau$,

$$\frac{h_a'(\varepsilon_F)}{h_a(\varepsilon_F)} = -\frac{\rho_{CT}\cdot\rho_C + \rho_{\tau T}\cdot\rho_\tau}{1+\varepsilon_F} < 0, \quad h_a''(\varepsilon) > 0, \tag{57}$$

(b) for $\varepsilon_F^\tau \leq \varepsilon_F < \varepsilon_F^C$,

$$\frac{h_b'(\varepsilon_F)}{h_b(\varepsilon_F)} = -\frac{\rho_{CT}\cdot\rho_C}{1+\varepsilon_F} < 0, \quad h_b''(\varepsilon) > 0, \tag{58}$$

(c) for $\varepsilon_F^C \leq \varepsilon_F < \infty$,

$$\frac{h_c'(\varepsilon_F)}{h_c(\varepsilon_F)} = 0, \tag{59}$$

implying that

$$-\frac{h_a'(\varepsilon_F)}{h_a(\varepsilon_F)} > -\frac{h_b'(\varepsilon_F)}{h_b(\varepsilon_F)} > \frac{h_c'(\varepsilon_F)}{h_c(\varepsilon_F)} = 0. \tag{60}$$

Figure 5 illustrates the *relative productivity curve*, $h(\varepsilon_F)$, over the three ranges of parameter ε_F when $\alpha_{CF}/\alpha_{CN} = 2$, $\alpha_{\tau F}/\alpha_{\tau N} = 1.5$ and $\rho_{CT}\cdot\rho_C = \rho_{\tau T}\cdot\rho_\tau = 0.25$.

Figure 5

As can been observed in the figure, in the (a) and (b) ranges, the relative productivity curve is decreasing and strictly convex, whereas it is flat over range (c). From Proposition 1, in range (a) joint work is conducted through F2F communication in both conceptual development and technical development, whereas in range (b) joint work is conducted through F2F interaction in the conceptual development phase, whereas joint work is conducted through the Net in the technical development phase. Moreover, increasing ε_F means a bigger obstacle for F2F communication, whereas *the cost of Net communication is independent of* ε_F. Hence, the relative productivity $h(\varepsilon_F)$ decreases with ε_F more sharply to the left of the boundary ε_F^τ than to the right of ε_F^τ, making the curve kinked at ε_F^τ. In contrast, in range (c), since joint work is conducted solely by using the Net, the relative productivity is not affected by the value of ε_F and hence it is flat; however, given that Net communication is less effective than F2F communication, the relative productivity is the lowest in range (c). The curve is also kinked at the boundary ε_F^C.

In Figure 5, the *broken curve*, $\tilde{h}_a(\varepsilon_F)$, to the right side of boundary ε_F^τ shows the *extension of curve* $h_a(\varepsilon_F)$ to the right of boundary ε_F^τ. That is,

$$\tilde{h}_a(\varepsilon_F) = (1+\varepsilon_F)^{-(\rho_{CT}\cdot\rho_C+\rho_{\tau T}\cdot\rho_\tau)}, \quad 0 \leq \varepsilon_F < \infty. \tag{61}$$

Recall that function $h_a(\varepsilon_F)$, has been derived by setting the communication mode so that the communication in joint work is conducted solely F2F in both the phase of conceptual development and the phase of technical development, without utilizing Net communication at all. In other words, the broken curve represents the relative productivity at each $\varepsilon_F \geq \varepsilon_F^\tau$ in *the fictitious situation when no net-technology was available.* In reality, however, in range (b), *optimal communication* requires use of the net, not F2F communication, in joint work for technical development. Hence, in range (b), the extended curve $\tilde{h}_a(\varepsilon_F)$ is based on non-optimal communication mode choice, and hence it is lower than the curve $h_b(\varepsilon_F)$. Finally, since the optimal communication mode in range (c) requires use of the net, not F2F communication, in both conceptual development and technical development, the extended productivity curve $\tilde{h}_a(\varepsilon)$ over range (c) is far lower than the curve $h_c(\varepsilon_F)$.

We may summarize the observations above as follows:

Proposition 2. Let $h(\varepsilon_F)$ be the relative productivity function defined by equation (53), which represents the relative productivity of joint work at each $\varepsilon_F \geq 0$ in comparison with that at $\varepsilon_F = 0$, in other words no lead time cost. Under Assumption 1, let (a), (b) and (c) represent the three ranges of ε_F as

defined in Figure 4. Then, the relative productivity function $h(\varepsilon_F)$ exhibits the following features, as illustrated in Figure 5:
(i) Over the ranges (a) and (b), the relative productivity curve is decreasing and strictly convex, and kinked at the boundary ε_F^τ. Given that in range (a) both conceptual joint work and technical joint work are conducted F2F, whereas in range (b) only conceptual joint work is conducted F2F, the relative productivity decreases with ε_F faster in range (a) than in range (b).
(ii) In range (c), since the joint work communication is conducted solely through the net, the relative productivity is insensitive to the value of ε_F, resulting in a flat curve. However, since no effective usage of F2F communication is possible in this range, the relative productivity is the lowest among the three ranges.
(iii) In Figure 5, the broken curve, $\tilde{h}_a(\varepsilon)$, represents the extension of the relative productivity curve $h_a(\varepsilon)$ to the right of boundary ε_F^τ, showing the relative productivity at each ε_F^τ under the *fictitious restriction* that no usage of the Net is permitted. The difference between the real relative productivity curve and this fictitious curve at each $\varepsilon_F > \varepsilon_F^\tau$ shows how much higher productivity in joint work is due to the effective use of net communication. Specifically, in range (c), although the broken curve decreases toward 0 as ε_F increases to ∞, the actual relative productivity $h_c(\varepsilon_F)$ stays relatively high due to the availability of net communication.

(Dear Marcus: You could try to make the statement of this proposition more compact.

Dear Masa: I have done so, but we might not want to make it more compact if we move most of the text preceding it to an appendix.)

Figure 5 can be interpreted as representing the impact of the development of net-communication technology, where the broken curve represents the relative productivity at each ε_F when *no net-technology is available*, whereas the real curve represents the relative productivity under the effective use of *modern net-technology*. The case where no net-technology is available is studied by Inoue et al (2022), since the internet didn't exist during the Spanish flu epidemic. Without this substitute for F2F, the effect on knowledge creation was severe.

In order to examine more closely the impact of the advances in net-communication technology in joint work, we add one more curve to Figure 5 and obtain Figure 6. In this new figure, the top solid curve and the bottom broken curve are respectively the same as those in Figure 5, where the horizontal axis in

Figure 6 is compressed to half of that in Figure 5. The new middle curve in Figure 6 represents the relative productivity curve when *old net-technology* was available. The middle curve has been obtained under the new set of communication-effectiveness parameters, $\{\alpha_{CF}, \alpha^o_{CN}, \alpha_{\tau F}, \alpha^o_{\tau N}\}$, reflecting the old net technology where the following assumption is satisfied:

Assumption 2. Let $\{\alpha_{CF}, \alpha_{CN}, \alpha_{\tau F}, \alpha_{\tau N}\}$ be the set of communication parameters used in Assumption 1, satisfying the relations in (47). Let $\{\alpha_{CF}, \alpha^o_{CN}, \alpha_{\tau F}, \alpha^o_{\tau N}\}$ be the communication parameters under the old Net technology, satisfying the following relations:

$$\frac{\alpha_{CF}}{\alpha^o_{CN}} > 1, \ \frac{\alpha_{\tau F}}{\alpha^o_{\tau N}} > 1, \text{ and } \frac{\alpha_{CF}}{\alpha^o_{CN}} > \frac{\alpha_{\tau F}}{\alpha^o_{\tau N}}. \tag{62}$$

and

$$\alpha_{CN} > \alpha^o_{CN}, \ \alpha_{\tau N} > \alpha^o_{\tau N}. \tag{63}$$

Next we explain this assumption. In the two sets of communication parameters above, α_{CF} and $\alpha_{\tau F}$ remain the same, reflecting that the effectiveness of F2F communication in joint work by itself is not affected by the development of Net technology. The conditions in (62) correspond to those in (47), whereas relations in (63) imply the *advancement* of net technology in conceptual development and in technical development. For both cases in Assumption 2, the weight parameters, $\rho_{CT} \cdot \rho_C$ and $\rho_{\tau T} \cdot \rho_\tau$ are assumed to remain the same.

Let ${}^o\varepsilon^C_F$ and ${}^o\varepsilon^\tau_F$ respectively be the values of ε_F such that

$$\frac{\alpha_{CF}}{\alpha^o_{CN}} \cdot \frac{1}{1 + {}^o\varepsilon^C_F} = 1, \text{ i.e., } {}^o\varepsilon^C_F = \frac{\alpha_{CF}}{\alpha^o_{CN}} - 1, \tag{64}$$

$$\frac{\alpha_{\tau F}}{\alpha^o_{\tau N}} \cdot \frac{1}{1 + {}^o\varepsilon^\tau_F} = 1, \text{ i.e., } {}^o\varepsilon^\tau_F = \frac{\alpha_{\tau F}}{\alpha^o_{\tau N}} - 1. \tag{65}$$

implying from (48), (49) and (63) that

$${}^o\varepsilon^C_F > \varepsilon^C_F, \ {}^o\varepsilon^\tau_F > \varepsilon^\tau_F. \tag{66}$$

The optimal communication frequencies $\{\lambda^*_{CF}, \lambda^*_{CN}, \lambda^*_{\tau F}, \lambda^*_{\tau N}\}$ under the old net technology for each of new ranges, (a^o), (b^o) and (c^o), are obtained by replacing ε^τ_F and ε^C_F by ${}^o\varepsilon^\tau_F$ and ${}^o\varepsilon^C_F$ respectively in (50), (51) and (52). The relative productivity function, $h^o(\varepsilon_F)$, for each range of ε_F is obtained by replacing ε^τ_F and ε^C_F by ${}^o\varepsilon^\tau_F$ and ${}^o\varepsilon^C_F$ respectively in (54), (55) and (56).

As illustrated in Figure 6, with the advance of net technology from the old to the modern, the *F2F-dominant range* (where F2F communication is exclusively used in both conceptual development and in technical development) *shrinks* from range (a^o) to range (a); the range of *mixed communication mode* (in which conceptual development is through F2F interaction whereas technical development is through the Net communication) also shrinks towards the origin from (b^o) to (b).[8] In contrast, the *net-dominant range expands towards the origin* from (c^o) to (c).

With such changes in the optimal communication mode with the advancement of Net technology, the productivity of joint work increases significantly in a wide range of parameter ε_F. To be precise, the relative productivity increases more steeply as parameter ε_F rises from ε_F^τ toward ${}^o\varepsilon_F^C$; and the relative proportion of productivity improvement reaches the maximum at ${}^o\varepsilon_F^C$, and it remains the same beyond ${}^o\varepsilon_F^C$.

We may summarize the observations above as follows:

Proposition 3. Suppose that the set of communication effectiveness parameters changes from the *old set* $\{\alpha_{CF}, \alpha_{CN}^o, \alpha_{\tau F}, \alpha_{\tau N}^o\}$ to the *new set* $\{\alpha_{CF}, \alpha_{CN}, \alpha_{\tau F}, \alpha_{\tau N}\}$, where all the conditions of Assumption 1 and Assumption 2 are satisfied; conditions (63) represent the advancement of Net technology. Then, as illustrated in Figure 6, the relative productivity curve under the new set of parameters shifts upward from that under the old set in a wide range of parameter ε_F. Specifically:
(i) With the advancement of Net technology, the three ranges of optimal communication mode change as follows: The *F2F-dominant range* shrinks toward the origin from range (a^o) to (a) as shown in Figure 6. The range of *mixed communication mode* also shrinks toward the origin from (b^o) to (b). In contrast, the *net-dominant range* expands toward the origin from (c^o) to (c).
(ii) Such changes in the optimal communication mode with the advancement of Net technology imply that the productivity of joint work increases significantly in a wide range of ε_F: To be specific, as shown in Figure 6, relative productivity increases more steeply as parameter ε_F moves from the new outer boundary ε_F^τ of F2F-dominant range toward the old inner boundary ${}^o\varepsilon_F^C$ of the net-dominant range: The relative proportion of productivity improvement reaches a maximum at ${}^o\varepsilon_F^C$, and it remains the same beyond ${}^o\varepsilon_F^C$.

As is apparent from Figure 6, although the relative productivity of joint work in the net-dominant range is always lower than in the F2F-dominant

[8] In Figure 6, the range, (b^o), significantly shrinks to (b). However, in order to know precisely how much range (b^o) shrinks, we need to know more specifically the relative magnitude of parameters in (63).

range, the former improves rapidly with the advance of net technology. Furthermore, with the advance of *transport technology* such as the appearance and improvement of rapid railways (e.g., the Shinkansen in Japan and China) and jet airplanes, the value of ε_F has been shrinking for the same geographical distance. This suggests that *with the advancement of net technology and transport technology, the location of potential partners for joint work expands*, as we will discuss further in Section 4.

Finally, we may reinterpret Figure 6 in a different context. Let us imagine that there are two groups of knowledge workers in terms of internet skills. One group consists of those workers who are have poor internet skills such as the "old generation" (like Masa) or people in living in environments that lack internet access, such as rural America. The other group consists of those workers who have mastered modern net technology (like Marcus). Then, we can naturally reinterpret Figure 6. The lower solid curve represents the relative productivity curve for workers with poor computer skills, whereas the upper solid curve represents that for workers with better skills. In this context, it is not difficult to imagine that restrictions on F2F communication such as those introduced during the Covid 19 pandemic will exert different impacts on the two groups, as we discuss further in Appendix A.

4 Knowledge composition, productivity, and choice of work mode

Given the joint output function A_{ij}^{J*} defined in equation (46), we have focused in Section 3 on the parameter ε_F specifying F2F lead time, and examined how communication mode and joint knowledge productivity are affected by a change in the value of ε_F, fixing the knowledge composition of joint workers. Conversely, in this section, we fix the value of parameter ε_F, and examine how knowledge composition of the pair of workers affects joint-work productivity. Furthermore, we specify the objective of each person in terms of knowledge creation, and define the rule used by each person to decide whether they create new ideas jointly or in isolation as a result of optimization of their objective.

To start, let us notice that in the joint output function A_{ij}^{J*} given by (46), the optimal combination of communication-mode frequencies, $\{\lambda_{CF}^*, \lambda_{CN}^*, \lambda_{\tau F}^*, \lambda_{\tau N}^*\}$, is left unspecified. To replace these unspecified variables with explicit functions of parameters, we suppose in this section that Assumption 1 holds, and hence Proposition 1 is valid. In this context, if we let $A_{ij}^{J*}(\varepsilon_F)$ be the value of A_{ij}^{J*} at each $\varepsilon_F \in [0, \infty)$, it follows from (53) that

$$A_{ij}^{J*}(\varepsilon_F) = A_{ij}^{J*}(0) \cdot h(\varepsilon_F),$$

where $h(\varepsilon_F)$ is the relative productivity function of joint work at parameter value ε_F. Furthermore, using (50) for $\varepsilon_F = 0$, we have:

$$\begin{aligned} & [\lambda_{CF}^*(0) \cdot \alpha_{CF} + \lambda_{CN}^*(0) \cdot \alpha_{CN}]^{\rho_{CT} \cdot \rho_C} \cdot [\lambda_{\tau F}^*(0) \cdot \alpha_{\tau F} + \lambda_{\tau N}^*(0) \cdot \alpha_{\tau N}]^{\rho_{\tau T} \cdot \rho_\tau} \\ = \; & [\alpha_{CF}]^{\rho_{CT} \cdot \rho_C} \cdot [\alpha_{\tau F}]^{\rho_{\tau T} \cdot \rho_\tau}. \end{aligned}$$

Hence, the optimized joint output function (46) can be expressed explicitly as a function of parameter ε_F as follows:

$$\begin{aligned} A_{ij}^{J*}(\varepsilon_F) \;=\; & n^{ij} \cdot \Phi_J \cdot \left[(1 - m_{ji}^d) + (1 - m_{ij}^d)\right]^{\rho_I + \rho_{CT} \cdot \rho_C + \rho_{\tau T} \cdot \rho_\tau} \\ & \times \left[(m_{ij}^c)^{1-\theta_C} \cdot (m_{ij}^d \cdot m_{ji}^d)^{\frac{\theta_C}{2}}\right]^{\rho_{CT} \cdot \rho_C} \cdot \left[(m_{ij}^c)^{1-\theta_\tau} \cdot (m_{ij}^d \cdot m_{ji}^d)^{\frac{\theta_\tau}{2}}\right]^{\rho_{\tau T} \cdot \rho_\tau} \\ & \times [\alpha_{CF}]^{\rho_{CT} \cdot \rho_C} \cdot [\alpha_{\tau F}]^{\rho_{\tau T} \cdot \rho_\tau} \cdot h(\varepsilon_F), \end{aligned} \tag{67}$$

where from (54), (55) and (56), the function $h(\varepsilon_F)$ is given by:

$$h(\varepsilon_F) = \begin{cases} (1 + \varepsilon_F)^{-(\rho_{CT} \cdot \rho_C + \rho_{\tau T} \cdot \rho_\tau)} & \text{for } 0 \leq \varepsilon_F < \varepsilon_F^\tau, \\ (1 + \varepsilon_F)^{-\rho_{CT} \cdot \rho_C} \cdot [\alpha_{\tau F}/\alpha_{\tau N}]^{-\rho_{\tau T} \cdot \rho_\tau} & \text{for } \varepsilon_F^\tau \leq \varepsilon_F < \varepsilon_F^C, \\ [\alpha_{CF}/\alpha_{CN}]^{-\rho_{CT} \cdot \rho_C} \cdot [\alpha_{\tau F}/\alpha_{\tau N}]^{-\rho_{\tau T}} & \text{for } \varepsilon_F^C \leq \varepsilon_F < \infty, \end{cases} \tag{68}$$

where ε_F^C and ε_F^τ are defined respectively by (48) and (49).

Given that parameter ε_F is fixed for the pair of joint workers, i and j, unless necessary, we express $A_{ij}^{J*}(\varepsilon_F)$ as just A_{ij}^{J*} for notational simplicity.

Before going further, we must specify the objective of each person in terms of knowledge creation. Specifically, we need to clarify how the output of the creation process, A_{ij}^{J*}, is split between i and j. In this connection, we introduced in our earlier papers two different specifications. In Berliant and Fujita (2008, 2009), assuming that A_{ij}^{J*} contributes directly to increasing each person's felicity (or instantaneous utility) at that time, A_{ij}^{J*} is not split between the two persons. It is a public good. In contrast, in Berliant and Fujita (2011, 2012), a fixed proportion of every collection of ideas created is assumed to become new patents, which are sold at the given market price at that time. The revenue from new patents is split evenly if persons i and j produce new ideas together. Although both specifications would lead to similar results, in this paper we adopt the latter specification of Berliant and Fujita (2011, 2012).

To go further, we must define the rule used by each person to decide whether they create new ideas jointly or in isolation. The rule is derived from optimization. For this purpose, we assume that income for each person derives from selling new ideas created as patents. *We assume that the revenue from new patents is split evenly if persons i and j produce new ideas together.*

Let $\delta_{ij}(t) \in \{0, 1\}$ be a function such that if

$$\delta_{ij}(t) = \delta_{ji}(t) = 1 \text{ for } j \neq i, \tag{69}$$

then joint knowledge creation occurs at time t. In contrast, if

$$\delta_{ii}(t) = \delta_{jj}(t) = 1, \tag{70}$$

then each person creates new knowledge in isolation. Notice that since the two persons must agree either to work jointly or to work in isolation each, either (69) or (70) can occur exclusively at time t.

Let $y_i(t)$ be the income of each person at time t, and let $\Pi(t)$ be the price of patents at time t. Then suppressing t for notational simplicity:

$$y_i = \Pi \cdot \left[\delta_{ii} \cdot A_{ii}^{I*} + \delta_{ij} \cdot A_{ij}^{J*}/2\right] \quad \text{for } j \neq i. \tag{71}$$

Since person i chooses δ_{ii} and δ_{ij} so as to maximize y_i and similarly for person

j, it follows that[9]

$$\delta_{ij} = \delta_{ji} = 1 \Longleftrightarrow A^{J*}_{ij}/2 > A^{I*}_{ii} \text{ and } A^{J*}_{ij}/2 > A^{I*}_{jj}, \tag{72}$$

$$\delta_{ii} = \delta_{jj} = 1 \Longleftrightarrow A^{J*}_{ij}/2 \leq A^{I*}_{ii} \text{ or } A^{J*}_{ij}/2 \leq A^{I*}_{jj}. \tag{73}$$

To study this system in greater detail, next we focus on the special case where persons i and j are *pairwise symmetric* in terms of knowledge heterogeneity. Specifically, suppose at time t that (suppressing t for notational simplicity):

$$n_i = n_j \equiv n. \tag{74}$$

By definition,

$$n^c \equiv n^c_{ij} = n^c_{ji}.$$

Then since $n^d_{ij} = n - n^c$, it follows that

$$n^d_{ij} = n^d_{ji} \equiv n^d, \tag{75}$$

and

$$n = n^c + n^d, \; n^{ij} = n^c + 2n^d = n + n^d, \tag{76}$$

implying that

$$\begin{aligned} m^d &\equiv m^d_{ij} = m^d_{ji} = \frac{n^d}{n^{ij}}, \; m^c \equiv m^c_{ij} = m^c_{ji} = \frac{m^c}{n^{ij}}, \\ m^c &\equiv m^c_{ij} = m^c_{ji} = \frac{n^c}{n^{ij}} = 1 - 2m^d. \end{aligned} \tag{77}$$

It also follows from (2) that

$$n = n^{ij} \cdot (1 - m^d), \tag{78}$$

Furthermore, from (7) and (8),

$$A^{I*}_{ii} = A^{I*}_{jj} = \Phi_I \cdot n. \tag{79}$$

Hence, relations (72) and (73) can be restated as

$$\delta_{ij} = \delta_{ji} = 1 \Longleftrightarrow A^{J*}_{ij}/2 > \Phi_I \cdot n, \tag{80}$$

$$\delta_{ii} = \delta_{jj} = 1 \Longleftrightarrow A^{J*}_{ij}/2 \leq \Phi_I \cdot n, \tag{81}$$

[9] Recall from equation (46) that $A^{J*}_{ij} = A^{J*}_{ji}$ by definition. In equations (72) and (73), we can use either strict inequality or weak inequality. However, since the case of ties is not important in the following analysis, for convenience we use strict inequality in (72) whereas we use weak inequality in (73).

where equation (67) for $A_{ij}^{J*} \equiv A_{ij}^{J*}(\varepsilon_F)$ is now given by

$$\begin{aligned} A_{ij}^{J*} &= n^{ij} \cdot \Phi_J \cdot \left[2 \cdot (1-m^d)\right]^{\rho_I+\rho_{CT}\cdot\rho_C+\rho_{\tau T}\cdot\rho_\tau} \cdot \left[(1-2m^d)^{1-\theta_C} \cdot (m^d)^{\theta_C}\right]^{\rho_{CT}\cdot\rho_C} \\ &\quad \times \left[(1-2m^d)^{1-\theta_\tau} \cdot (m^d)^{\theta_\tau}\right]^{\rho_{\tau T}\cdot\rho_\tau} \cdot [\alpha_{CF}]^{\rho_{CT}\cdot\rho_C} \cdot [\alpha_{\tau F}]^{\rho_{\tau T}\cdot\rho_\tau} \cdot h(\varepsilon_F). \end{aligned} \tag{82}$$

Using (78) and rearranging terms, let us define *the knowledge growth rate per person for joint creation,*

$$\begin{aligned} g_J(m^d) \equiv \tfrac{A_{ij}^{J*}/2}{n} &= \Omega \cdot (1-m^d)^{\rho_I+\rho_{CT}\cdot\rho_C+\rho_{\tau T}\cdot\rho_\tau-1} \cdot (1-2m^d)^{(1-\theta_C)\cdot\rho_{CT}\cdot\rho_C+(1-\theta_\tau)\cdot\rho_{\tau T}\cdot\rho_\tau} \\ &\quad \times (m^d)^{\theta_C\cdot\rho_{CT}\cdot\rho_C+\theta_\tau\cdot\rho_{\tau T}\cdot\rho_\tau} \cdot h(\varepsilon_F) \\ &= \Omega \cdot (1-m^d)^a \cdot (1-2m^d)^b \cdot (m^d)^c \cdot h(\varepsilon_F), \end{aligned} \tag{83}$$

where

$$\begin{aligned} \Omega &\equiv \Phi_J \cdot 2^a \cdot [\alpha_{CF}]^{\rho_{CT}\cdot\rho_C} \cdot [\alpha_{\tau F}]^{\rho_{\tau T}\cdot\rho_\tau}, \qquad (84) \\ a &\equiv \rho_I + \rho_{CT}\cdot\rho_C + \rho_{\tau T}\cdot\rho_\tau - 1, \\ b &\equiv (1-\theta_C)\cdot\rho_{CT}\cdot\rho_C + (1-\theta_\tau)\cdot\rho_{\tau T}\cdot\rho_\tau, \\ c &\equiv \theta_C\cdot\rho_{CT}\cdot\rho_C + \theta_\tau\cdot\rho_{\tau T}\cdot\rho_\tau. \end{aligned}$$

We note that the knowledge growth rate defined by equation (83) represents the "public aspect" of the knowledge creation and accumulation process. That is, A_{ij}^{J*} in equation (83) represents newly created "formal knowledge" to be accumulated in the "Server" in Figure 1 as public documents, which are accessible by any person in the future. In contrast, as discussed in Section 5, "inside the brain" of each person i and j, the tacit knowledge generated in the process of creating A_{ij}^{J*} is also accumulated for future knowledge creation. Furthermore, in equation (83), A_{ij}^{J*} is divided by 2 to account for the number of "patents" created per person.

By definition,

$$a + b + c = 0.$$

Since the function $g_J(m^d)$ contains only m^d as a variable, differentiating $g_J(m^d)$ yields

$$g_J'(m^d) = g_J(m^d) \cdot \frac{c-(b+2c)\cdot m^d}{m^d\cdot(1-m^d)\cdot(2-m^d)}. \tag{85}$$

Let us define

$$\begin{aligned} m^B &\equiv \frac{1}{2+\frac{b}{c}} \qquad (86) \\ &= \frac{1}{2+\frac{(1-\theta_C)\cdot\rho_{CT}\cdot\rho_C+(1-\theta_\tau)\cdot\rho_{\tau T}\cdot\rho_\tau}{\theta_C\cdot\rho_{CT}\cdot\rho_C+\theta_\tau\cdot\rho_{\tau T}\cdot\rho_\tau}} \end{aligned}$$

which is smaller than 1/2 since $\theta_C < 1$ and $\theta_\tau < 1$. Thus, (85) implies that

$$g'_J(m^d) \gtreqless 0 \text{ as } m^d \lesseqgtr m^B \text{ for } m^d \in \left(0, \frac{1}{2}\right) \tag{87}$$

Hence, $g_J(m^d)$ is strictly quasi-concave on $[0, 1/2]$, achieving its maximal value at m^B, which we call the "Bliss Point."

Next, for the case of knowledge creation *in isolation*, from equations (7) and (8), *the knowledge growth rate for each person* is obtained as follows:

$$\frac{A_{ii}^{I*}}{n_i} = \frac{A_{jj}^{I*}}{n_j} = \Phi_I, \tag{88}$$

which is constant for the two persons.

Overall, the knowledge growth rate per person is given by:

$$g(m^d) = \max\left\{\Phi_I, g_J(m^d)\right\} \tag{89}$$

Using function $g_J(m^d)$, the selection rule (72) and (73) can be restated as

$$\delta_{ij} = \delta_{ji} = 1 \Longleftrightarrow g_J(m^d) > \Phi_I \tag{90}$$

$$\delta_{ii} = \delta_{jj} = 1 \Longleftrightarrow g_J(m^d) \leq \Phi_I \tag{91}$$

Figure 7 depicts the shape of the function $g_J(m^d)$ for parameters:

(To be specified later)

together yielding

$$m_B = 0.4, \;\; g_J(m^B) = 1.0.$$

Figure 7

In Figure 7, the horizontal line with height

$$\Phi_I = 0.5$$

represents the knowledge growth rate when i and j are *working separately in isolation*, which is obtained from the set of parameters:

$$\alpha_I = \alpha_{IS} = \alpha_{IT} = 1, \; \rho_{IS} = \rho_{IT} = \frac{1}{2}.$$

This horizontal line intersects the $g_J(m^d)$ curve at points E and H. Hence,

$$\begin{aligned}\delta_{ij} &= \delta_{ji} = 1 \text{ for } m^d \in (m^E, m^H), &(92)\\ \delta_{ii} &= \delta_{jj} = 1 \text{ for } m^d \in [0, m^E] \cup [m^H, \frac{1}{2}]. &(93)\end{aligned}$$

At this point, let us recall that the knowledge growth function defined by equation (83) contains parameter ε_F. Hence, focusing on parameter ε_F, let us represent the knowledge growth rate function as $g_J(m^d; \varepsilon_F)$ for $\varepsilon_F \in [0, \infty)$. In this general context, relation (85) becomes

$$g'_J(m^d; \varepsilon_F) = g_J(m^d; \varepsilon_F) \cdot \frac{c - (b + 2c) \cdot m^d}{m^d \cdot (1 - m^d) \cdot (2 - m^d)}.$$

Since the last term of the equation above is the same as that of equation (85), we can conclude that *all knowledge growth rate curves, $g_J(m^d; \varepsilon_F)$ for $\varepsilon_F \in [0, \infty)$, share the same bliss point given by* (86). This is not surprising if we observe in equation (83) that the value of function $h(\varepsilon_F)$ simply shifts each knowledge growth rate curve vertically in the same proportion.

In Figure 8, three representative knowledge growth rate curves are depicted. The top curve, $g_J(m^d; 0)$, depicts the knowledge growth rate curve corresponding to $\varepsilon_F = 0$. The bottom curve, $g_J(m^d; \varepsilon_F^C)$, shows the knowledge growth rate curve at $\varepsilon_F = \varepsilon_F^C$, where ε_F^C represents the inner boundary of *net-dominant range*, given by (48). Since $g_J(m^d; \varepsilon_F) = g_J(m^d; \varepsilon_F^C)$ for all $\varepsilon_F \geq \varepsilon_F^C$, the bottom curve shows the knowledge growth rate curve for all ε_F in the net dominant range in Figure 5.

Figure 8

As ε_F increases from 0 to ε_F^C, the knowledge growth rate curve shifts down continuously from the top curve to the bottom curve. The middle curve, $g_J(m^d; \varepsilon_F)$, corresponds to a representative value of ε_F between 0 and ε_F^C.

Comparing the knowledge growth rate curves in Figure 8, we can make a few important observations. First, *in terms of vertical proportion*, all knowledge growth rate curves $g_J(m^d; \varepsilon_F)$ for $\varepsilon_F \in [0, \infty)$ locate within *a relatively small range*, i.e., between $h(0)$ and $h(\varepsilon_F^C)$. This reflects the fact that the relative productivity curve $h(\varepsilon_F)$ in Figure 5 locates within a relatively narrow belt. Second, in contrast, any knowledge growth rate curve in Figure 8, say $g_J(m^d; \varepsilon_F)$, *exhibits a huge variation in terms of horizontal axis parameter*

m^d. Furthermore, as noted previously in connection with Proposition 3, with the advancement of net technology and transport technology, the difference between $h(0)$ and $h(\varepsilon_F^C)$ shrinks. This suggests that in seeking for the most productive partner in knowledge creation, *knowledge composition becomes increasingly more important than the geographical distance* due to technological improvements.

For example, in the context of Figure 8, let us suppose that researcher i has three potential partners, j_1, j_2 and j_3. In terms of "knowledge composition m^d" and "distance ε_F", researcher i has the following combination with each potential partner; $(m^d_{ij_1}, 0)$, $(m^d_{ij_2}, \varepsilon_F)$ and $(m^d_{ij_3}, \varepsilon_F^C)$ such that

$$m^d_{ij_1} < m^d_{ij_2} < m^d_{ij_3} \quad \text{and} \quad 0 < \varepsilon_F < \varepsilon_F^C,$$

as shown in Figure 8. Then, the knowledge growth rate is lowest at [1] when i partners with person j_1, who locates nearest at $\varepsilon_F = 0$ and has the lowest degree of knowledge differentiation at $m^d_{ij_1}$. Conversely, the knowledge growth rate is highest at [3] in Figure 8, when i partners with person j_3, who locates farthest from i at ε_F^C and has the largest degree of knowledge differentiation at $m^d_{ij_3}$. Thus, provided that j_3 also agrees to work jointly with i, researcher i will choose j_3, who locates farthest from i.

In the next section, we treat m^d as an endogenous variable, and examine how m^d changes over time when a pair of researchers continue to work jointly. And we show that when a pair of researchers continues working jointly, the proportion of their differentiated knowledge tends to shrink over time.

We may summarize the observations above as follows:

Proposition 4. Keeping Assumption 1, we further assume that persons i and j have the same size of knowledge, n, at the time in question. In this context, let $g_J(m^d; \varepsilon_F) \equiv g_J(m^d)$ be the knowledge growth rate for joint creation per person, defined by equation (83). Then, the knowledge growth rate function $g_J(m^d; \varepsilon_F)$ exhibits the following features, as illustrated in Figure 8:
(i) For each $\varepsilon_F \in [0, \infty)$, $g_J(0; \varepsilon_F) \equiv g_J(1/2; \varepsilon_F) = 0$, and $g_J(m^d; \varepsilon_F)$ is strictly quasi-concave on $[0, 1/2]$.
(ii) As ε_F increases from 0 to ε_F^C, the knowledge growth rate curve $g_J(m^d; \varepsilon_F)$ shifts vertically down in the same proportion as the value of $h(\varepsilon_F)$, where $h(\varepsilon_F)$ is the relative productivity function defined by (68), and where ε_F^C represents the inner boundary of the net-dominant range given by (48). For all $\varepsilon_F \geq \varepsilon_F^C$, $g_J(m^d; \varepsilon_F) \equiv g_J(m^d; \varepsilon_F^C)$, and hence the bottom curve in Figure 8 shows the knowledge growth rate curve for all ε_F in the net dominant range in Figure 5.

(iii) Consequently, all knowledge growth rate curves achieve their maximum at the same Bliss Point m^B given by (86).
(iv) In terms of vertical proportions in Figure 8, all knowledge growth rate curves $g_J(m^d; \varepsilon_F)$ for $\varepsilon_F \in [0, \infty)$ locate within a relatively small range, i.e., between $h(0)$ and $h(\varepsilon_F^C)$. In contrast, any knowledge growth rate curve in Figure 8 exhibits a huge variation in terms of horizontal axis parameter m^d. Furthermore, as noted in connection with Proposition 3, with the advancement of net and transport technology, the difference between $h(0)$ and $h(\varepsilon_F^C)$ shrinks, whereas the horizontal variation of each knowledge growth rate curve in terms of m^d remains huge. This suggests that when a researcher seeks their best partner, their knowledge composition becomes increasingly more important than their geographical distance as technology improves.

5 Dynamics of the two-person system

Thus far, through Sections 2 to 4, the size of each person's knowledge as well as the relative composition of two persons' knowledge have been treated parametrically. In this section, we examine how the size of each person's knowledge as well as the relative composition of two persons' knowledge change endogenously over time when they continue interacting by sequentially choosing either to work alone or to work together for knowledge creation.

Next we shall define tacit knowledge and then explain how it arises in our context. The the ideas behind "tacit knowledge" originate with Polanyi (1958). Polanyi (1966, p.4) famously states, "We know more than we can tell." Nonaka and Takeuchi (1995, p. viii) develop these ideas further, defining tacit knowledge:

> "...we classify human knowledge into two kinds. One is *explicit knowledge*, which can be articulated in formal language including grammatical statements, mathematical expressions, specifications, manuals, and so forth. This kind of knowledge can be transmitted across individuals formally and easily. This has been the dominant mode of knowledge in the Western philosophical tradition. However, we shall argue, a more important type of knowledge is *tacit knowledge*, which is hard to articulate with formal language. It is personal knowledge embedded in individual experience and involves intangible factors such as personal belief, perspective, and

the value system. Tacit knowledge has been overlooked as a critical component of collective human behavior."

Borrowing one of their examples,[10] when designing and building a new automobile, certainly a large portion of the value is embodied in the final product sold to consumers. However, in the process of innovating the new design, the members of the team creating the vehicle acquire new tacit knowledge related to design and manufacturing that cannot be easily articulated. It comes with the experience of creation, either together or alone, and it is a part of human capital. Tacit knowledge of a researcher expands with their innovation over time.

Here we attempt to construct a model of (explicit and) tacit knowledge dynamics. Explicit knowledge can be bought and sold. It is difficult to market tacit knowledge, except as a package with a worker. Here we do not model the labor market in the context of tacit knowledge, leaving that to future work. Given our assumptions, tacit knowledge represents a dynamic externality in the following sense. The choice of a knowledge worker at a given time, in our case whether to work alone or in tandem with another, will affect their tacit knowledge acquisition during this work. This choice of one worker will affect the future payoff of that worker as well as others since tacit knowledge acquisition will affect future knowledge productivity either working alone or with others. In fact, since agents are assumed to be myopic, they do not account for the effect of tacit knowledge acquisition on their own future payoffs, much less the payoffs of others. So inefficient equilibria are to be expected. In contrast, Berliant and Fujita (2011) find that, under the same framework but with neither multimodal communication nor tacit knowledge, equilibria can have efficiency properties.

To study dynamics of the two-person system, in what follows we focus on the special case where persons i and j have the same size of knowledge at the initial time, $t = 0$; namely

$$n_i(0) = n_j(0) \equiv n(0). \tag{94}$$

It holds by definition that

$$n^c_{ij}(0) = n^c_{ji}(0) \equiv n^c(0),$$

[10] The particular example is the innovation involved in designing and producing the Honda City car.

thus as shown in equations (74), (75) and (76), condition (94) means

$$n_{ij}^d(0) = n_{ji}^d(0) \equiv n^d(0), \tag{95}$$

and

$$m_{ij}^d(0) = m_{ji}^d(0) \equiv m^d(0). \tag{96}$$

Given that the initial state of knowledge is symmetric for persons i and j as above, as seen below, it turns out that the equilibrium configuration at any time also maintains pairwise symmetry between the two persons; at any time $t \in [0, \infty)$,[11]

$$n_i(t) = n_j(t) \equiv n(t), \tag{97}$$

and hence,

$$n_{ij}^d(t) = n_{ji}^d(t) \equiv n^d(t), \tag{98}$$

$$m_{ij}^d(t) = m_{ji}^d(t) \equiv m^d(t), \tag{99}$$

whereas, by definition,

$$n_{ij}^c(t) = n_{ji}^c(t) \equiv n^c(t). \tag{100}$$

In this context of a pairwise-symmetric equilibrium path, our main goal is to examine how *the knowledge growth rate per person* changes in the long-run. As defined in equation (89), given $m^d \in [0, \frac{1}{2}]$, the knowledge growth rate per person, $g(m^d)$, is given by

$$g(m^d) = \max\left\{\Phi_I, g_J(m^d)\right\}. \tag{101}$$

To make the analysis interesting, let us consider the case

$$g_J(m^B) > \Phi_I. \tag{102}$$

Then, as depicted in Figure 7, on the horizontal axis m^d, there exist m^E and m^H such that $0 < m^E < m^B < m^H < \frac{1}{2}$ and

$$g_J(m^d) > \Phi_I \text{ for } m^d \in (m^E, m^H),$$

$$\Phi_I > g_J(m^d) \text{ for } m^d \in \left[0, m^E\right) \cup (m^H, \frac{1}{2}]$$

Let $m^d(0)$ be the initial value of m^d at time 0. First, let us assume that

$$g_J(m^d(0)) > \Phi_I. \tag{103}$$

[11] This can be seen from the fact that when condition (97) is met at any time t (say, $t = 0$), the right hand side of the differential equation specifying $\dot{n}(t)$ (also $\dot{n}^d(t)$ and $\dot{m}^d(t)$) does not involve any variable that is specific to i or j. For a more precise explanation of this point, please see Berliant and Fujita (2011).

That is, $m^d(0)$ locates between m^E and m^H in Figure 7. Then, the two persons will continue to work jointly (i.e., $\delta_{ij} = \delta_{ji} = 1$) as long as

$$g_J(m^d(t)) > \Phi_I. \tag{104}$$

In order to know which direction $m^d(t)$ moves and how long the two persons will continue working jointly, let us examine the dynamics of m^d when condition (104) holds. To do so, recalling the definition $m^d \equiv n^d/n^{ij}$ and dropping t for simplicity, we have that:

$$\begin{aligned} \dot{m}^d &= \frac{\dot{n}^d}{n^{ij}} - \frac{n^d \cdot \dot{n}^{ij}}{(n^{ij})^2} \\ &= (n^{ij})^{-1} \cdot [\dot{n}^d - \dot{n}^{ij} \cdot m^d], \end{aligned}$$

Since $n^{ij} \equiv n^c + 2n^d$, it follows that

$$\dot{m}^d = (n^{ij})^{-1} \cdot [\dot{n}^d - (\dot{n}^c + 2\dot{n}^d) \cdot m^d] \tag{105}$$

implying that

$$\dot{m}^d \gtreqless 0 \Leftrightarrow \dot{n}^d - (\dot{n}^c + 2\dot{n}^d) \cdot m^d \gtreqless 0. \tag{106}$$

Hence, to identify the sign of $\dot{m}^d$, we must know the values of $\dot{n}^d$ and $\dot{n}^c$. That is, when $m^d(t)$ locates between m^E and m^H in Figure 7, the knowledge growth rate $g_J(m^d(t))$ is realized for each person, and therefore we must calculate how much *differential knowledge* ($\dot{n}^d(t)$) for each person and how much *common knowledge* ($\dot{n}^c(t)$) for the two persons has been generated at that moment, t. We calculate the values of $\dot{n}^d(t)$ and $\dot{n}^c(t)$ through a sequence of steps as follows:

In the first step, let us recall Figure 3, which represents the structural relationship among all basic activities and intermediate activities used to generate the final output, A_{ij}^{J*}, as a tree. Table 1 provides further description of this activity tree.

Table 1

The activity tree depicted in Figure 3 consists of *three tiers* of knowledge creation activities. We may note that in describing a tree in mathematics, the term "level" is used instead of "tier." In the present context, however, the diagram shown in Figure 3 resembles the structure of a *supply chain* that produces a final industrial product (e.g., automobiles) by sequentially assembling various parts produced through *multiple tiers* of many factories. Thus, here we use the term "tier." A major difference between "automobile production" and the present "knowledge production" must be noted. That is, all parts used in the production process of a car "disappear" into the car at the end of the assembly process. In contrast, the tacit knowledge created and used in the knowledge production process is accumulated for knowledge creation in the future.

In Figure 3, each node denoted by a black square ■ or a black dot • represents a *joint activity* by persons i and j, whereas each node denoted by a white circle ∘ represents an *independent activity* conducted by either i or j alone. The symbol $*$ attached to a node means that the frequency variables (ω's, λ's and/or μ's) have been chosen optimally.

In the second step, the output of each activity is represented by an equation, which is listed in the third column of Table 1. Although each equation in this column of Table 1 represents new tacit knowledge generated by each activity, different pieces of new tacit knowledge cannot simply be added up because each one has different importance for the creation of new knowledge in the future. Thus, in the third step, we attach a "weight" to each specific output, say x. For the weight on x, following the tradition of microeconomics, we use the value $\partial A_{ij}^{J*}/\partial x$, representing the marginal contribution of x to the final output A_{ij}^{J*}. Then, we calculate the (total) contribution of x to the final output A_{ij}^{J*} by

$$\frac{\partial A_{ij}^{J*}}{\partial x} \cdot x, \tag{107}$$

which we call the *imputed value* of x, using the terminology of microeconomics. Here, it must be noted that in each calculation of (107), the *pairwise symmetry* stated in equation (97), namely $n_i(t) = n_j(t) \equiv n(t)$, is taken into account. The imputed value of each output is listed in the last column of Table 1. (Please see Appendix B for the actual calculation of imputed values.)

Next, using the terms listed in the last column of Table 1, we calculate the values of $\dot{n}^d(t)$ and $\dot{n}^c(t)$. Before doing so, however, let us observe from the last column of Table 1 that

$$\begin{aligned}
& \text{the sum of imputed values in the second tier} \\
= \;& A_{ij}^{J*} \cdot \frac{\rho_I}{2} + A_{ij}^{J*} \cdot \frac{\rho_I}{2} + A_{ij}^{J*} \cdot \rho_F + A_{ij}^{J*} \cdot \rho_N \\
= \;& A_{ij}^{J*} \cdot (\rho_I + \rho_F + \rho_N) \\
= \;& A_{ij}^{J*},
\end{aligned} \tag{108}$$

since $\rho_I + \rho_F + \rho_N = 1$. Likewise, it can be readily confirmed that

$$\begin{aligned} & \text{the sum of imputed values in the third tier} \\ = \ & A_{ij}^{J*}. \end{aligned} \tag{109}$$

Therefore, when the final output A_{ij}^{J*} is created, the same size of tacit knowledge (in terms of imputed values) is generated in tier 2 as well as in tier 3.[12]

In calculating the values of $\dot{n}^d(t)$ and $\dot{n}(t)$, we may take into account all imputed values in tier 2 and tier 3. However, the contents of tacit knowledge generated in tier 2 seem rather close to those in tier 3. Thus, in order to avoid double counting of similar pieces of tacit knowledge, we consider only the tacit knowledge generated in tier 3.[13] In this context, the value of $\dot{n}^d$ for each person, say i, is obtained by summing up the imputed values of all third-tier activities *by person* i that are *identified by* $\circ$ *circles* in the second column of Table 1 as follows:

$$\begin{aligned} \dot{n}^d &= A_{ij}^{J*} \cdot \frac{\rho_I \cdot \rho_{JIS}}{2} + A_{ij}^{J*} \cdot \frac{\rho_I \cdot \rho_{JIT}}{2} + A_{ij}^{J*} \cdot \frac{\rho_F \cdot \rho_{ISF}}{2} + A_{ij}^{J*} \cdot \frac{\rho_N \cdot \rho_{ISN}}{2} \\ &= \frac{A_{ij}^{J*}}{2} \cdot (\rho_I + \rho_F \cdot \rho_{ISF} + \rho_N \cdot \rho_{ISN}), \end{aligned} \tag{110}$$

using the identity $\rho_{JIS} + \rho_{JIT} = 1$. This equation represents the size of the *differential knowledge* generated by each person at the time. Next, the value of $\dot{n}^c$ is obtained as the sum of final output A_{ij}^{J*} and the imputed values of third-tier *joint activities* by i and j as follows:

$$\begin{aligned} \dot{n}^c &= A_{ij}^{J*} + A_{ij}^{J*} \cdot \rho_C \cdot \rho_{CT} \cdot \frac{\lambda_{CF}^* \cdot a_{ij}^{CTF} + \lambda_{CN}^* \cdot a_{ij}^{CTN}}{\lambda_{CF}^* \cdot a_{ij}^{CTF} + \lambda_{CN}^* \cdot a_{ij}^{CTN}} \\ &\quad + A_{ij}^{J*} \cdot \rho_\tau \cdot \rho_{\tau T} \cdot \frac{\lambda_{\tau F}^* \cdot a_{ij}^{\tau TF} + \lambda_{\tau N}^* \cdot a_{ij}^{\tau TN}}{\lambda_{\tau F}^* \cdot a_{ij}^{\tau TF} + \lambda_{\tau N}^* \cdot a_{ij}^{\tau TN}} \\ &= A_{ij}^{J*} \cdot (1 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T}). \end{aligned} \tag{111}$$

This equation represents the size of *common knowledge* generated by the joint work of i and j.

[12]If we recall the microeconomics of production, this result is not surprising. That is, "all production functions / sub-production functions" in each tier are composed of Cobb-Douglas functions, and each function displays constant returns to scale. Hence, the sum of all imputed values in each tier equals the value of the final output, the price of which is normalized to be 1.

[13]Even if we take into account all pieces of tacit knowledge in tier 2 and tier 3, the main results below would not change qualitatively.

Finally, substituting (110) and (111) into equation (105) and noting that $\dot{n}^c + 2\dot{n}^d = 2A_{ij}^{J*}$, we have that

$$\dot{m}^d = (n^{ij})^{-1} \cdot \frac{A_{ij}^{J*}}{2} \cdot [(\rho_I + \rho_F \cdot \rho_{ISF} + \rho_N \cdot \rho_{ISN}) - 4m^d] \quad \text{for } m^d \in (m^E, m^H). \tag{112}$$

Let us define

$$\tilde{m} \equiv \frac{\rho_I + \rho_F \cdot \rho_{ISF} + \rho_N \cdot \rho_{ISN}}{4}, \tag{113}$$

which represents the *stationary point* of $\dot{m}^d$. Then, it follows from (112) that

$$\dot{m}^d \gtreqless 0 \text{ as } m^d \lesseqgtr \tilde{m} \quad \text{for } m^d \in (m^E, m^H). \tag{114}$$

Notice that both the Bliss Point m^B and the stationary point $\tilde{m}$ are independent of parameter ε_F. It is also clear from (113) that $\tilde{m} < 1/4$.

Although we have identified the sign of $\dot{m}^d$ when the two persons work jointly, we must also consider the case when *two persons work independently.* Fortunately, this case is rather simple. First, by definition of $m^d \equiv n^d/n^{ij}$, the dynamics of m^d are given by equation (105) also in the present context of each person working in isolation. Hence, to identify the sign of $\dot{m}^d$, we need to know the values of $\dot{n}^d$ and $\dot{n}^c$ in the present context. To do so, we focus on person i. Then, the activity tree for the knowledge created by person i in isolation can be represented as in Figure 2.

The top node $\Box$ in Figure 2 represents the final output created by person i in isolation. The second tier consists of two basic activities: a_i^{IT} and a_i^{IS}. From (6), the imputed value of each intermediate output can be obtained as follows:

$$\frac{\partial A_{ii}^{I*}}{\partial a_i^{IT}} \cdot a_i^{IT} = A_{ii}^{I*} \cdot \rho_{IT}, \ \frac{\partial A_{ii}^{I*}}{\partial a_i^{IS}} \cdot a_i^{IS} = A_{ii}^{I*} \cdot \rho_{IS}. \tag{115}$$

implying that:

$$\begin{aligned} & \text{the sum of imputed value in the second tier} \\ = \ & A_{ii}^{I*} \cdot \rho_{IT} + A_{ii}^{I*} \cdot \rho_{IS} \\ = \ & A_{ii}^{I*}, \end{aligned} \tag{116}$$

since $\rho_{IT} + \rho_{IS} = 1$.

By assumption, since person i is working in isolation (from j), the formal output A_{ii}^{I*} and imputed values of second-tier activities by person i become *differential knowledge* of person i (from j).[14] Hence, the value of $\dot{n}^d$ for person i is obtained as follows:

$$\begin{aligned}\dot{n}^d &= A_{ii}^{I*} + A_{ii}^{I*} \cdot \rho_{IS} + A_{ii}^{I*} \cdot \rho_{IT} \\ &= 2A_{ii}^{I*}, \end{aligned} \tag{117}$$

whereas no common knowledge is generated for the two persons:

$$\dot{n}^c = 0. \tag{118}$$

Substituting (116) and (117) into (105) yields

$$\dot{m}^d = (n^{ij})^{-1} \cdot 2A_{ii}^{I*} \cdot (1 - 2m^d) > 0 \text{ for } m^d \in [0, \frac{1}{2}). \tag{119}$$

Therefore, whenever each person is working in isolation, the proportion of differential knowledge, m^d, will keep increasing.

Now, combining the two situations of working jointly and working in isolation by two persons, we can derive the overall dynamics of m^d. As explained in Appendix C, depending on the relative position of $\tilde{m}$ and m^B on the m^d-axis, we have three different diagrams for the dynamics of two-person system, which are depicted in Figure C. However, given that the three diagrams in Figure C are qualitatively rather similar, we focus in the following on the representative dynamics presented in Figure 9.

Figure 9

The diagram in Figure 9 is based on the assumption that

$$m^E < \tilde{m} < m^B. \tag{120}$$

As depicted in this diagram, if the initial position, $m^d(0)$, at time 0 is to the left of $\tilde{m}$, then $m^d(t)$ gradually increases towards $\tilde{m}$. If $\tilde{m} < m(0) < m^H$, then $m^d(t)$ gradually decreases towards $\tilde{m}$. Only when $m^d(0) < m^H$, does $m^d(t)$ moves away from $\tilde{m}$ toward 1/2. Hence, whenever $m^d(0) < m^H$, $m^d(t)$ approaches the *sink point*, $\tilde{m}$, which is to the left of bliss point m^B.

[14]The formal output A_{ii}^{I*} by person i will be registered in the knowledge stock in the Server. Given that the size of the total stock of knowledge in the Server is almost infinite, it can be safely assumed that the probability of person j finding A_{ii}^{I*} in the Server is nearly zero (and vice versa for i). Even if we assume that a part of A_{ii}^{I*} and A_{ij}^{I*} is transferred to each other by some mechanism, it will not change the sign of $\dot{m}^d$ below.

Notice that the value of $\tilde{m}$ defined by equation (113) is always less that $1/4$. On the other hand, as explained in Appendix C, the bliss point m^B is close to $1/2$ when the weights on differential knowledge in K-subproduction functions (9) and (29) are sufficiently large. In such a case, as illustrated in Figure 9, the K-knowledge growth rate at the sink point $\tilde{m}$ is much lower than that at the bliss point. In other words, the research partners are *in a low-productivity sink point trap.* Thus, it is important to ask: When the research partners are in a low-productivity sink point trap, what possible mechanism could enable the partners to escape from such a trap and attain much higher productivity? Before examining this question further in the last section, however, we need to ask another fundamental question here.

Recall the discussion near the end of Section 4 stating that A^{J*}_{ij} in equation (83) represents newly created *formal knowledge* to be accumulated in the Server as public documents. However, we have just seen that the knowledge production process also yields a large amount of *tacit knowledge* which will be accumulated *inside the brain* of each person for future knowledge creation activity. Thus, we need to know the growth rate of *total knowledge* per person, including both formal knowledge and tacit knowledge. Furthermore, we must examine the relationship between the knowledge growth rate defined by equation (83) and the growth rate of total knowledge for each person.

To answer these questions, let us first calculate the growth rate of total knowledge, $\dot{n}_i$, for person i *when working in isolation.* From (117) and (118), we have that

$$\begin{aligned} \dot{n}_i &= \dot{n}^d + \dot{n}^c \\ &= 2A^{I*}_{ii}. \end{aligned} \tag{121}$$

Thus, recalling (88), the growth rate of total knowledge for person i when working in isolation is given by

$$G_I \equiv \frac{\dot{n}_i}{n_i} = \frac{2A^{I*}_{ii}}{n_i} = 2\Phi_I. \tag{122}$$

Next, when person i is working jointly with j, from (110) and (111) we have that

$$\begin{aligned} \dot{n}_i &= \dot{n}^c_i + \dot{n}^d_i = \dot{n}^c + \dot{n}^d \\ &= A^{J*}_{ij} \cdot (1 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN}) + \frac{A^{J*}_{ij}}{2} \cdot (\rho_I + \rho_F \cdot \rho_{ISF} + \rho_N \cdot \rho_{ISN}) \\ &= A^{J*}_{ij} \cdot (1 + \frac{1 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN}}{2}), \end{aligned} \tag{123}$$

by using the relations $\rho_{CT} + \rho_{CS} = 1$ and $\rho_{\tau T} + \rho_{\tau S} = 1$. Hence, the growth rate of total knowledge for person i when working jointly with j is given by

$$\begin{aligned} G_J(m^d) &\equiv \frac{\dot{n}_i}{n_i} = \frac{A_{ij}^{J*}}{n_i} \cdot (1 + \frac{1 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN}}{2}) \\ &= \frac{A_{ij}^{J*}/2}{n_i} \cdot (3 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN}) \\ &= g_J(m^d) \cdot (3 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN}), \end{aligned} \tag{124}$$

recalling definition (83) and setting $n_i = n$.

Putting together (122) and (124), *the growth rate of total knowledge per person* is given by:

$$G(m^d) = \begin{cases} 2\Phi_I & \text{when } \Phi_I \geq g_J(m^d), \\ g_J(m^d) \cdot (3 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN}) & \text{when } g_J(m^d) > \Phi_I. \end{cases} \tag{125}$$

Let us recall equation (89) which defines *the growth rate of public knowledge* per person. By comparing equations (89) and (125), we can see that the two functions are significantly different, but they also have a close relationship. First, we can observe that the upper term in equation (125) is twice the corresponding term in equation (89), reflecting the fact that the total knowledge newly created in isolation includes the same amount of tacit knowledge (given by (116)) as formal knowledge. The lower term in equation (125) is more than the triple of the corresponding term in equation (89), reflecting the fact that the total knowledge newly acquired by person i through joint work with person j includes the *full amount* of newly created formal knowledge, A_{ij}^{J*}, and the tacit knowledge represented by equations (110) and (111).[15]

Next, observe from equation (124) that since function $G_J(m^d)$ is the product of $g_J(m^d)$ and a constant, it follows that

$$\begin{aligned} m^B &\equiv \text{the bliss point of function } g_J(m^d) \\ &= \text{the bliss point of function } G_J(m^d), \end{aligned} \tag{126}$$

where m^B is defined by (86), meaning that the maximum values of two functions, $g_J(m^d)$ and $G_J(m^d)$, are attained at the same point, $m^B \in (0, 1/2)$.

We are now ready to summarize the dual dynamics of formal-K and tacit-K in the two-person system. Notice that along the symmetric equilibrium path,

[15] Recall that in the definition of the function $g_J(m^d)$ in (83), the formal K-output A_{ij}^{J*} created by joint work is *divided by* 2 to account for the number of "patents" created *per person*. In contrast, in deriving equation (122), the full amount of knowledge embodied in these "patents," A_{ij}^{J*}, is considered a part of *new knowledge acquired by each person.*

at any $t \in [0, \infty)$, if values of $n_i(t) = n_j(t) \equiv n(t) > 0$ and $m^d(t) \in (0, 1/2)$ are known, then values of all the rest of structural variables, $n^c(t) \equiv n^c_{ij}(t) = n^c_{ji}(t)$, $n^d(t) \equiv n^d_{ij}(t) = n^d_{ji}(t)$, $n^{ij}(t) \equiv n(t) + n^d(t)$, $m^c(t) = 1 - 2m^d(t)$, $A^{J*}_{ij}(t)$, $A^{I*}_{ii}(t)$, $g_J(m^d(t))$, $g_I(m^d(t))$ are uniquely determined, together with the values of choice rules,

$$\delta_{ii}(t) = \delta_{jj}(t) = 1 \text{ if } \Phi_I \geq g_J(m^d(t)), \;\; \delta_{ij}(t) = \delta_{ji}(t) = 1 \text{ if } \Phi_I < g_J(m^d(t)). \tag{127}$$

Hence, given the initial values, $n(0) > 0$ and $m^d(0) \in (0, 1/2)$, solving the system of differential equations, (112) and (119) for $\dot{m}^d$, and (121) and (123) for $\dot{n} = \dot{n}_i = \dot{n}_j$, together with the choice rule (127), the equilibrium path of $m^d(t)$ and $n(t)$, together with all the rest of the structural variables, is uniquely determined for $t \in [0, \infty)$.

For example, let us take the case of diagram (a) depicted in Figure C, which is qualitatively the same as Figure 9. In the context of diagram (a), Figure 10 synthesizes the dual dynamics of formal-K and tacit-K in the two-person system.

Figure 10

In the bottom part of Figure 10, diagram (a) in Figure C is duplicated, representing the dynamics of formal-K for the case of $m^E < \tilde{m} < m^B$. The top part of Figure 10 depicts the dynamics of total-K. Comparing the bottom part and the top part of Figure 10, we can recognize similarities and differences between the two dynamics. Since the upper curve $G_J(m^d)$ is the product of the lower curve $g_J(m^d)$ with the constant $(3 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN})$, the two curves share the same bliss point, m^B, and the same sink point, $\tilde{m}$, on the m^d-axis. Hence, in the phase of joint work by i and j, and hence when $m^E < m^d(t) < m^H$, the two dynamics are essentially parallel, leading respectively to the sink points S and S^* at the same $\tilde{m}$. However, since the constant of multiplication, $(3 + \rho_F \cdot \rho_{JTF} + \rho_N \cdot \rho_{JTN})$, is more than 3, the upper curve $G_J(m^d)$ is much higher than the lower curve $g_J(m^d)$. This indicates that by working jointly for knowledge creation, the accumulation of total knowledge by each person is much higher than that of formal knowledge created jointly.

In contrast with joint work, during the phase of knowledge creation in isolation by each person independently, the dynamics with and without tacit knowledge are significantly different. In the bottom part of Figure 10, formal-K dynamics without tacit knowledge happens when $m^d(0) < m^E$. The switch from working in isolation to joint work occurs at the point E, where the horizontal line of height Φ_I crosses the $g_J(m^d)$ curve. Since m^E on the m^d-axis represents the real switching point based on the switching rule noted in (127), in the top part of Figure 10 *the switch to joint work is realized at point* E' *when the growth rate of total knowledge,* $2\Phi_I$*, from working in isolation is much lower than that from working together.* For efficient growth of total knowledge in joint work, the switch should occur much earlier at E'', where the horizontal line with height $2\Phi_I$ crosses the $G_J(m^d)$ curve. The opposite situation happens at point m^H on the m^d-axis. Hence, from the viewpoint of total knowledge accumulation, the *myopic switching rule* (127) is not optimal.

To see more clearly why the switching rule (127) is not optimal from the viewpoint of total knowledge growth, let us compare switching at E with switching at E'' in Figure 10. Based on *per capita production of formal-*K, switching occurs at point E where $\Phi_I = g_J(m^E)$. Since $A_{ii}^{I*}(m^d) = \Phi_I \cdot n$ for all $m^d \in [0, \frac{1}{2}]$, recalling (83) we can rewrite this relationship at E as follows:

$$A_{ii}^{I*}(m^E) = A_{ij}^{J*}(m^E) \cdot \frac{1}{2}. \tag{128}$$

In contrast, *for achieving efficient growth of total-*K*, switching should occur at point* E'' where $2\Phi_I = G_J(\hat{m})$, that is, when

$$2A_{ii}^{I*}(\hat{m}) = \frac{A_{ij}^{J*}(\hat{m})}{2} \cdot (3 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T})$$

or

$$A_{ii}^{I*}(\hat{m}) = A_{ij}^{J*}(\hat{m}) \cdot \frac{(3 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T})}{4}. \tag{129}$$

Switching rule (128) means that switching from work in isolation to joint work should occur *at point* E where *one half of joint output is equal to the output of work in isolation.* On the other hand, switching rule (129) means that switching from the work in isolation to joint work should occur *at point* $\hat{m}$ *where* $(3 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T})/4$ *of joint output is equal to the output of work in isolation.*

The difference between the two switching rules becomes clear when we consider this situation in the context of the academic world. An evaluation committee always faces the difficult problem of how to evaluate the contribution of each author when papers are written jointly. A simple, drastic rule would be that when a paper is written by 2 authors, the contribution of each author just equals one half of the contribution of the paper. Equation (128) represents this drastic rule. In the actual academic world, however, an evaluation committee tends to adopt a more generous rule which is represented qualitatively by equation (129). That is, *if we consider efficient growth of the total knowledge of each (young) researcher in the long-run*, when a paper is written by 2 persons, the evaluation committee is justified to allocate much more than one half of the paper's contribution to each author. That is because tacit knowledge accumulates during the knowledge production process; such knowledge is invisible in the final product, namely the paper itself.

To understand more intuitively the suboptimality of the switching rule based on equation (128), let us suppose as indicated by (128) that at a given time, *one isolated researcher* writes *one paper* per unit time, whereas *two joint-researchers* write *two papers* per unit time. In this situation, each of the two joint-researchers accumulates its formal-K (from two papers) at a rate twice that of isolated research; furthermore, each of the joint-researchers accumulates informal-K (from developing two papers) much faster than an isolated researcher [to be precise, $(1 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T})$ times that of the isolated researcher]. Therefore, when relation (128) holds, the total-K of each joint-research accumulates much faster than that of an isolated researcher.

Returning to our case where each researcher gets its revenue from the sales of newly created patents (per person), the optimal switching rule (129) justifies a government (or public agent) subsidy for a significant portion of the revenue of each joint researcher; more precisely, the subsidy should be targeted in the following proportion:

$$\frac{3 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T}}{4} - \frac{1}{2} = \frac{1 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T}}{4}. \tag{130}$$

We may summarize the observations above as follows:

Proposition 5. To study the dynamics of knowledge creation in the context of a two-person system, let us focus on two persons i and j who share the same fixed value of F2F lead time ε_F. Suppose further that persons i and j have the same quantity of knowledge at initial time 0 as indicated by (94). Then, it turns out that the equilibrium knowledge configuration of the two persons maintains pairwise symmetry as described by the relations (97), (98), (99) and (100) at any time $t \in [0, \infty)$. In this context of symmetry with two persons, the equilibrium dynamics of knowledge creation exhibit the following features: (i) When $m^d(t)$ is between m^E and m^H in Figure 7, and hence the two persons continue working jointly, the value $\dot{m}^d$ is given by equation (112). This has been obtained by using the identity (105) and taking into account the final output A_{ij}^{J*} and the imputed value (listed in the last column of Table 1) of the output of each basic activity in the third-tier of Figure 3. Likewise, when

person i is working in isolation (from j), the value of $\dot{m}^d$ is given by equation (119). The dynamics of joint work represented by equation (112) has a unique stationary point $\tilde{m}$ given by equation (113). From equation (119), the value of working in isolation is always positive for $m^d \in [0, 1/2)$.
(ii) Based on equations (112) and (119), representative dynamics for $m^d(t)$ are shown in Figure 9. This corresponds to the case where $m^E < \tilde{m} < m^B$. In this case, as shown in Figure 9, whenever $m^d(0) < m^H$, $m^d(t)$ approaches the sink point $\tilde{m}$.
(iii) On the one hand, the value of $\tilde{m}$ defined by equation (113) is always less than 1/4. On the other hand, the bliss point m^B is close 1/2 when each weight on the differential knowledge in K-subproduction functions (9) and (29) is sufficiently large. In such a case, as shown in Figure 9, the K-growth rate at the sink point $\tilde{m}$ is much lower than that at the bliss point. In other words, the two research partners eventually fall into a trap of a low-productivity sink point.
(iv) In the knowledge growth rate function $g_J(m^d)$ defined by equation (83), A_{ij}^{J*} represents newly created *formal knowledge* to be accumulated in the Server as public documents. However, the knowledge production process also yields a large amount of tacit knowledge created (and maintained) at each node identified by $\circ$ or $\bullet$ in Figure 3. Thus, we need to know the growth rate of *total knowledge* per person, *including both formal knowledge and tacit knowledge.* Similar to deriving equations for $\dot{m}^d$, the rate of increase in total knowledge of each person, say i, is given by equations (121) and (123), that is:

$$\begin{aligned}
\dot{n}_i &= 2A_{ii}^{I*} \quad \text{when } i \text{ is working in isolation,} \\
\dot{n}_i &= A_{ij}^{J*} \cdot \left(1 + \frac{1 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T}}{2}\right) \quad \text{when } i \text{ is jointly working with } j.
\end{aligned}$$

Hence the growth rate of total knowledge for i is given by equations (122) and (124), that is:

$$\begin{aligned}
G_J &= \frac{\dot{n}_i}{n_i} = 2\Phi_I \quad \text{when } i \text{ is in isolation,} \\
G_J(m^d) &= \frac{\dot{n}_i}{n_i} = g_J(m^d) \cdot (3 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T})
\end{aligned}$$

(v) Thus, keeping the context of Figure 9 or diagram (a) in Figure C, the dual dynamics of formal-K and total-K for the two-person system can be represented by Figure 10, as explained previously. Since the upper curve $G_J(m^d)$ is the product of the lower curve with the constant $(3 + \rho_C \cdot \rho_{CT} + \rho_\tau \cdot \rho_{\tau T})$, the two dynamics are essentially parallel when $m^E < m^d(t) < m^H$,

leading respectively to the sink points S and S^* at the same $\tilde{m}$. However, when $m^d(0) < m^E$, based on the switching rule (127), the switch from working in isolation to joint work occurs at point E; in terms of the growth of total knowledge represented by the upper curve, the switching is realized at point E' when the growth rate of total knowledge, $2\Phi_I$, from working isolation is much lower than that from working together. For efficient growth of total knowledge in joint work should occur much earlier at E'' when the horizontal line which height $2\Phi_I$ crossed the $G_J(m^d)$ curve. Hence from the view-point of total knowledge accumulation, the myopic switching rule (127) is not optimal.
(vi) To remedy the inefficiency caused by the myopic switching rule (127), when the income of each researcher comes from the sales of newly created patents per person, a public agent may subsidize a significant portion of the revenue, in a proportion given by (130).

6 Conclusions

Building on our earlier work, we have developed a model of knowledge creation in the context of two persons when multiple modes of communication are available, and knowledge workers can independently use the internet for the purpose of search. As in the empirical literature, we separate the development of new ideas into conceptual and technical phases, allowing researchers a choice between modes of communication, face to face or remote, in each phase. Tacit knowledge, missing from earlier models, plays a huge role in our analysis of the dynamics of the system. We apply our framework to analyze consequences of Covid policies and firm requirements for office presence. When considering long run dynamics, the sink point of the system generally yields a suboptimal configuration of knowledge differentiation. Finally, we find that if knowledge workers use a myopic rule to determine whether they work jointly or independently, they will not internalize tacit knowledge creation in their decisions, leading to a second source of inefficiency. In summary, we have seen the importance of tacit knowledge when considering the process of knowledge creation in the long run.

How can these two inefficiencies be addressed? For the first type of inefficiency, illustrated in Figure 8(a,c), the solution would be to have larger research group sizes as pointed out in our earlier work. This prevents the dynamic buildup of knowledge in common, so the bliss point can be decentralized using the myopic core. Our restriction to two persons here does not allow that, so the extension of the model to more than two persons would have to be considered.

For the second inefficiency, the situation illustrated in Figure 8(b), again we should consider an extension of the model to more knowledge workers. If the workers have a diversity of backgrounds in terms of differential knowledge relative to each other at time 0, initial partners for knowledge creation could be chosen so that this inefficiency is excluded. In other words, it would be best if a person's initial partner were chosen so that knowledge in common is large. Both partners would find this optimal in the long run, and thus could be optimal for forward looking agents. Another way to justify this is to notice that in Figure 8(b), every knowledge state to the right of the bliss point can be matched in terms of knowledge growth by a state to the left of the bliss point, but the (absolute value of the) slope of the knowledge creation function is *higher to the left of the bliss point* for each level of growth rate or point on the vertical axis. So even if agents are myopic, the *derivative* of the rate of growth to the left of the bliss point is higher at the same level of knowledge growth as the corresponding point to the right of the bliss point. Given a choice of backgrounds for their initial partner, myopic agents would choose an initial partner with more knowledge in common. That is to say, it is a weakly dominant myopic strategy. The assumption here is that if there is an initial partner with a profile to the right of the bliss point available, then there is also one to the left of the bliss point with the same or higher initial K-growth rate available.

Regarding the second inefficiency, a result of the myopic rule and tacit knowledge, the use of forward looking agents should be investigated.

In the future, it would be interesting to let θ vary *endogenously* with the particular knowledge profile of the two agents who are meeting to produce new knowledge. That is, the agents who meet know their respective profiles of common and differential knowledge, but choose a project summarized by θ to maximize their output. Then, the output curve would be the upper envelope over θ of productivity at the bliss points.

Of course, it is important to extend the model to many people in a spatial context to examine patterns of knowledge creation in an academic society. It would also be of interest to allow other forms of agent heterogeneity, such as large and small city residents, or CBD and suburban residents. Intergenerational transmission of knowledge and improvements in internet search productivity due to the increasing stock of knowledge over time should be investigated.

To sum up, there is much further work to be done to analyze the microeconomic dynamics of knowledge creation in settings with tacit knowledge and multiple modes of communication.

7 Appendix A: The impact of regulating communication modes on joint knowledge productivity

> "The World Health Organization officially declared Covid-19 a pandemic on March 11. Within a few weeks, an estimated 16 million U.S. knowledge workers had switched to working remotely to flatten the curve of the health crisis, according to a new survey by Slack.
>
> This amounts to nearly one-quarter of all knowledge workers in the U.S., and that proportion has climbed even higher as more states have urged citizens to stay home."
>
> Hanson (2020)

During the Covid-19 pandemic, the intensity of use of communication channels between researchers changed, as many switched to work from home, for instance. How does the use of electronic and face to face communication change from an exogenous event, and how does it affect the patterns and volume of knowledge creation? How do pandemic restrictions change the use of communication modes as well as the characteristics of the knowledge creation process?

In this appendix, we analyze the effect of restrictions on face to face communication in our model.

There is a large and rapidly expanding literature on the economic effects of the Covid19 pandemic. Here we focus on *the effects of the pandemic on knowledge creation activity when multiple channels of communication are present*, and in particular on the productivity of researchers under pandemic restrictions. How do they change their choice of joint or individual work, and how do they change their modality of joint work (face to face or internet communication) with the imposition of pandemic restrictions? And how does this interact with the differing commuting cost faced by workers living at various distances from work?

There is some interesting empirical work associated with these questions. Morikawa (2020) finds significant effects, as well as significant heterogeneity in effects, of pandemic restrictions on workers in Japan. For example, worker productivity was reduced by 30-40% when working from home as opposed to commuting to work. We shall discuss in more detail below how further empirical results from this paper support our theory. Inoue et al (2022) examine the effect of the Spanish flu pandemic from the early 20th century on patent productivity in industries where face to face communication was important, and find a huge effect. Of course, this pre-dates the internet. Finally, Yamauchi et al (2022) find a significant negative shock to patent applications as a result of the Covid-19 pandemic. Interestingly, they find that shocks are fine tuned even to the timing of the waves of the pandemic in Japan,[16] suggesting upheavals in the modes of collaborations used in R&D. Gibbs et al (2023) find a big change in worker productivity due to an abrupt imposition of work from home restrictions.

A number of papers consider the effect on productivity of Covid restrictions requiring remote work, related to our theoretical application below where we model this effect. Yang et al (2022) study the effect of such restrictions on Microsoft employees, and find that information sharing decreased. Liu and Su (2023) conclude that working from home reduces the urban wage premium along with agglomeration externalities in cities, consistent with a reduction in returns to interpersonal skills. Thus, a reduction in face to face contact in turn reduces productivity in the conceptual phase of knowledge creation, where tacit knowledge is important. Monte et al (2023) examine how, in a monocentric spatial model with multiple steady state equilibria, a shock to commuting such as Covid can shift the economy from one steady state equilibrium to another, thereby resulting in long term effects. The multiple equilibria are caused by a coordination problem: two workers must be in the office to interact, and if only one shows up, there is no interaction. Thus, there could be equilibria where nobody shows up at the office, and where everyone does. In contrast, we use a myopic core solution concept, so there is no prospect of a noncooperative coordination failure. Thus, the inefficiencies we identify are distinct from theirs. Using micro data that includes the Covid time period, Emanuel et al (2023) find that face to face interaction increases long run human capital at the expense of short term payoffs. Physical proximity is more important to women and mentor-mentee relationships. Gibbs et al (2023) find that productivity fell during the pandemic, as communication over the internet was more costly than face to face communication. In an empirical study, Xiao et al (2021) find that commuting distance decreases the productivity of inventors.

In contrast with this other work, we attempt here to model microfounda-

[16] See their Figure 3.

tions for this literature, *breaking down research into conceptual and technical phases.* By making the communication mode an endogenous choice, and disaggregating this choice into two phases, we are able to address the subtleties of creative activity, such as the role of tacit knowledge.

In the formulation (44), the combination of communication modes, $\{\lambda_{CF}, \lambda_{CN}, \lambda_{\tau F}, \lambda_{\tau N}\}$, has been freely chosen. In reality, however, there exist numerous restrictions on the use of communication modes; some are explicit whereas others are implicit. For example, old metropolises, such as London, New York and Tokyo, have big CBDs that formed a long time ago when commuting railways were predominant. In these metropolises, before the Covid-19 pandemic, a large proportion of workers were forced to commute to CBD offices to ease F2F communications. Likewise, professors and students at traditional universities have been forced to commute to their main campuses for ease of F2F communications. Conversely, during the recent Covid-19 pandemic, many governments introduced regulations that forced a large proportion of workers and students in big cities to work at their homes through the net while discouraging F2F meetings. Specifically, in the early period of the Covid-19 pandemic, the Japanese government asked offices in large cities to reduce worker commuting by 80%, switching to WFH (working from home) to avoid transmission of the corona virus through F2F communications.

Given such examples of restrictions on communication mode in the real world, let us try to apply the model we have developed to study the impact of such restrictions on joint knowledge productivity. For this purpose, focusing on the optimal combination of communication frequencies $(\lambda^*_{CF}, \lambda^*_{CN}, \lambda^*_{\tau F}, \lambda^*_{\tau N})$ in equation (44), we represent the left side of equation (44) explicitly as:

$$A^{J*}_{ij}(\lambda^*_{CF}, \lambda^*_{CN}, \lambda^*_{\tau F}, \lambda^*_{\tau N}). \tag{131}$$

And, given any feasible combination of communication frequencies, $\{\lambda_{CF}, \lambda_{CN}, \lambda_{\tau F}, \lambda_{\tau N}\}$, let us substitute this general combination for the optimized combination $\{\lambda^*_{CF}, \lambda^*_{CN}, \lambda^*_{\tau F}, \lambda^*_{\tau N}\}$ on the right side of equation (44), and denote the resulting equation by

$$A^{J}_{ij}(\lambda_{CF}, \lambda_{CN}, \lambda_{\tau F}, \lambda_{\tau N}). \tag{132}$$

This represents the final output of joint work under the combination of communication frequencies $\{\lambda_{CF}, \lambda_{CN}, \lambda_{\tau F}, \lambda_{\tau N}\}$. Then, taking the ratio of (132) to (131), define:

$$\begin{aligned}\mathcal{H}(\lambda_{CF},\lambda_{CN},\lambda_{\tau F},\lambda_{\tau N}) &\equiv \frac{A_{ij}^{J}(\lambda_{CF},\lambda_{CN},\lambda_{\tau F},\lambda_{\tau N})}{A_{ij}^{J*}(\lambda_{CF}^{*},\lambda_{CN}^{*},\lambda_{\tau F}^{*},\lambda_{\tau N}^{*})} \\ &= \frac{[\lambda_{CF}\cdot\alpha_{CF}+\lambda_{CN}\cdot\alpha_{CN}]^{\rho_{CT}\cdot\rho_{C}}\cdot[\lambda_{\tau F}\cdot\alpha_{\tau F}+\lambda_{\tau N}\cdot\alpha_{\tau N}]^{\rho_{\tau T}\cdot\rho_{\tau}}}{[\lambda_{CF}^{*}\cdot\alpha_{CF}+\lambda_{CN}^{*}\cdot\alpha_{CN}]^{\rho_{CT}\cdot\rho_{C}}\cdot[\lambda_{\tau F}^{*}\cdot\alpha_{\tau F}+\lambda_{\tau N}^{*}\cdot\alpha_{\tau N}]^{\rho_{\tau T}\cdot\rho_{\tau}}},\end{aligned} \tag{133}$$

representing the productivity under the combination $\{\lambda_{CF},\lambda_{CN},\lambda_{\tau F},\lambda_{\tau N}\}$ relative to the optimal combination of communication frequencies. In the rest of this Appendix, we adopt Assumption 1. In order to reduce the number of variables in equation (133), using the feasibility constraints (13) and (33) on frequencies, let us rewrite the terms inside of each bracket in (133) as follows:

$$\begin{aligned}\lambda_{CF}\cdot\alpha_{CF}+\lambda_{CN}\cdot\alpha_{CN} &= \lambda_{CF}\left[\frac{\alpha_{CF}}{\alpha_{CN}}\cdot\frac{1}{1+\varepsilon_F}-1\right]\cdot\alpha_{CN}\cdot(1+\varepsilon_F)+\alpha_{CN} \\ &\equiv f_C(\lambda_{CF},\varepsilon_F),\end{aligned} \tag{134}$$

$$\begin{aligned}\lambda_{\tau F}\cdot\alpha_{\tau F}+\lambda_{\tau N}\cdot\alpha_{\tau N} &= \lambda_{\tau F}\left[\frac{\alpha_{\tau F}}{\alpha_{\tau N}}\cdot\frac{1}{1+\varepsilon_F}-1\right]\cdot\alpha_{\tau N}\cdot(1+\varepsilon_F)+\alpha_{\tau N} \\ &\equiv f_\tau(\lambda_{\tau F},\varepsilon_F),\end{aligned} \tag{135}$$

whereas from Proposition 1,

$$\lambda_{CF}^{*}\cdot\alpha_{CF}+\lambda_{CN}^{*}\cdot\alpha_{CN}=\lambda_{CF}^{*}(\varepsilon_F)\cdot\alpha_{CF}+\lambda_{CF}^{*}(\varepsilon_F)\cdot\alpha_{CN} \tag{136}$$

$$\lambda_{\tau F}^{*}\cdot\alpha_{\tau F}+\lambda_{\tau N}^{*}\cdot\alpha_{\tau N}=\lambda_{\tau F}^{*}(\varepsilon_F)\cdot\alpha_{\tau F}+\lambda_{\tau N}^{*}(\varepsilon_F)\cdot\alpha_{\tau N} \tag{137}$$

where $\{\lambda_{CF}^{*}(\varepsilon_F),\lambda_{CF}^{*}(\varepsilon_F),\lambda_{\tau F}^{*}(\varepsilon_F),\lambda_{\tau N}^{*}(\varepsilon_F)\}$ represents the optimal combination of frequencies at each $\varepsilon_F\geq 0$, as specified by either (50), (51) or (52), depending on which of the three ranges ε_F belongs to. Substituting (134), (135), (136) and (137) into (133), the relative productivity function can be expressed as follows:

$$\begin{aligned}&\mathcal{H}(\lambda_{CF},\lambda_{\tau F};\varepsilon_F) \\ =\;& \frac{f_c(\lambda_{CF},\varepsilon_F)^{\rho_{CT}\cdot\rho_C}\cdot f_\tau(\lambda_{\tau F},\varepsilon_F)^{\rho_{\tau T}\cdot\rho_\tau}}{[\lambda_{CF}^{*}(\varepsilon_F)\cdot\alpha_{CF}+\lambda_{CN}^{*}(\varepsilon_F)\cdot\alpha_{CN}]^{\rho_{CT}\cdot\rho_C}\cdot[\lambda_{\tau F}^{*}(\varepsilon_F)\cdot\alpha_{\tau F}+\lambda_{\tau N}^{*}(\varepsilon_F)\cdot\alpha_{\tau N}]^{\rho_{\tau T}\cdot\rho_\tau}}.\end{aligned} \tag{138}$$

Next, in association with the optimal frequencies $\{\lambda_{CF}^{*}(\varepsilon_F),\lambda_{CF}^{*}(\varepsilon_F),\lambda_{\tau F}^{*}(\varepsilon_F),\lambda_{\tau N}^{*}(\varepsilon_F)\}$ at ε_F, let

$$\lambda_F^{*}(\varepsilon_F)\equiv\rho_{CT}\cdot\rho_C\cdot\lambda_{CF}^{*}(\varepsilon_F)+\rho_{\tau T}\cdot\rho_\tau\cdot\lambda_{\tau F}^{*}(\varepsilon_F) \tag{139}$$

be the *total frequency of F2F communication* per unit of time that has been realized at ε_F either in conceptual development or in technical development. Likewise, in association with any feasible combination of frequencies, $\{\lambda_{CF}, \lambda_{CN}, \lambda_{\tau F}, \lambda_{\tau N}\}$, at any given $\varepsilon_F \geq 0$, let

$$\lambda_F \equiv \rho_{CT} \cdot \rho_C \cdot \lambda_{CF} + \rho_{\tau T} \cdot \rho_\tau \cdot \lambda_{\tau F} \tag{140}$$

be the total frequency of F2F communication.

For concreteness, let us imagine a large city in which a pair of knowledge workers, i and j, reside and work. Let ε_F represent the *commuting time of i and j to the common CBD office* where they work together F2F. Here, ε_F is assumed to be the same for i and j, and it is treated parametrically. Now, reflecting real world experience during the Covid-19 pandemic, let us suppose the following restriction has been imposed (on this pair of workers as well as on all CBD workers) by the government:

Restriction: Given the pair of workers i and j with common commuting time ε_F, their choice of communication mode frequencies, $\{\lambda_{CF}, \lambda_{CN}, \lambda_{\tau F}, \lambda_{\tau N}\}$, must satisfy the following restriction.[17]

$$\lambda_F = \eta \cdot \lambda_F^*(\varepsilon_F) \tag{141}$$

where $\eta \in [0, 1)$ is a constant chosen by the government.

That is, the government requires the workers to reduce the total F2F frequency by $(1 - \eta)\%$ from the present $\lambda_F^*(\varepsilon_F)$, that has been chosen optimally under no restrictions.

We are assuming throughout that communicating using the net means working from home, whereas communicating F2F means working together at the office.

Given restriction (141), the workers must now solve the following problem:

$$\max \Big\{ \ \mathcal{H}(\lambda_{CF}, \lambda_{\tau F}; \varepsilon_F, \eta) \big| \, \rho_{CT} \cdot \rho_C \cdot \lambda_{CF} + \rho_{\tau T} \cdot \rho_\tau \cdot \lambda_{\tau F} = \eta \cdot \lambda_F^*(\varepsilon_F),$$

$$0 \leq \lambda_{CF} \leq \frac{1}{1 + \varepsilon_F}, \ 0 \leq \lambda_{\tau F} \leq \frac{1}{1 + \varepsilon_F} \Big\}. \tag{142}$$

[17] We can generalize restriction (141) as follows:

$$\lambda_F \leq \eta \cdot \lambda_F^*(\varepsilon_F), \ 0 \leq \eta < 1$$

However, since the workers wish to make λ_F as close to $\lambda_F^*(\varepsilon_F)$ as possible, we can replace this inequality with the equality in (141).

Let $\{\hat{\lambda}_{CF}(\varepsilon_F,\eta),\hat{\lambda}_{\tau F}(\varepsilon_F,\eta)\}$ be the solution to problem (142). Then, using constraints (13) and (33), we obtain the associated net frequencies as follows:

$$\hat{\lambda}_{CN}(\varepsilon_F,\eta) = 1-(1+\varepsilon_F)\cdot\hat{\lambda}_{CF}(\varepsilon_F,\eta), \tag{143}$$

$$\hat{\lambda}_{\tau N}(\varepsilon_F,\eta) = 1-(1+\varepsilon_F)\cdot\hat{\lambda}_{\tau F}(\varepsilon_F,\eta). \tag{144}$$

Together,

$$\{\hat{\lambda}_{CF}(\varepsilon_F,\eta),\hat{\lambda}_{CN}(\varepsilon_F,\eta),\hat{\lambda}_{\tau F}(\varepsilon_F,\eta),\hat{\lambda}_{\tau N}(\varepsilon_F,\eta)\} \tag{145}$$

represents the optimal choice of communication-frequencies *under regulation.*

Substituting $\hat{\lambda}_{CF}(\varepsilon_F,\eta)$ and $\hat{\lambda}_{\tau F}(\varepsilon_F,\eta)$ for λ_{CF} and $\lambda_{\tau F}$ in (138), we denote the maximized value of $\mathcal{H}(\lambda_{CF},\lambda_{\tau F};\varepsilon_F)$ by

$$\mathcal{H}(\hat{\lambda}_{CF}(\varepsilon_F,\eta),\hat{\lambda}_{\tau F}(\varepsilon_F,\eta);\varepsilon_F,\eta)\equiv\hat{\mathcal{H}}(\varepsilon_F,\eta). \tag{146}$$

Next, we study the characteristics of solution (145), which depends on where the parameter ε_F belongs in terms of three ranges, (a), (b) and (c), as defined in Figure 5. We start with range (c) since it is easiest. Then we move to range (b), the next easiest. Finally, we examine range (a), the most complex and interesting range.

(c) Consider the *net-dominant the range*, $\left[\varepsilon_F^C,\infty\right)$: From Proposition 1, at each $\varepsilon_F\in\left[\varepsilon_F^C,\infty\right)$, the optimal combination of communication mode frequencies under no restriction is given by

$$\{\lambda^*_{CF}(\varepsilon_F),\lambda^*_{CN}(\varepsilon_F),\lambda^*_{\tau F}(\varepsilon_F),\lambda^*_{\tau N}(\varepsilon_F)\}=\{0,1,0,1\}. \tag{147}$$

Hence, in this net-dominant range, from definition (139),

$$\lambda^*_F(\varepsilon_F)=0, \tag{148}$$

for $\varepsilon_F\in\left[\varepsilon_F^C,\infty\right)$. Substitution of (148) into (142) yields the following equality constraint:

$$\rho_{CT}\cdot\rho_C\cdot\lambda_{CF}+\rho_{\tau T}\cdot\rho_\tau\cdot\lambda_{\tau F}=0 \;\text{ for any } \eta\in[0,1).$$

which has a unique feasible (i.e., non-negative) solution such that

$$\hat{\lambda}_{CF}(\varepsilon_F,\eta)=0,\quad \hat{\lambda}_{\tau F}(\varepsilon_F,\eta)=0,$$

and hence, from (143) and (144),

$$\hat{\lambda}_{CN}(\varepsilon_F,\eta)=1,\quad \hat{\lambda}_{\tau N}(\varepsilon_F,\eta)=1.$$

In sum, the unique solution to problem (142) is given by

$$\begin{aligned}
& \{\hat{\lambda}_{CF}(\varepsilon_F,\eta), \hat{\lambda}_{CN}(\varepsilon_F,\eta), \hat{\lambda}_{\tau F}(\varepsilon_F,\eta), \hat{\lambda}_{\tau N}(\varepsilon_F,\eta)\} \\
= \; & \{0,1,0,1\} \\
= \; & \{\lambda^*_{CF}(\varepsilon_F), \lambda^*_{CF}(\varepsilon_F), \lambda^*_{\tau F}(\varepsilon_F), \lambda^*_{\tau N}(\varepsilon_F)\}, \qquad (149)
\end{aligned}$$

for $\varepsilon_F \in \left[\varepsilon_F^C, \infty\right)$ and $\eta \in [0,1)$. Hence, recalling (133) and (138), we can conclude that

$$\hat{\mathcal{H}}(\varepsilon_F,\eta) = 1 \;\text{ for } \varepsilon_F \in [\varepsilon_F, \infty)\,, \eta \in [0,1). \qquad (150)$$

In other words, in the net-dominant range of ε_F, the restriction of the total F2F frequency in (141) has no effect on the modal choice of communication.

(b) Consider the range, $\varepsilon_F \in \left[\varepsilon_F^\tau, \varepsilon_F^C\right)$, of *mixed communication mode*: From Proposition 1, at each $\varepsilon_F \in \left[\varepsilon_F^\tau, \varepsilon_F^C\right)$, the optimal combination of communication mode frequencies under no restriction is given by

$$\{\lambda^*_{CF}(\varepsilon_F), \lambda^*_{CN}(\varepsilon_F), \lambda^*_{\tau F}(\varepsilon_F), \lambda^*_{\tau N}(\varepsilon_F)\} = \left\{\frac{1}{1+\varepsilon_F}, 0, 0, 1\right\} \qquad (151)$$

implying that

$$\lambda^*_F(\varepsilon_F) = \rho_{CT}\cdot\rho_C\cdot\lambda^*_{CF}(\varepsilon_F) = \frac{\rho_{CT}\cdot\rho_C}{1+\varepsilon_F}. \qquad (152)$$

Thus, the maximization problem (142) becomes:

$$\max\left\{\mathcal{H}(\lambda_{CF},\lambda_{\tau F};\varepsilon_F,\eta) \mid \rho_{CT}\cdot\rho_C\cdot\lambda_{CF} + \rho_{\tau T}\cdot\rho_\tau\cdot\lambda_{\tau F} = \frac{\eta\cdot\rho_{CT}\cdot\rho_C}{1+\varepsilon_F},\right.$$

$$\left. 0 \le \lambda_{CF} \le \frac{1}{1+\varepsilon_F},\; 0 \le \lambda_{\tau F} \le \frac{1}{1+\varepsilon_F}\right\}. \qquad (153)$$

Observing that $\partial\mathcal{H}/\partial\lambda_{CF} > 0$ for $\lambda_{CF} \in (0, 1/(1+\varepsilon_F))$ when $\varepsilon_F \in \left[\varepsilon_F^\tau, \varepsilon_F^C\right)$, we can readily obtain the solution of this maximization problem as follows:

$$\{\hat{\lambda}_{CF}(\varepsilon_F,\eta), \hat{\lambda}_{CN}(\varepsilon_F,\eta), \hat{\lambda}_{\tau F}(\varepsilon_F,\eta), \hat{\lambda}_{\tau N}(\varepsilon_F,\eta)\} = \left\{\frac{\eta}{1+\varepsilon_F}, 1-\eta, 0, 1\right\} \qquad (154)$$

for $\varepsilon_F \in \left[\varepsilon_F^\tau, \varepsilon_F^C\right)$ and $\eta \in [0,1)$, which in turn implies that

$$\begin{aligned}
\hat{\mathcal{H}}(\varepsilon_F,\eta) &= \left[\eta + (1+\varepsilon_F)\cdot(1-\eta)\cdot\frac{\alpha_{CN}}{\alpha_{CF}}\right]^{\rho_{CT}\cdot\rho_C} \\
&= \left\{\eta\cdot\left[1 - \frac{(1+\varepsilon_F)\cdot\alpha_{CN}}{\alpha_{CF}}\right] + \frac{(1+\varepsilon_F)\cdot\alpha_{CN}}{\alpha_{CF}}\right\}^{\rho_{CT}\cdot\rho_C} \\
\text{for } \varepsilon_F &\in \left[\varepsilon_F^\tau, \varepsilon_F^C\right) \text{ and } \eta \in [0,1) \qquad (155)
\end{aligned}$$

Hence, we have that for $\varepsilon_F \in \left[\varepsilon_F^{\tau}, \varepsilon_F^{C}\right)$ and $\eta \in [0,1)$,

$$\partial\hat{\mathcal{H}}(\varepsilon_F,\eta)/\partial\varepsilon_F > 0,\ \hat{\mathcal{H}}(\varepsilon_F^C,\eta) = 1 \ \text{ for any } \eta \in [0,1), \tag{156}$$

$$\partial\hat{\mathcal{H}}(\varepsilon_F,\eta)/\partial\eta > 0,\ \hat{\mathcal{H}}(\varepsilon_F,0) = \left[\frac{(1+\varepsilon_F)\cdot\alpha_{CN}}{\alpha_{CF}}\right]^{\rho_{CT}\cdot\rho_C},\ \hat{\mathcal{H}}(\varepsilon_F,1) = 1. \tag{157}$$

Comparing (151) and (154), we can see that in the (b) range of mixed-communication modes, imposing a restriction on the total F2F frequency in the form of (141) reduces, at each $\varepsilon_F \in \left[\varepsilon_F^{\tau}, \varepsilon_F^{C}\right)$, the F2F frequency in the conceptual development phase from the optimal frequency, $1/(1+\varepsilon_F)$, to $\eta/(1+\varepsilon_F)$. In turn, this increases the net use frequency for conceptual development from 0 to $1-\eta$. As indicated in (156) and (157), the impact of such changes in communication frequencies on the relative productivity $\hat{\mathcal{H}}(\varepsilon_F,\eta)$ is *smaller* (i.e., a relatively smaller reduction in relative productivity) either as ε_F increases or as η increases (i.e., the restriction is weaker).

(a) Consider the *F2F dominant range*, $\varepsilon_F \in [0, \varepsilon_F^{\tau})$: For simplicity of notation in the analysis of this case, let us assume that[18]

$$\rho_{CT}\cdot\rho_C = \rho_{\tau T}\cdot\rho_\tau \equiv \sigma_T. \tag{158}$$

From Proposition 1, at each $\varepsilon_F \in [0, \varepsilon_F^{\tau})$, the optimal combination of communication frequencies under no restriction is given by

$$\{\lambda_{CF}^*(\varepsilon_F), \lambda_{CN}^*(\varepsilon_F), \lambda_{\tau F}^*(\varepsilon_F), \lambda_{\tau N}^*(\varepsilon_F)\} = \left\{\frac{1}{1+\varepsilon_F}, 0, \frac{1}{1+\varepsilon_F}, 0\right\}, \tag{159}$$

Recalling definition (139) and using (157) yields

$$\lambda_F^*(\varepsilon_F) = \frac{2\sigma_T}{1+\varepsilon_F},$$

and hence the restriction (141) becomes

$$\sigma_T(\lambda_{CF}+\lambda_{\tau F}) = \eta\cdot\frac{2\sigma_T}{1+\varepsilon_F},$$

or

$$\lambda_{CF}+\lambda_{\tau F} = \frac{2\eta}{1+\varepsilon_F}. \tag{160}$$

Thus, the maximization problem (142) becomes:

$$\max\left\{\mathcal{H}(\lambda_{CF},\lambda_{\tau F};\varepsilon_F,\eta)\middle|\ \lambda_{CF}+\lambda_{\tau F} = \frac{2\eta}{1+\varepsilon_F}, 0 \le \lambda_{CF} \le \frac{1}{1+\varepsilon_F},\right.$$

[18] Without this assumption, we can conduct the analysis in this section and obtain qualitatively the same result. However, without this assumption, the notation in the analysis becomes awfully complex.

$$0 \leq \lambda_{\tau F} \leq \frac{1}{1+\varepsilon_F}, \Bigg\}. \tag{161}$$

To obtain the solution of this maximization problem, let us neglect for a while the inequality constraints in (161), and consider the following simpler problem:

$$\max \left\{ \mathcal{H}(\lambda_{CF}, \lambda_{\tau F}; \varepsilon_F, \eta) | \, \lambda_{CF} + \lambda_{\tau F} = \frac{2\eta}{1+\varepsilon_F} \right\}. \tag{162}$$

To get the first-order condition for the solution of (162), let us set

$$\left. \frac{\partial \mathcal{H}(\lambda_{CF}, \lambda_{\tau F}; \varepsilon_F, \eta)}{\partial \lambda_{CF}} \right|_{\lambda_{\tau F} = \frac{2\eta}{1+\varepsilon_F} - \lambda_{CF}} = 0. \tag{163}$$

From straight forward calculation using (138), we can obtain the unique solution of (163) as follows:

$$\tilde{\lambda}_{CF}(\varepsilon_F, \eta) \equiv \frac{1}{1+\varepsilon_F} \cdot \left\{ \eta + \frac{q(\varepsilon_F)}{2} \right\}, \tag{164}$$

where

$$q(\varepsilon_F) \equiv \frac{v_C(\varepsilon_F) - v_\tau(\varepsilon_F)}{v_C(\varepsilon_F) \cdot v_\tau(\varepsilon_F)}, \tag{165}$$

$$v_C(\varepsilon_F) \equiv \frac{\alpha_{CF}}{\alpha_{CN}} \cdot \frac{1}{1+\varepsilon_F} - 1, \;\; v_\tau(\varepsilon_F) \equiv \frac{\alpha_{\tau F}}{\alpha_{\tau N}} \cdot \frac{1}{1+\varepsilon_F} - 1. \tag{166}$$

From Assumption 1, we always have for $\varepsilon_F \in [0, \varepsilon_F^\tau)$ that

$$v_C(\varepsilon_F) > 0, \;\; v_\tau(\varepsilon_F) > 0, \;\; q(\varepsilon_F) > 0, \tag{167}$$

and hence $\tilde{\lambda}_{CF}(\varepsilon_F, \eta) > 0$. Furthermore, it can be readily shown that *the left hand side of equation (163) is positive (respectively, negative) for $\lambda_{CF} < \tilde{\lambda}_{CF}$ (respectively, $\lambda_{CF} > \tilde{\lambda}_{CF}$).* Hence, the function, $\mathcal{H}(\lambda_{CF}, \lambda_{\tau F}; \varepsilon_F, \eta)$ subject to $\lambda_{\tau F} = 2\eta/(1+\varepsilon_F) - \lambda_{CF}$, achieves its maximum at $\tilde{\lambda}_{CF}$. However, in general, for $\{\lambda_{CF}, \lambda_{\tau F} = [2\eta/(1+\varepsilon_F) - \lambda_{CF}]\}$ to be the solution to maximization problem (161), it must satisfy also all the inequality conditions in (161). The last inequality condition in (161) can be rewritten as:

$$0 \leq \left\{ \lambda_{\tau F} = \frac{2\eta}{1+\varepsilon_F} - \lambda_{CF} \right\} \leq \frac{1}{1+\varepsilon_F}$$

that is,

$$\lambda_{CF} \leq \frac{2\eta}{1+\varepsilon_F} \text{ and } \frac{2\eta - 1}{1+\varepsilon_F} \leq \lambda_{CF}. \tag{168}$$

Based on the observations above, let $\hat{\lambda}_{CF}(\varepsilon_F, \eta)$ and $\hat{\lambda}_{\tau F}(\varepsilon_F, \eta)$ be defined respectively as

$$\begin{aligned} \hat{\lambda}_{CF}(\varepsilon_F, \eta) &= \min\left\{\frac{1}{1+\varepsilon_F}, \frac{2\eta}{1+\varepsilon_F}, \tilde{\lambda}(\varepsilon_F, \eta)\right\}, && (169) \\ \hat{\lambda}_{\tau F}(\varepsilon_F, \eta) &= \frac{2\eta}{1+\varepsilon_F} - \hat{\lambda}(\varepsilon_F, \eta). && (170) \end{aligned}$$

where $\tilde{\lambda}_{CF}(\varepsilon_F, \eta)$ is defined by (164). Then, it is not difficult to confirm that for $\varepsilon_F \in [0, \varepsilon_F^{\tau})$ and $\eta \in [0, 1)$, *the pair* $\{\hat{\lambda}_{CF}(\varepsilon_F, \eta), \hat{\lambda}_{\tau F}(\varepsilon_F, \eta)\}$ *is the unique solution of the maximization problem* (161) that satisfies, of course, all the constraints in (161), including the second inequality constraint in (168).

Although (169) and (170) together define the solution to problem (161) *implicitly*, to obtain the explicit solution of (161), we must consider the following three cases separately:

$$\text{Case 1:} \quad \tfrac{1}{1+\varepsilon_F} \le \min\left\{\tfrac{2\eta}{1+\varepsilon_F}, \tilde{\lambda}(\varepsilon_F, \eta)\right\}, \qquad (171)$$

$$\text{Case 2:} \quad \tfrac{2\eta}{1+\varepsilon_F} \le \min\left\{\tfrac{1}{1+\varepsilon_F}, \tilde{\lambda}(\varepsilon_F, \eta)\right\}, \qquad (172)$$

$$\text{Case 3:} \quad \tilde{\lambda}(\varepsilon_F, \eta) \le \min\left\{\tfrac{1}{1+\varepsilon_F}, \tfrac{2\eta}{1+\varepsilon_F}\right\}. \qquad (173)$$

Let us study each case in turn.

(i) Case 1 means the following two inequalities hold:

$$\begin{aligned} \bullet \frac{1}{1+\varepsilon_F} &\le \frac{2\eta}{1+\varepsilon_F} \Leftrightarrow \frac{1}{2} \le \eta \\ \bullet \frac{1}{1+\varepsilon_F} &\le \tilde{\lambda}_{CF}(\varepsilon_F, \eta) \equiv \frac{1}{1+\varepsilon_F} \cdot \left\{\eta + \frac{q(\varepsilon_F)}{2}\right\} \Leftrightarrow 1 - \frac{q(\varepsilon_F)}{2} \le \eta \end{aligned}$$

Hence,

$$\text{Case 1} \Leftrightarrow \eta \ge \max\left\{\frac{1}{2}, 1 - \frac{q(\varepsilon_F)}{2}\right\} \qquad (174)$$

On the other hand, from (169) and (170), by using (143) and (144) the optimal combination of communication frequencies under restriction (160) for Case 1 is given by:

$$\{\hat{\lambda}_{CF}(\varepsilon_F, \eta), \hat{\lambda}_{CN}(\varepsilon_F, \eta), \hat{\lambda}_{\tau F}(\varepsilon_F, \eta), \hat{\lambda}_{\tau N}(\varepsilon_F, \eta)\} = \left\{\frac{1}{1+\varepsilon_F}, 0, \frac{2\eta - 1}{1+\varepsilon_F}, 2(1-\eta)\right\}. \qquad (175)$$

Comparing *the optimal combination* (159) *under no restrictions* and *the restricted optimum* (175) above, we can see that

$$\{\lambda^*_{CF}(\varepsilon_F, \eta), \lambda^*_{CN}(\varepsilon_F, \eta)\} = \{\hat{\lambda}_{CF}(\varepsilon_F, \eta), \hat{\lambda}_{CN}(\varepsilon_F, \eta)\} = \left\{\frac{1}{1+\varepsilon_F}, 0\right\} \qquad (176)$$

$$\lambda^*_{\tau F}(\varepsilon_F,\eta) = \frac{1}{1+\varepsilon_F} > \frac{2\eta-1}{1+\varepsilon_F} = \hat{\lambda}_{\tau F}(\varepsilon_F,\eta), \tag{177}$$

$$\lambda^*_{\tau N}(\varepsilon_F,\eta) = 0 < 2(1-\eta) = \hat{\lambda}_{\tau N}(\varepsilon_F,\eta). \tag{178}$$

That is, under a restriction in the form of (160), the optimal frequency-combination in the conceptual development phase is maintained; whereas in the technical development phase, the restriction (160) reduces the F2F-frequency while increasing the net-frequency. This result arises for the following reason: From Assumption 1, F2F communication is more effective in the conceptual development phase than in the technical development phase. On the other hand, condition (174) implies that $\eta \geq \frac{1}{2}$, i.e., the restriction is not strong. Thus, it is efficient to absorb the impact of the total F2F-frequency restriction solely in the technical development phase.

Next, based on (159) and (175), we have that

$$\hat{\mathcal{H}}(\varepsilon_F,\eta) = \left[(2\eta-1) + \frac{\alpha_{\tau N}}{\alpha_{\tau F}}\cdot(1+\varepsilon_F)\cdot 2\cdot(1-\eta)\right]^{\sigma_T} \tag{179}$$

for $\varepsilon_F \in [0,\varepsilon_F^\tau)$ under condition (174). Hence,

$$\frac{\partial\hat{\mathcal{H}}(\varepsilon_F,\eta)}{\partial\eta} > 0,\ \hat{\mathcal{H}}(\varepsilon_F,1)=1,\ \hat{\mathcal{H}}(\varepsilon_F,\frac{1}{2}) = \left[\frac{\alpha_{\tau N}}{\alpha_{\tau F}}\cdot(1+\varepsilon_F)\right]^{\sigma_T} < 1 \tag{180}$$

$$\frac{\partial\hat{\mathcal{H}}(\varepsilon_F,\eta)}{\varepsilon_F} > 0,\ \hat{\mathcal{H}}(0,\eta) = \left[(2\eta-1)+\frac{\alpha_{\tau N}}{\alpha_{\tau F}}\cdot 2\cdot(1-\eta)\right]^{\sigma_T} < 1, \tag{181}$$

for $\eta < 1$ as expected.

(ii) Case 2 means the following two inequalities hold:

- $\dfrac{2\eta}{1+\varepsilon_F} \leq \dfrac{1}{1+\varepsilon_F} \Leftrightarrow \eta \leq \dfrac{1}{2},$
- $\dfrac{2\eta}{1+\varepsilon_F} \leq \tilde{\lambda}(\varepsilon_F,\eta) \equiv \dfrac{1}{1+\varepsilon_F}\cdot\left\{\eta + \dfrac{q(\varepsilon_F)}{2}\right\} \Leftrightarrow \eta \leq \dfrac{q(\varepsilon_F)}{2}.$

Hence,

$$\text{Case 2} \Leftrightarrow \eta \leq \frac{1}{2} \text{ and } \eta \leq \frac{q(\varepsilon_F)}{2}. \tag{182}$$

On the other hand, from (169) and (170) and using (143) and (144), the optimal combination of communication frequencies under restriction (160) for Case 2 is given by:

$$\{\hat{\lambda}_{CF}(\eta,\varepsilon_F),\hat{\lambda}_{CN}(\eta,\varepsilon_F),\hat{\lambda}_{\tau F}(\eta,\varepsilon_F),\hat{\lambda}_{\tau N}(\eta,\varepsilon_F)\} = \left\{\frac{2\eta}{1+\varepsilon_F}, 1-2\eta, 0, 1\right\} \tag{183}$$

Comparing *the optimal combination* (159) *under no restrictions* and *the restricted optimum* above, we can see that *a strong restriction* in the form of (160) with $\eta \leq \frac{1}{2}$ elicits huge changes in communication mode: First, given that F2F communication can be used more effectively in the conceptual development phase than in the technical development phase, the F2F frequency in the technical development phase is reduced from $1/(1+\varepsilon_F)$ ($\equiv$ the non restricted optimum) to 0, whereas the net frequency in the technical development phase is increased from 0 to 1. Such changes in the communication frequencies in the technical development phase allows the full permissible capacity of F2F frequency to be utilized in the conceptual development phase.

Next, using (159) and (183), we have that

$$\hat{\mathcal{H}}(\varepsilon_F, \eta) = \left[2\eta + (1-2\eta) \cdot \frac{\alpha_{CN}}{\alpha_{CF}} \cdot (1+\varepsilon_F)\right]^{\sigma_T} \cdot \left[\frac{\alpha_{\tau N}}{\alpha_{\tau F}} \cdot (1+\varepsilon_F)\right]^{\sigma_T}, \quad (184)$$

for $\varepsilon_F \in [0, \varepsilon_F^{\tau})$ under condition (182). Thus,

$$\begin{aligned}
\frac{\partial \hat{\mathcal{H}}(\varepsilon_F, \eta)}{\partial \eta} &> 0, \ \hat{\mathcal{H}}(\varepsilon_F, \frac{1}{2}) = \left[\frac{\alpha_{\tau N}}{\alpha_{\tau F}} \cdot (1+\varepsilon_F)\right]^{\sigma_T}, \\
\hat{\mathcal{H}}(\varepsilon_F, 0) &= \left[\frac{\alpha_{CN}}{\alpha_{CF}} \cdot (1+\varepsilon_F)\right]^{\sigma_T} \cdot \left[\frac{\alpha_{\tau N}}{\alpha_{\tau F}} \cdot (1+\varepsilon_F)\right]^{\sigma_T}, \qquad (185) \\
\frac{\partial \hat{\mathcal{H}}(\varepsilon_F, \eta)}{\varepsilon_F} &> 0, \ \hat{\mathcal{H}}(0, \eta) = \left[2\eta + (1-2\eta) \cdot \frac{\alpha_{CN}}{\alpha_{CF}}\right]^{\sigma_T} \cdot \left[\frac{\alpha_{\tau N}}{\alpha_{\tau F}}\right]^{\sigma_T}. \qquad (186)
\end{aligned}$$

for $\eta < \frac{1}{2}$.

(iii) Case 3. Recalling (164), Case 3 means the following two inequalities hold:

$$\begin{aligned}
\bullet \tilde{\lambda}(\varepsilon_F, \eta) &\leq \frac{1}{1+\varepsilon_F} \Leftrightarrow \eta \leq 1 - \frac{q(\varepsilon_F)}{2}, \\
\bullet \tilde{\lambda}(\varepsilon_F, \eta) &\leq \frac{2\eta}{1+\varepsilon_F} \Leftrightarrow \frac{q(\varepsilon_F)}{2} \leq \eta.
\end{aligned}$$

Thus,

$$\text{Case 3} \Leftrightarrow \frac{q(\varepsilon_F)}{2} \leq \eta \leq 1 - \frac{q(\varepsilon_F)}{2} \qquad (187)$$

implying that $q(\varepsilon_F) \leq \frac{1}{2}$.

On the other hand, from (169) and (170), using (143) and (144), the optimal combination of communication frequencies under restriction (160) for Case 3 is given by

$$\begin{aligned}
&\{\hat{\lambda}_{CF}(\eta, \varepsilon_F), \hat{\lambda}_{CN}(\eta, \varepsilon_F), \hat{\lambda}_{\tau F}(\eta, \varepsilon_F), \hat{\lambda}_{\tau N}(\eta, \varepsilon_F)\} \\
= &\left\{\frac{1}{1+\varepsilon_F} \cdot \left(\eta + \frac{q(\varepsilon_F)}{2}\right), 1 - \left(\eta + \frac{q(\varepsilon_F)}{2}\right), \frac{1}{1+\varepsilon_F} \cdot \left(\eta - \frac{q(\varepsilon_F)}{2}\right), 1 - \left(\eta - \frac{q(\varepsilon_F)}{2}\right)\right\}
\end{aligned} \qquad (188)$$

implying that in the domain of $\{\eta, q(\varepsilon_F)\}$ where the strict inequalities hold in relation (187), *each component in (188) is strictly positive*; that is, each frequency combination given by (188) represents an *interior solution.* In contrast, the optimal combination (159) under no restriction represents a *corner solution.* The constrained optimum (188) implies that in order to absorb effectively the constraint (160) on the total F2F frequency, in each of the phases of conceptual development and technical development, only a part of F2F frequency in the optimal solution (159) has been switched to the net. When we compare the three cases below, it will become clear that Case 3 tends to happen when F2F communication is much more effective than net communication in both the phase of conceptual development and the phase of technical development.

Next, based on (159) and (188), we have that

$$\begin{aligned} \hat{\mathcal{H}}(\varepsilon_F, \eta) &= \left[\left(\eta + \frac{q(\varepsilon_F)}{2}\right) \cdot \left(1 - \frac{\alpha_{CN}}{\alpha_{CF}}\right) \cdot (1 + \varepsilon_F) + \frac{\alpha_{\tau N}}{\alpha_{\tau F}} \cdot (1 + \varepsilon_F)\right]^{\sigma_T} \\ &\quad \times \left[\left(\eta - \frac{q(\varepsilon_F)}{2}\right) \cdot \left(1 - \frac{\alpha_{\tau N}}{\alpha_{\tau F}}\right) \cdot (1 + \varepsilon_F) + \frac{\alpha_{\tau N}}{\alpha_{\tau F}} \cdot (1 + \varepsilon_F)\right]^{\sigma_T}, \end{aligned} \tag{189}$$

for $\varepsilon_F \in [0, \varepsilon_F^\tau)$ and under condition (187). Hence,

$$\frac{\partial \hat{\mathcal{H}}(\varepsilon_F, \eta)}{\partial \eta} > 0. \tag{190}$$

It is also not difficult to see from (189) that

$$\frac{\partial \hat{\mathcal{H}}(\varepsilon_F, \eta)}{\varepsilon_F} > 0, \tag{191}$$

under condition (187).

(iv) Synthesis. Thus far, we have examined each case separately. Putting the results of the three cases together, Figure A shows which case happens under what range of parameters.

Figure A

In the right half of Figure A, the (η, q)-space is partitioned into three cases. In the left half of Figure A, the $q(\varepsilon_F)$-curve is depicted by taking two examples of α-parameters:

$$\text{Example 1:} \qquad \frac{\alpha_{CN}}{\alpha_{CF}} = 2, \ \frac{\alpha_{\tau N}}{\alpha_{\tau F}} = 1.5 \tag{192}$$

$$\text{Example 2:} \qquad \frac{\alpha_{CN}}{\alpha_{CF}} = 3, \ \frac{\alpha_{\tau N}}{\alpha_{\tau F}} = 2. \tag{193}$$

First, let us consider Example 1. Substituting the parameter values in Example 1 into (165) and (166) yields:

$$q_1(\varepsilon_F) = \frac{\frac{1/2}{1+\varepsilon_F}}{(\frac{2}{1+\varepsilon_F} - 1) \cdot (\frac{3/2}{1+\varepsilon_F} - 1)}, \tag{194}$$

implying that

$$q_1(0) = 1, \; q_1'(\varepsilon_F) > 0, \; q_1\left(\frac{1}{2}\right) = \infty,$$

as depicted in Figure A. Since $q_1(\varepsilon_F) > 1$ for all $\varepsilon_F \in (0, 1/2)$, only two cases can happen: Case 1 when $\eta \in [1/2, 1)$, and Case 2 when $\eta \in [0, 1/2)$. Notice that the parameter values in Example 1 correspond to those in Figure 4.

Next, let us consider Example 2, which yields the following function:

$$q_2(\varepsilon_F) = \frac{\frac{1}{1+\varepsilon_F}}{(\frac{3}{1+\varepsilon_F} - 1) \cdot (\frac{2}{1+\varepsilon_F} - 1)}, \tag{195}$$

implying that

$$q_2(0) = \frac{1}{2}, \; q_2'(\varepsilon_F) > 0, \; q_2(0.268) = 1, \; q_2(1) = \infty,$$

as shown in Figure A. Since $q_2(\varepsilon_F) < 1$ for $\varepsilon_F \in [0, 0.268)$, Case 3 can happen for each $\varepsilon_F \in [0, 0.268)$ with intermediate values of η. Notice that the parameter values in Example 2 correspond to those for Case 2 in Figure 5. As discussed in the last paragraph of Section 3, relative values of α's in Example 2 (i.e., Case 2 in Figure 5) correspond to those knowledge workers who have poor internet skills (or poor internet access). For such workers, in both the phases of conceptual development and technical development, F2F communications are much more effective than net communication; hence, these workers retain a large fraction of F2F frequencies in both phases when η is an intermediate value.

Finally, let us recall that the *relative productivity function*, $\hat{\mathcal{H}}(\varepsilon_F, \eta)$, has been defined by (146), which shows the ratio of {the maximized value of joint work output at ε_F under the F2F frequency constraint η} over {the maximized value of joint work output at ε_F without a frequency constraint}. We have derived the following set of relative productivity functions:

(c) function (150) for the *net dominant range*, $[\varepsilon_F^\tau, \infty)$,

(b) function (155) for the *mixed communication-mode range*, $\varepsilon_F \in \left[\varepsilon_F^\tau, \varepsilon_F^C\right)$,

(a) for the *F2F dominant range*, $\varepsilon_F \in [0, \varepsilon_F^\tau)$;
function (179) for Case 1,
function (184) for Case 2,
function (189) for Case 3,

Based on these relative productivity functions, Figure B presents a set of relative productivity curves $\hat{\mathcal{H}}(\varepsilon_F, \eta)$ as a function of the total F2F frequency constraint η at each selected value of ε_F.

Figure B

In drawing Figure B, the α-parameters are based on those in Example 2 in (193), implying that

$$\varepsilon_F^{\tau} = 1, \ \varepsilon_F^{C} = 2.$$

Furthermore, the values of the ρ-parameters are set as

$$\rho_{CT} \cdot \rho_C = \rho_{\tau T} \cdot \rho_{\tau} = 0.25.$$

Drawing on Figure B, we make the following observations:

• All relative productivity curves meet at $(1, 1)$ since $\eta = 1$ means no F2F frequency constraint.

• In the *net-dominant range* $\varepsilon_F \in \left[\varepsilon_F^C, \infty\right)$, it holds that $\hat{\mathcal{H}}(\varepsilon_F, \eta) = 1$ for all $\eta \in [0, 1]$ since any F2F frequency constraint is not binding in the net dominant range.

• As ε_F *decreases* from ε_F^C toward 0, the relative productivity curve rotates counterclockwise around the pivot $(1, 1)$, implying that

$$\begin{aligned} \frac{\partial \hat{\mathcal{H}}(\varepsilon_F, \eta)}{\partial \varepsilon_F} &> 0 \ \text{ for } \varepsilon_F \in \left(0, \varepsilon_F^C\right), && (196) \\ \frac{\partial \hat{\mathcal{H}}(\varepsilon_F, \eta)}{\partial \eta} &> 0 \ \text{ for } \ \eta \in (0, 1). && (197) \end{aligned}$$

• Relation (196) arises for the following reason. As can be seen from (50), the optimal values of λ_{CF}^* and $\lambda_{\tau F}^*$ decrease as ε_F increases; likewise, from (51), the value of λ_{CF}^* decreases as ε_F increases. That is, when there is no constraint on the total F2F meeting frequency, communication in joint work depends less on F2F as ε_F increases. Hence, as ε_F increases, the negative impact of a given total frequency constraint η on the joint work output is relatively smaller. Conversely, as ε_F *decreases toward* 0, *the negative impact of a given constraint* η *becomes relatively greater*; and hence *the relative productivity* $\hat{\mathcal{H}}(\varepsilon_F, \eta)$ *becomes lower as* ε_F *decreases toward* 0.

• Relation (197) arises for the simple reason that when ε_F is fixed, the negative impact of F2F constraint η on the joint-work output becomes weaker as η

increases (i.e., the constraint becomes less binding), and hence *the relative productivity $\hat{\mathcal{H}}(\varepsilon_F, \eta)$ increases as η increases towards* 1.

• Based on (196) and (197), we see that in Figure B, *the relative productivity $\hat{\mathcal{H}}(\varepsilon_F, \eta)$ decreases most significantly in the F2F-dominant range (a) when the value of η is small (i.e., the total F2F constraint is strong)*. Specifically, we can see from Figure B that at $(\eta, \varepsilon_F) = (0, 0)$, the relative productivity $\hat{\mathcal{H}}(0, 0)$ equals 0.64. That is, *when no F2F communication is permitted (i.e., all communications are forced to be through the net only), and when $\varepsilon_F = 0$ (i.e., no lead-time for F2F communication), the productivity of joint work decreases 36% in comparison with the situation without a total-F2F-constraint.* Under the specification $(\varepsilon_F, \eta) = (0, 1/4)$, the relative productivity $\hat{\mathcal{H}}(0, 1/4)$ is 0.76, i.e., a 24% reduction in joint-work productivity.

• As shown in Figure B, in the F2F dominant range (a), there are three cases of optimal combinations of communication-mode frequencies under a total F2F constraint. In the area for Case 3 (below the doted boundary line curve), optimal frequencies of F2F and net are both positive in both the phases of conceptual development and technical development. In this way, the negative impact of a total F2F constraint on joint-work productivity is mitigated most effectively.

• Recall that in the context of a metropolitan area, ε_F represents the commuting time to the common CBD office. Thus, the results above indicate that under the same total F2F restriction η, workers residing near the CBD see their joint-work productivity decrease relatively more than those residing far from the CBD.

• Figure B is based on the relative values of α's given in Example 2 of (193). In contrast, let the relative values of α's be set at those given in Example 1 of (192). Then, it can be readily shown that $\hat{\mathcal{H}}(0, 0) = 0.75$. Thus, in terms of Figure B, the lowest relative productivity point, $\hat{\mathcal{H}}(0, 0)$, moves upward to 0.75 from the present 0.64. Similarly, when relative values of α's are changed from those in Example 2 to those in Example 1, all relative productivity curves in Figure B shift upward significantly for all $\eta < 1$. Recall that relative values of α's in Example 2 correspond to workers who have poor computer skills (or poor internet access). In contrast, relative values of α's in Example 1 correspond to workers who have better skills (or better internet access).

• Finally, notice that the relative values of α's in Example 1 of (192) (vs. Example 2 of (193)) correspond to those values in Case 1 (vs. Case 2) in Figure 6. Thus, the results mentioned just above suggest that *the negative* *impact of a total F2F constraint on relative productivity becomes weaker with the advancement of net-technology.*

We may summarize our observations as follows:

Proposition A. In the context of a metropolitan area, let us imagine that the government (facing the Covid-19 pandemic) introduces the *restriction* that each pair of joint knowledge workers must reduce the total F2F frequency by $(1-\eta)\%$ from the present optimal choice. To be precise, at each ε_F, the restriction is defined by equation (141) together with (139) and (140). In this context, let $\hat{\mathcal{H}}(\varepsilon_F, \eta)$ be *the relative productivity function* defined by (142) and (146), which represents the ratio of {the maximized value of joint work output at ε_F under the F2F frequency constraint η} over {the maximized value of joint work output at ε_F without a frequency constraint}. For illustration, let the relative values of communication-effectiveness parameters, α's, be set as those in Example 2 of (193), implying that the ε_F-parameter range is divided into the following three sub-ranges:

(c) the net dominant range: $\left[\varepsilon_F^C, \infty\right) = [2, \infty)$,

(b) the range of mixed communication mode: $\left[\varepsilon_F^\tau, \varepsilon_F^C\right) = [1, 2)$,

(a) the F2F dominant range: $[0, \varepsilon_F^\tau) = [0, 1)$.

In this context, the impact of the total F2F frequency restriction on the relative productivity, $\hat{\mathcal{H}}(\varepsilon_F, \eta)$, can be represented as in Figure B. Specifically:

(i) Since $\eta = 1$ means no F2F frequency constraint, all *relative productivity curves*, $\hat{\mathcal{H}}(\varepsilon_F, \eta)$ for each fixed ε_F, meet at $(1, 1)$.

(ii) In the net-dominant range (c), since F2F frequency is not binding, the F2F frequency constraint has no impact on the relative productivity, implying that $\hat{\mathcal{H}}(\varepsilon_F, \eta) = 1$ for all $\varepsilon_F \in \left[\varepsilon_F^C, \infty\right)$ and $\eta \in [0, 1]$.

(iii) As ε_F decreases from ε_F^C toward 0, the relative productivity curve rotates counterclockwise around the pivot $(1, 1)$, implying that $\partial\hat{\mathcal{H}}(\varepsilon_F, \eta) > 0$ for $\varepsilon_F \in (0, \varepsilon_F^C)$ and $\partial\hat{\mathcal{H}}(\varepsilon_F, \eta) > 0$ for $\eta \in (0, 1)$.

(iv) Hence, as can be seen in Figure B, the relative productivity $\hat{\mathcal{H}}(\varepsilon_F, \eta)$ decreases most significantly in the F2F-dominant range, (a), when the value of η is small (i.e., the total F2F constraint is strong). Specifically, we can see from Figure B that at $(\eta, \varepsilon_F) = (0, 0)$, the relative productivity $\mathcal{H}(0, 0)$ equals 0.64; that is, the productivity of joint work decreases 36% in comparison with the situation of no total frequency constraint. For another specification, at $(\varepsilon_F, \eta) = (0, 1/4)$, the relative productivity $\hat{\mathcal{H}}(0, 1/4)$ equals 0.76, implying a 24% reduction in joint-work productivity.

(v) In the context of a metropolitan area, the results above indicate that under the same total F2F restriction η, workers residing near the CBD experience a relatively higher decrease in their joint-work productivity compared with those residing far from the CBD.

(vi) Figure B is based on the relative values of α's given in Example 2 of (193). In contrast, let the relative values of α's be set at those given in Example 1 of (192). Then, it can be readily shown that in Figure B, the lowest relative productivity point, $\hat{\mathcal{H}}(0,0)$, moves upward to 0.75 from the present 0.64; furthermore, all relative productivity curves shift upward significant for all $\eta < 1$. In other words, the relative productivity loss from the same F2F frequency constraint is smaller for workers with better computer skills (or better internet access) than that for workers with poor computer skills (or poor internet access).

(vii) Finally, comparing relative values of α's in Example 1 and Example 2, we can also reinterpret the results above as follows: the negative impact of a total F2F constraint on relative productivity becomes weaker with the advancement of net technology.

8 Appendix B: Calculation of imputed values in Table 1

For example, let us calculate the imputed value of a_{ij}^{CTF} (the top of the third tier in Table 1). Observe that

$$\frac{\partial A_{ij}^{J*}}{\partial a_{ij}^{CTF}} = \frac{\partial A_{ij}^{J*}}{\partial a_{ij}^{C*}} \cdot \frac{\partial a_{ij}^{C*}}{\partial a_{ij}^{CTF}} \tag{198}$$

First, from (43),

$$\begin{aligned} \log A_{ij}^{J*} &= \rho_C \cdot [\log \rho_C + \log a_{ij}^{C*}] + \rho_I \cdot [\log \rho_I + \log(a_{ii}^{I*} + a_{ij}^{I*})] \\ &\quad + \rho_\tau \cdot [\log \rho_\tau + \log a_{ij}^{\tau *}]. \end{aligned}$$

Hence,

$$\frac{\partial A_{ij}^{J*}}{\partial a_{ij}^{C*}} = A_{ij}^{J*} \cdot \rho_C \cdot \frac{1}{a_{ij}^{C*}}. \tag{199}$$

Next, from (19),

$$\begin{aligned} \log a_{ij}^{C*} &= \log \alpha_{ij}^{C} + \rho_{CT} \cdot [\log \rho_{CT} + \log(\lambda_{CF}^{*} \cdot a_{ij}^{CTF} + \lambda_{CN}^{*} \cdot a_{ij}^{CTN})] \\ &\quad + \rho_{CS} \cdot [\log \rho_{CS} + \log(a_{i}^{CSI} + a_{j}^{CSI})]. \end{aligned}$$

Thus,

$$\frac{\partial a_{ij}^{C*}}{\partial a_{ij}^{CTF}} = a_{ij}^{C*} \cdot \frac{\rho_{CT} \cdot \lambda_{CF}^{*}}{\lambda_{CF}^{*} \cdot a_{ij}^{CTF} + \lambda_{CN}^{*} \cdot a_{ij}^{CTN}}. \tag{200}$$

Substituting (199) and (200) into (198) yields:

$$\begin{aligned} \frac{\partial A_{ij}^{J*}}{\partial a_{ij}^{CTF}} &= A_{ij}^{J*} \cdot \rho_C \cdot \frac{1}{a_{ij}^{C*}} . a_{ij}^{C*} \cdot \frac{\rho_{CT} \cdot \lambda_{CF}^{*}}{\lambda_{CF}^{*} \cdot a_{ij}^{CTF} + \lambda_{CN}^{*} \cdot a_{ij}^{CTN}} \\ &= A_{ij}^{J*} \cdot \rho_C \cdot \rho_{CT} \cdot \frac{\lambda_{CF}^{*}}{\lambda_{CF}^{*} \cdot a_{ij}^{CTF} + \lambda_{CN}^{*} \cdot a_{ij}^{CTN}}. \end{aligned}$$

Hence,

$$\frac{\partial A_{ij}^{J*}}{\partial a_{ij}^{CTF}} \cdot a_{ij}^{CTF} = A_{ij}^{J*} \cdot \rho_C \cdot \rho_{CT} \cdot \frac{\lambda_{CF}^{*} \cdot a_{ij}^{CTF}}{\lambda_{CF}^{*} \cdot a_{ij}^{CTF} + \lambda_{CN}^{*} \cdot a_{ij}^{CTN}}$$

which gives the imputed value of a_{ij}^{CTF} listed in the fourth column of Table 1.

Likewise, we can obtain the imputed value of each activity output listed in the fourth column of Table 1.

9 Appendix C: Three possible dynamics of the two-person system

In Figure C, diagrams (a), (b) and (c) illustrate each of three possible dynamics for m^d, whereas diagram (d) explains the determination of the relationship between $\tilde{m}$ and m^B.[19]

Figure C

First, diagram (a) shows the dynamics of m^d for the case, $m^E < \tilde{m} < m^B$, which is also the case illustrated in Figure 9. As depicted in this diagram, if the initial position, $m^d(0)$, at time 0 is to the left of $\tilde{m}$, then $m^d(t)$ gradually increases towards $\tilde{m}$. If $\tilde{m} < m(0) < m^H$, then $m^d(t)$ gradually decreases towards $\tilde{m}$. Only when $m^d(0) > m^H$ does $m^d(t)$ move away from $\tilde{m}$ toward 1/2. Hence, whenever $m^d(0) < m^H$, $m^d(t)$ approaches *the sink point*, $\tilde{m}$, which is to the left of the bliss point m^B.

Next, diagram (b) depicts the dynamics of m^d for the case $m^B < \tilde{m} < m^H$. In this case, except when $m^d(0) > m^H$, $m^d(t)$ approaches the sink point $\tilde{m}$ to the right of the bliss point m^B.

Third, diagram (c) describes the case where $\tilde{m} < m^E$, which happens when $g_J(\tilde{m}) < \Phi_I$. In this case, except when $m^d(0) > m^H$, $m^d(t)$ approaches the sink point m^E, where the K-growth rate is much lower than at the bliss point m^B.

Diagram (d) in Figure C explains when each of the three possible cases of dynamics happens. For the purpose of intuitive understanding of when each case happens, we here focus on the special situation where

$$\theta_C = \theta_\tau \equiv \theta. \tag{C-1}$$

That is, both in equations (10) and (11) and equations (29) and (30) (the K-production function for joint thinking through the Net), the weight on differential knowledge has the same value. In this case, from (86), we have

[19]Theoretically speaking, there exists the possibility of a fourth case where $m^B < m^H < \tilde{m} < 1/2$. This can happen only in the extreme situation where both θ_F and θ_N are close zero, and Φ_I and $g_J(m^B)$ are nearly equal. Hence, in the following discussion, we neglect this fourth case. Actually, it is possible that $\tilde{m} = m^B$. But this case happens on a set of measure zero, so we neglect it as well.

that

$$m^B = \frac{1}{1+\frac{1}{\theta}} \equiv m^B(\theta), \tag{C-2}$$

meaning that the value of m^B is uniquely determined by the value of parameter θ. In diagram (d) of Figure C, function $m^B(\theta)$ is depicted by the bold curve for $\theta \in [0, 1]$.

In contrast, the value of $\tilde{m}$ is determined, as in (113), by the values of ρ's, independent of θ. Hence, here we treat $\tilde{m}$ as a single parameter. By definition, the maximum value of $\tilde{m}$ equals 1/4:

$$\max \tilde{m} = \frac{1}{4}.$$

Given each $\tilde{m} \in (0, 1/4]$, as shown in diagram (d), by setting

$$\tilde{m} = m^B(\theta), \tag{C-3}$$

the value of $\theta(\tilde{m})$ is uniquely determined by $\tilde{m} = m^B$.

Now, given $\tilde{m} \in (0, 1/4]$, we can see from diagram (d) that

$$\tilde{m} < m^B(\theta) \text{ for } \theta \in (\theta(\tilde{m}), 1/2). \tag{C-4}$$

That is, given $\tilde{m} \in (0, 1/4]$, if we take the value of θ sufficiently large so that $\theta > \theta(\tilde{m})$, then the corresponding bliss point, $m^B(\theta)$, locates to the right of $\tilde{m}$, which corresponds to diagrams (a) and (c) in Figure C. By definition, $\theta > \theta(\tilde{m})$ means that *the weight on differential knowledge in knowledge sub-production functions (10) and (29) is sufficiently large.* Furthermore, when θ is close $\theta(\tilde{m})$, then we have diagram (a). On the other hand, when θ is much larger than $\theta(\tilde{m})$, then the corresponding bliss point m^B is far to the right of $\tilde{m}$, yielding diagram (c).

In contrast to (C-4), we can see from diagram (d) that for each $\tilde{m} \in (0, 1/4)$,

$$\tilde{m} > m^B(\theta) \text{ for } \theta \in (0, \theta(\tilde{m})). \tag{C-5}$$

That is, given $\tilde{m} \in (0, 1/4)$, if the value of θ is sufficiently small so that $\theta < \theta(\tilde{m})$, then the corresponding bliss point, $m^B(\theta)$, locates to the left of $\tilde{m}$, which corresponds to diagram (b) in Figure C. By definition, $\theta < \theta(\tilde{m})$ means that *the weight on common knowledge* in sub-production functions (10) and (29) *is sufficiently large.*

References

[1] Aghion, P., Jones, B.F. and C.I. Jones, 2017. "Artificial Intelligence and Economic Growth," in The Economics of Artificial Intelligence: An Agenda, ed. by A. Agrawal, J. Gans and A Goldfarb, pp. 237-282. National Bureau of Economic Research, Cambridge, MA.

[2] Atkin, D., Chen, M.K. and Popov, A., 2022. "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley." Working Paper. http://www.nber.org/papers/w30147

[3] Battiston, D., i Vidal, J.B. and T. Kirchmaier, 2021. "Face-to-Face Communication in Organizations," The Review of Economic Studies 88, No. 2, 574–609.

[4] Berliant, M. and M. Fujita, 2008. "Knowledge Creation as a Square Dance on the Hilbert Cube," International Economic Review 49, No.4, pp.1251-1295.

[5] Berliant, M. and M. Fujita, 2009. "Dynamics of Knowledge Creation and Transfer: The Two Person Case," International Journal of Economic Theory 5, No.2, pp.155-179.

[6] Berliant, M. and M. Fujita, 2011. "The Dynamics of Knowledge Diversity and Economic Growth," Southern Economic Journal 77, No.4, pp.856-884.

[7] Berliant, M. and M. Fujita, 2012. "Culture and Diversity in Knowledge Creation," Regional Science and Urban Economics 42, No.4, pp.648-662.

[8] Bloom, N., Jones, C.I., Van Reenen, J. and M. Webb, 2020. "Are Ideas Getting Harder to Find?" American Economic Review 110, 1104-1144. https://doi.org/10.1257/aer.20180338

[9] Ceci, F., Lazoi, M., Lezzi, M. and H. Mohammad, 2021. "Fostering Knowledge Sharing in the Innovation Process: Information and Communication Technology-Based versus Face-to-Face Relationships," Knowledge Process and Management 28, 302-316.

[10] Emanuel, N., Harrington E. and A. Pallais, 2023. "The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today?" Working Paper. https://ssrn.com/abstract=4638214

[11] Finke, R.A., Ward, T.B. and S.M. Smith, 1992. Creative Cognition: Theory, Research, and Applications. MIT Press: Cambridge, MA.

[12] Gibbs, M., Mengel, F. and C. Siemroth, 2023. "Work from Home and Productivity: Evidence from Personnel and Analytics Data on Information Technology Professionals," Journal of Political Economy Microeconomics 1, 7-41.

[13] Gordon, P., 2013. "Thinking About Economic Growth: Cities, Networks, Creativity and Supply Chains for Ideas," Annals of Regional Science 50, 667-684.

[14] Hanson, R.S., 2020. "Report: Remote work in the age of Covid-19," Slack https://slack.com/intl/zh-tw/blog/collaboration/report-remote-work-during-coronavirus

[15] Inoue, H., Nakajima, K., Okazaki, T. and Y.U. Saito, 2022. "The Role of Face-to-face Contact in Innovation: The Evidence from the Spanish Flu Pandemic in Japan." RIETI Discussion Paper 22-E-026.

[16] Krugman, P., 1991. Geography and Trade. MIT Press: Cambridge, MA.

[17] Lin, Y., Frey, C.B. and L. Wu, 2022. "Remote Collaboration Fuses Fewer Breakthrough Ideas," Nature 623, 987-991. https://doi.org/10.1038/s41586-023-06767-1

[18] Liu, S. and Y. Su, 2023. "The Effect of Working from Home on the Agglomeration Economies of Cities: Evidence from Advertised Wages." Working Paper. https://dx.doi.org/10.2139/ssrn.4109630

[19] Monte, F., Porcher, C. and E. Rossi-Hansberg, 2023. "Remote Work and City Structure." NBER Working Paper 31494. http://www.nber.org/papers/w31494

[20] Morikawa, M., 2020. "Productivity of Working from Home during the COVID-19 Pandemic: Evidence from an Employee Survey." RIETI Discussion Paper Series 20-E-073.

[21] Nonaka, I. and H. Takeuchi, 1995. The Knowledge-Creating Company. Oxford University Press: Oxford.

[22] Polanyi, M., 1958. Personal Knowledge: Towards a Post-Critical Philosophy. University of Chicago Press: Chicago, IL.

[23] Polanyi, M., 1966. The Tacit Dimension. University of Chicago Press: Chicago, IL.

[24] Panahi, S., Watson, J. and H. Partridge, 2012. "Social Media and Tacit Knowledge Sharing: Developing a Conceptual Model," World Academy of Science, Engineering and Technology 64, 1095-1102.

[25] Xiao, H., Wu, A. and J. Kim, 2021. "Commuting and Innovation: Are Closer Inventors More Productive?" Journal of Urban Economics 121, 103300. https://doi.org/10.1016/j.jue.2020.103300

[26] Yamauchi, I., Nagaoka, S. and D. Miyazaki, 2022. "Impacts of COVID-19 on R&D and Patenting Activities in Japan: Demand Shocks, Application Delays, and Patent Option Value." RIETI Policy Discussion Paper 22-P-013.

[27] Yang, L., Holtz, D., Jaffe, S., Suri, S., Sinha, S., Weston, J., Joyce, C., Shah, N., Sherman, K., Hecht, B. and J. Teevan, 2022. "The Effects of Remote Work on Collaboration among Information Workers," Nature Human Behavior 6, 43-54.

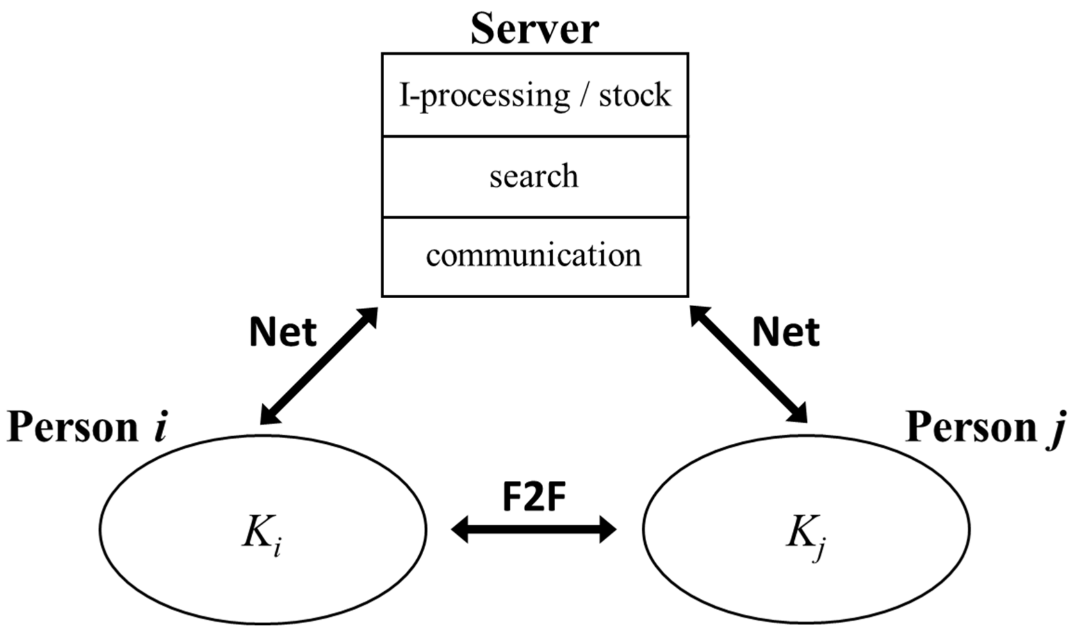

Figure 1. Knowledge creation through multiple modes of communication

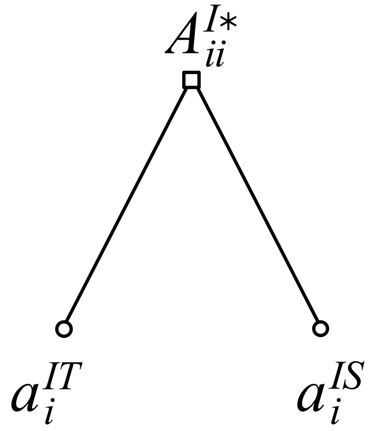

Figure 2. The activity tree for knowledge creation by person i in Isolation.

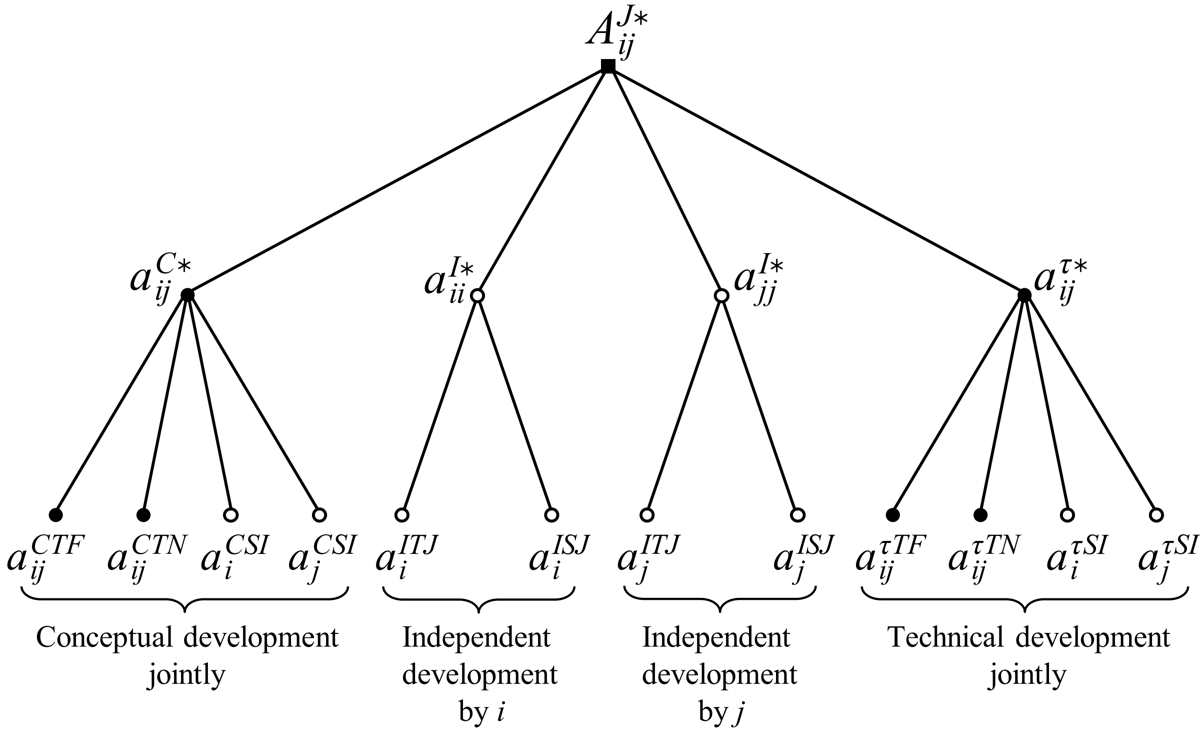

Figure 3. The activity tree for joint knowledge creation:

■ the final output of joint work

● representing a joint activity with subscript ij,

○ representing an independent activity for the purpose of joint creation with subscript i, ii, j or jj.

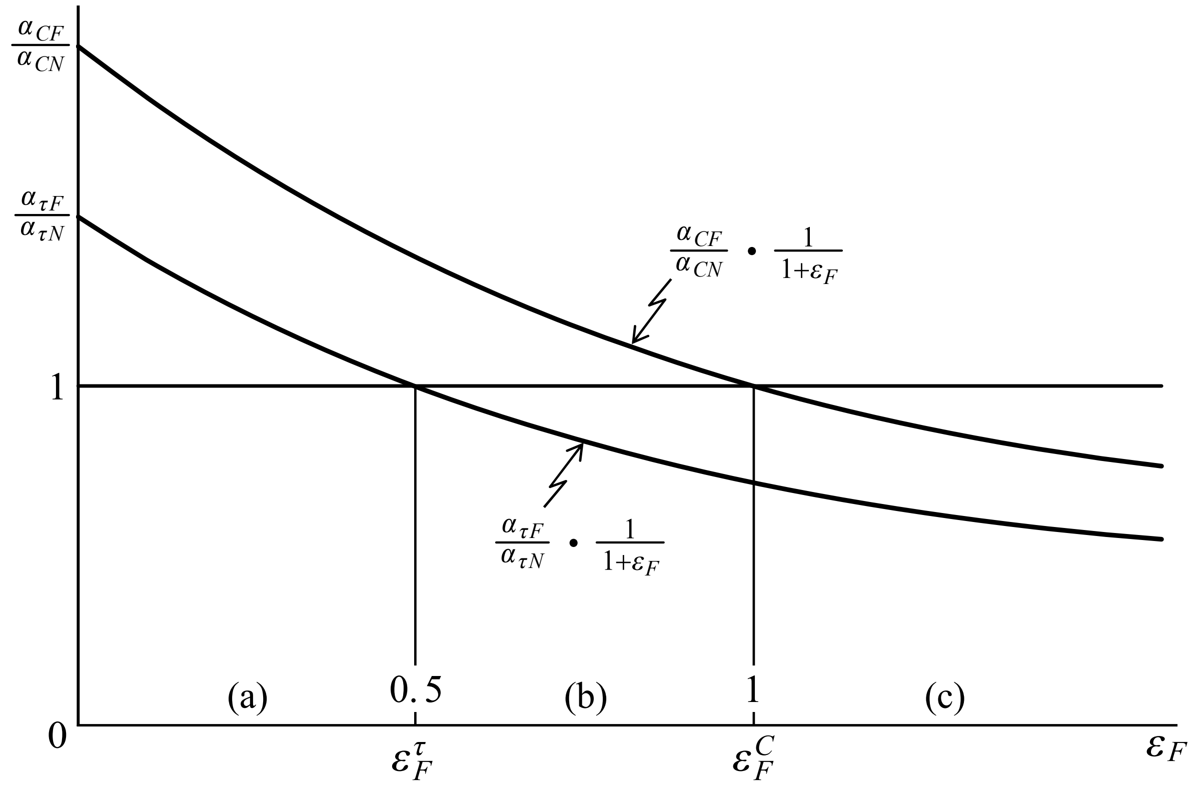

Figure 4. The parameter value of F2F lead-time ε_F and the three ranges of communication mode when $\frac{\alpha_{CF}}{\alpha_{CN}} = 2$ and $\frac{\alpha_{\tau F}}{\alpha_{\tau N}} = 1.5$.

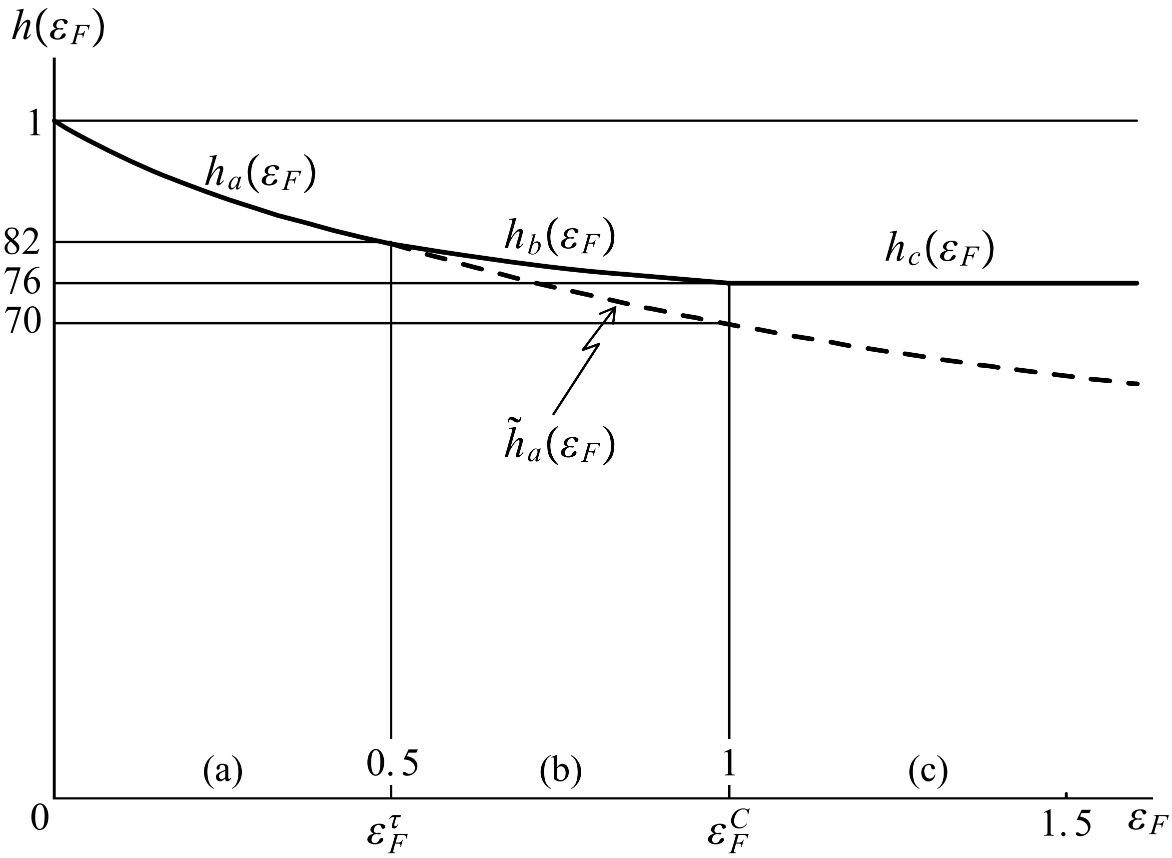

Figure 5. The relative productivity curve $h(\varepsilon_F)$ in the three ranges of ε_F when $\frac{\alpha_{CF}}{\alpha_{CN}} = 2$ and $\frac{\alpha_{\tau F}}{\alpha_{\tau N}} = 1.5$, $\rho_{CT} \cdot \rho_C = \rho_{\tau T} \cdot \rho_\tau = 0.25$.

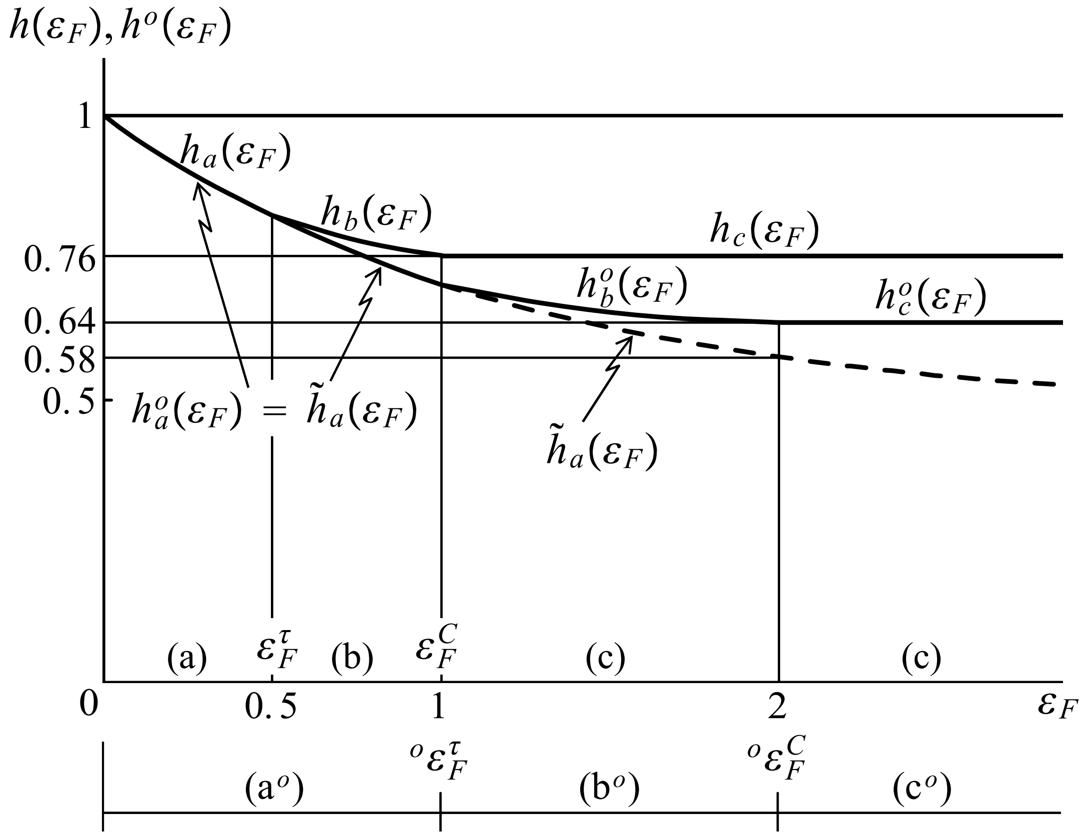

Figure 6. The impact of the advancement of net-technology on the relative productivity curve with the comparison of two cases where

Case 1: The relative productivity curve $h(\varepsilon_F)$ in the three ranges (a), (b) and (c), of ε_F when $\frac{\alpha_{CF}}{\alpha_{CN}} = 2$ and $\frac{\alpha_{\tau F}}{\alpha_{\tau N}} = 1.5$, $\rho_{CT} \cdot \rho_C = \rho_{\tau T} \cdot \rho_\tau = 0.25$.

Case 2: The relative productivity curve $h^o(\varepsilon_F)$ in the three ranges (a^o), (b^o) and (c^o), of ε_F when $\frac{\alpha_{CF}}{\alpha_{CN}^o} = 3$ and $\frac{\alpha_{\tau F}}{\alpha_{\tau N}^o} = 2$, $\rho_{CT} \cdot \rho_C = \rho_{\tau T} \cdot \rho_\tau = 0.25$.

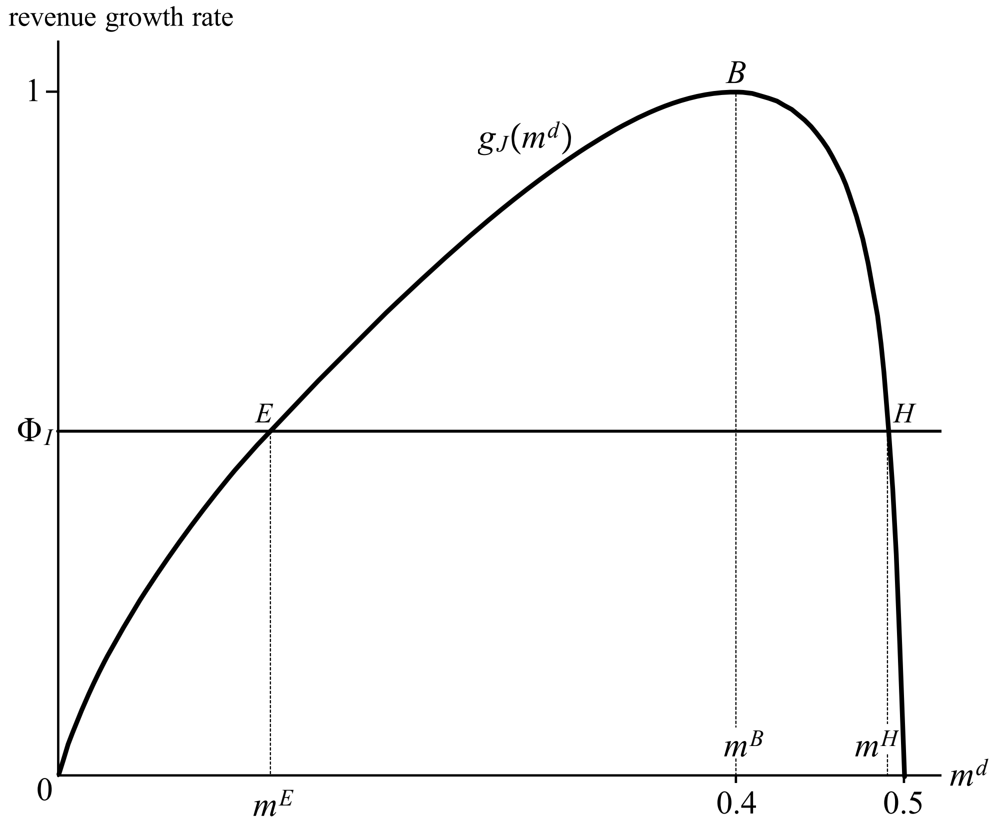

Figure 7. The knowledge growth rate curve $g_J(m^d)$ and the Bliss Point m^B.

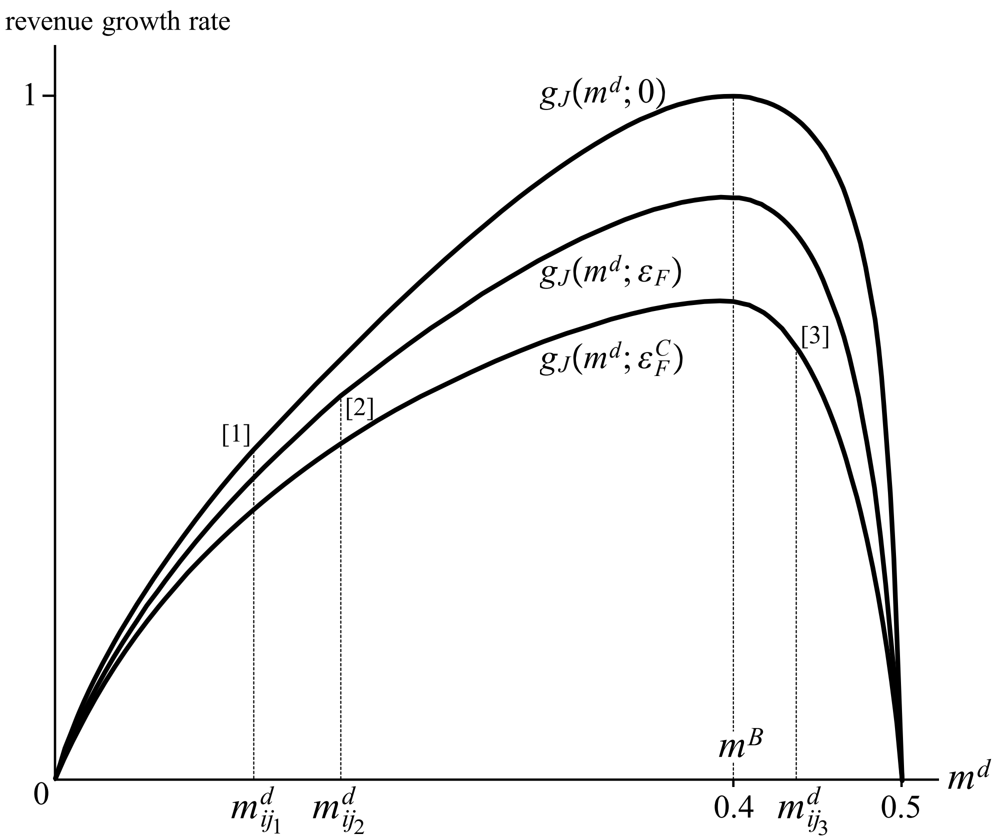

Figure 8. Knowledge growth rate curves $g_J(m^d;\varepsilon_F)$ for $\varepsilon_F = 0 < \varepsilon_F < \varepsilon_F^C$, sharing the same Bliss Point m^B, and the share of differential knowledge for each of three potential partners, (i,j_1), (i,j_2), (i,j_3).

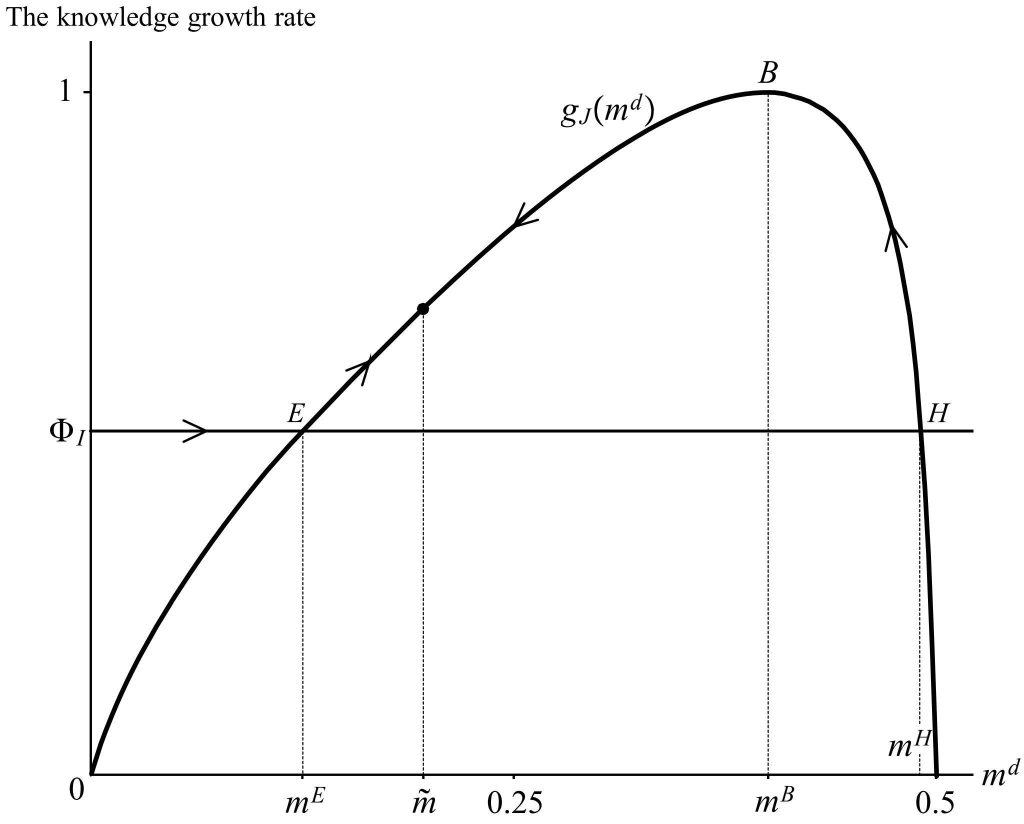

Figure 9. The dynamics of two-person system when $m^E < \tilde{m} < m^B$.

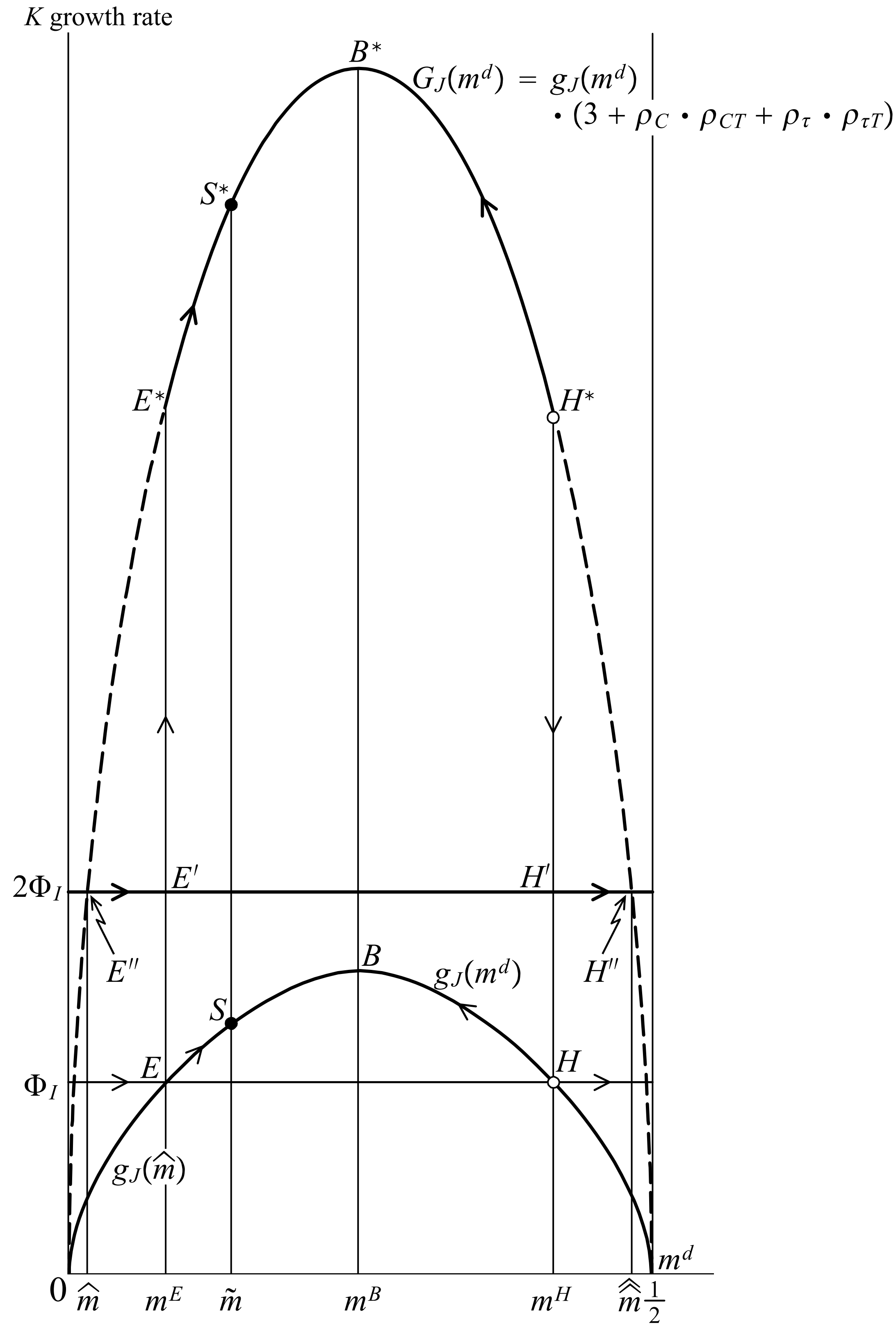

Figure 10. Dual dynamics of formal-*K* and total-*K* for the two-person system.

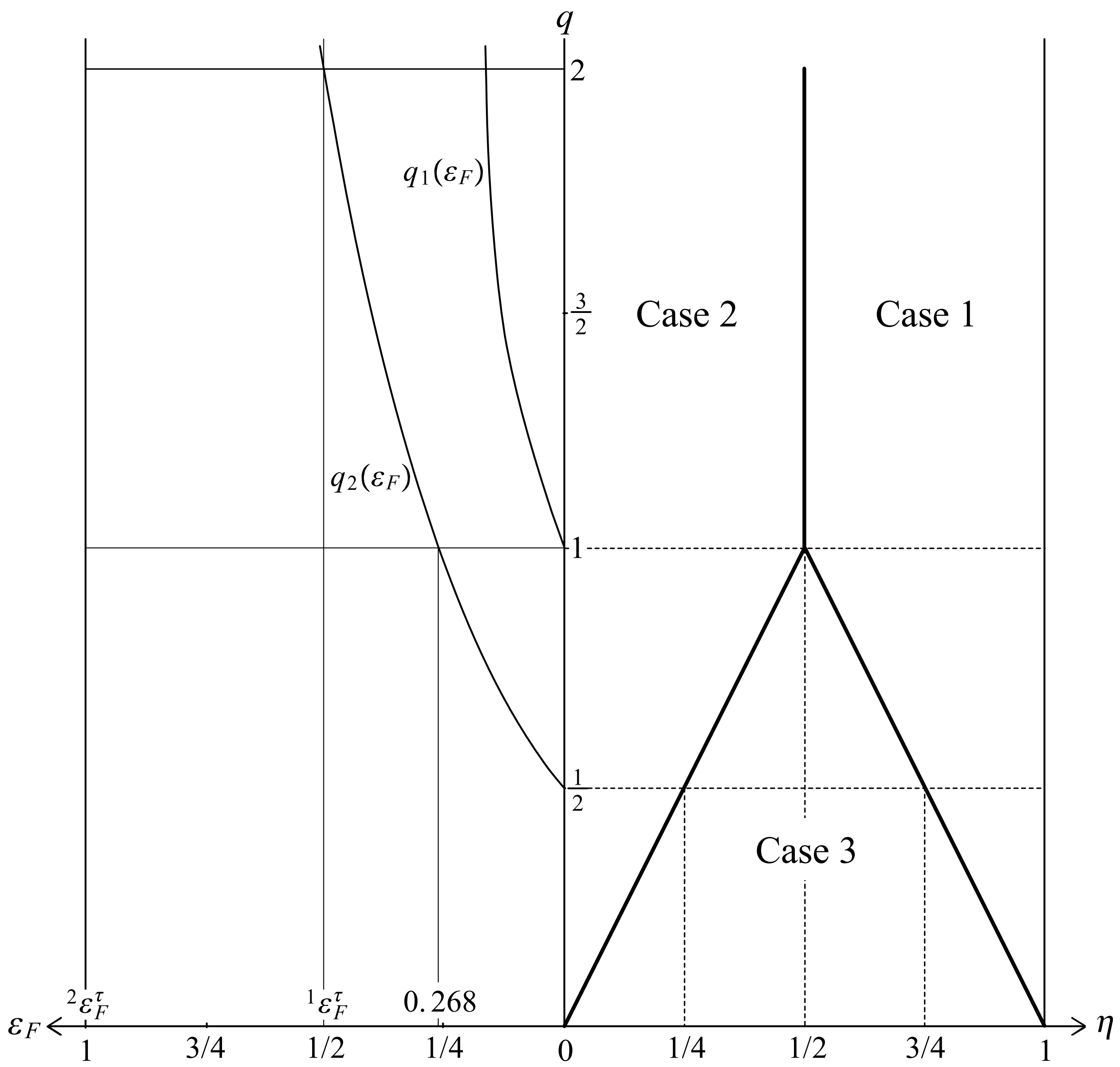

Figure A. Parameter ranges for the three cases of the impact of communication regulation with two examples of $q(\varepsilon_F)$ function:

Example 1: $q_1(\varepsilon_F)$ when $\frac{\alpha_{CF}}{\alpha_{CN}} = 2$ and $\frac{\alpha_{\tau F}}{\alpha_{\tau N}} = 1.5$

Example 2: $q_2(\varepsilon_F)$ when $\frac{\alpha_{CF}}{\alpha_{CN}} = 3$ and $\frac{\alpha_{\tau F}}{\alpha_{\tau N}} = 2.$

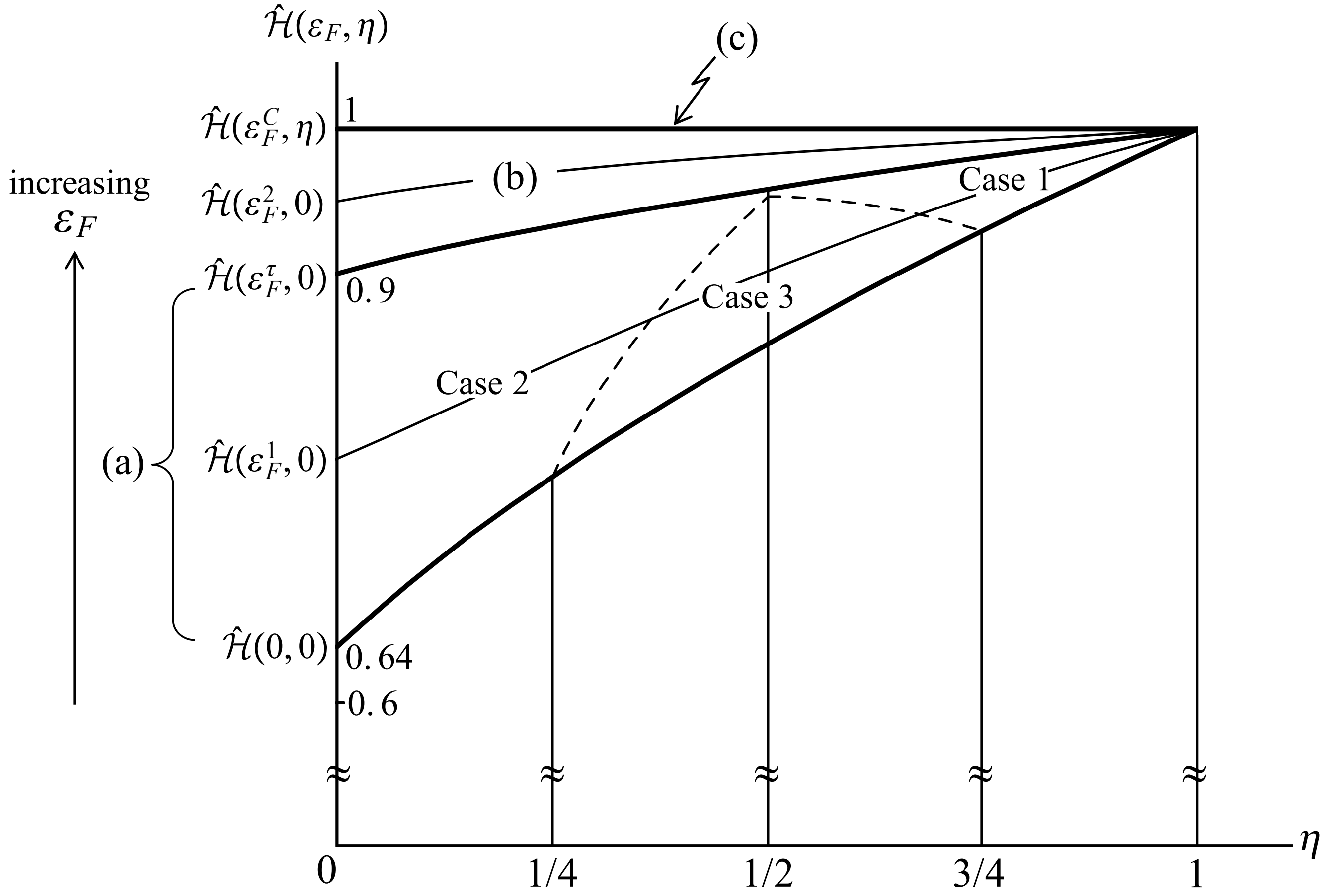

Figure B. Relative productivity curve $\widehat{\mathcal{H}}(\varepsilon_F, \eta)$ as a function of total F2F constraint η with parametric ε_F such that $0 < \varepsilon_F^1 < \varepsilon_F^\tau = 1 < \varepsilon_F^2 < \varepsilon_F^C = 2$; where $\frac{\alpha_{CF}}{\alpha_{CN}} = 3$, $\frac{\alpha_{\tau F}}{\alpha_{\tau N}} = 2$ and $\rho_{CT} \cdot \rho_C = \rho_{\tau T} \cdot \rho_\tau = 0.25$.

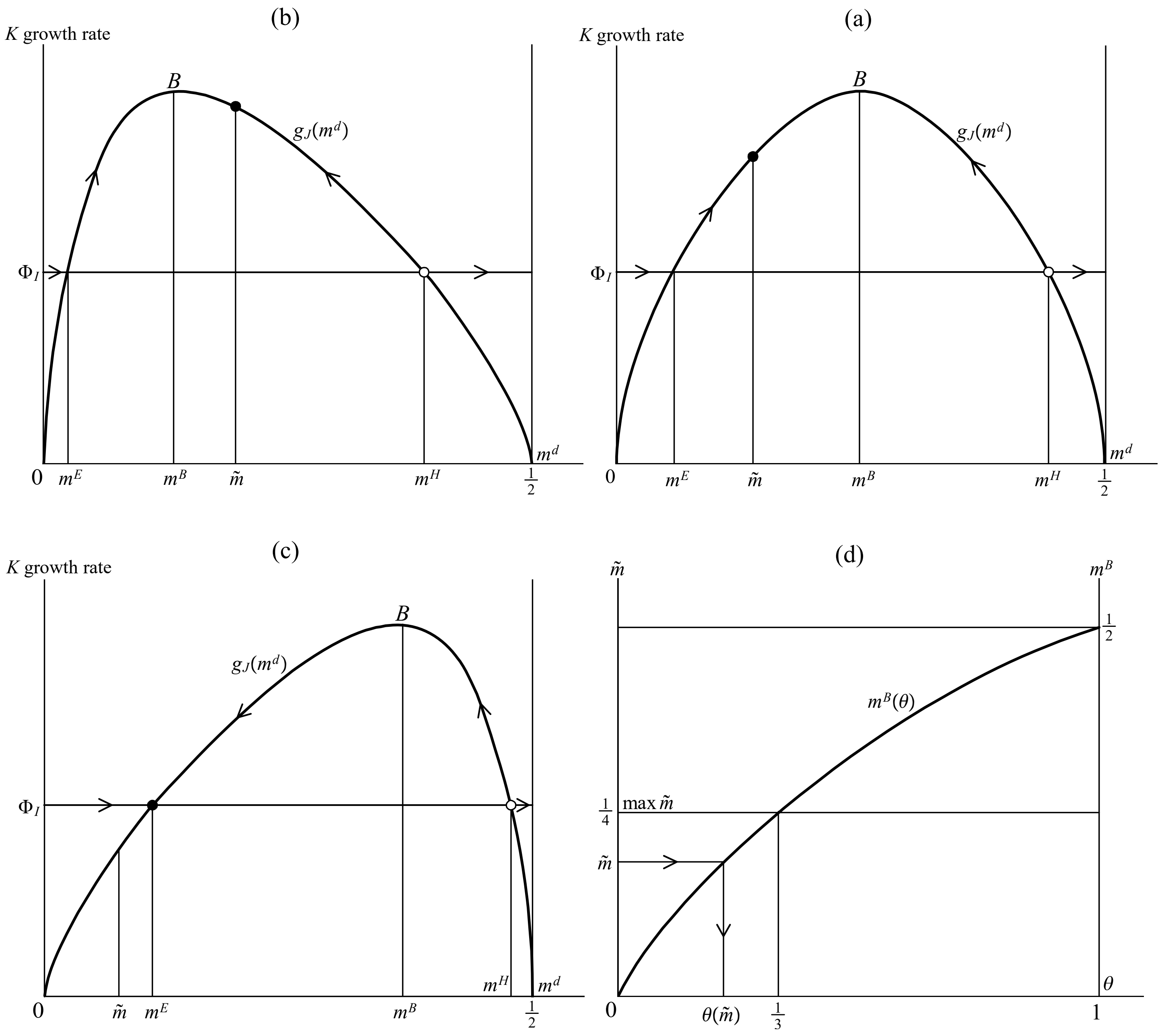

Figure C. The three possible cases, (a), (b) and (c), for the dynamics of two-person system; and diagram (d) for explaining the relationship between $\tilde{m}$ and m^B.

tier	activity output	equation of x	imputed value of x: $(\partial A_{ij}^{J*}/\partial x) \cdot x$
top	■ A_{ij}^{J*}	(43)	A_{ij}^{J*}
second tier	● a_{ij}^{C*}	(19)	$A_{ij}^{J*} \cdot \rho_C$
	○ a_{ii}^{I*}	(25)	$A_{ij}^{J*} \cdot \frac{\rho_I}{2}$
	○ a_{jj}^{I*}	(28)	$A_{ij}^{J*} \cdot \frac{\rho_I}{2}$
	● $a_{ij}^{\tau*}$	(39)	$A_{ij}^{J*} \cdot \rho_\tau$
third tier	● a_{ij}^{CTF}	(9)	$A_{ij}^{J*} \cdot \rho_C \cdot \rho_{CT} \cdot \frac{\lambda_{CF}^* \cdot a_{ij}^{CTF}}{\lambda_{CF}^* \cdot a_{ij}^{CTF} + \lambda_{CN}^* \cdot a_{ij}^{CTN}}$
	● a_{ij}^{CTN}	(10)	$A_{ij}^{J*} \cdot \rho_C \cdot \rho_{CT} \cdot \frac{\lambda_{CN}^* \cdot a_{ij}^{CTN}}{\lambda_{CF}^* \cdot a_{ij}^{CTF} + \lambda_{CN}^* \cdot a_{ij}^{CTN}}$
	○ a_i^{CSI}	(11)	$A_{ij}^{J*} \cdot \frac{\rho_C \cdot \rho_{CS}}{2}$
	○ a_j^{CSI}	(11)	$A_{ij}^{J*} \cdot \frac{\rho_C \cdot \rho_{CS}}{2}$
	○ a_i^{ITJ}	(21)	$A_{ij}^{J*} \cdot \frac{\rho_I \cdot \rho_{ITJ}}{2}$
	○ a_i^{ISJ}	(22)	$A_{ij}^{J*} \cdot \frac{\rho_I \cdot \rho_{ISJ}}{2}$
	○ a_j^{ITJ}	(21) with j	$A_{ij}^{J*} \cdot \frac{\rho_I \cdot \rho_{ITJ}}{2}$
	○ a_j^{ISJ}	(22) with j	$A_{ij}^{J*} \cdot \frac{\rho_I \cdot \rho_{ISJ}}{2}$
	● $a_{ij}^{\tau TF}$	(29)	$A_{ij}^{J*} \cdot \rho_\tau \cdot \rho_{\tau T} \cdot \frac{\lambda_{\tau F}^* \cdot a_{ij}^{\tau TF}}{\lambda_{\tau F}^* \cdot a_{ij}^{\tau TF} + \lambda_{\tau N}^* \cdot a_{ij}^{\tau TN}}$
	● $a_{ij}^{\tau TN}$	(30)	$A_{ij}^{J*} \cdot \rho_\tau \cdot \rho_{\tau T} \cdot \frac{\lambda_{\tau N}^* \cdot a_{ij}^{\tau TN}}{\lambda_{\tau F}^* \cdot a_{ij}^{\tau TF} + \lambda_{\tau N}^* \cdot a_{ij}^{\tau TN}}$
	○ $a_i^{\tau SI}$	(31)	$A_{ij}^{J*} \cdot \frac{\rho_\tau \cdot \rho_{\tau S}}{2}$
	○ $a_j^{\tau SI}$	(31)	$A_{ij}^{J*} \cdot \frac{\rho_\tau \cdot \rho_{\tau S}}{2}$

Table 1. The output x at each node of activity tree in Figure 6, and the imputed value of x.