Knowledge Creation as a Square Dance on the Hilbert Cube*

Marcus Berliant** and Masahisa Fujita±

March 12, 2007

Abstract

This paper presents a micro-model of knowledge creation through the interactions among a group of people. Our model incorporates two key aspects of the cooperative process of knowledge creation: (i) heterogeneity of people in their state of knowledge is essential for successful cooperation in the joint creation of new ideas, while (ii) the very process of cooperative knowledge creation affects the heterogeneity of people through the accumulation of knowledge in common. The model features myopic agents in a pure externality model of interaction. Surprisingly, in the general case for a large set of initial conditions we find that the equilibrium process of knowledge creation converges to the most productive state, where the population splits into smaller groups of optimal size; close interaction takes place within each group only. This optimal size is larger as the heterogeneity of knowledge is more important in the knowledge production process. Equilibrium paths are found analytically, and they are a discontinuous function of initial heterogeneity. JEL Classification Numbers: D83, O31 Keywords: knowledge creation, knowledge externalities, dynamic R and D, endogenous agent heterogeneity

---

*The first author is grateful for funding from the Kyoto Institute of Economic Research at Kyoto University, from Washington University in St. Louis, and from the American Philosophical Society. The second author is grateful for Grants Aid for Scientific Research Grant S 13851002 from the Japanese Ministry of Education and Science. We thank two anonymous referees, Alex Anas, Bob Hunt, Katsuhito Iwai, Felix Kubler, Fan-chin Kung, Charles Leung, David Levine, Koji Nishikimi, Joseph Perktold, Richard Rogerson, Tony Smith, and Takatoshi Tabuchi for helpful comments. Evidently, the authors alone are responsible for any remaining errors and for the views expressed herein.

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

±Institute of Economic Research, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-01 Japan Phone: (81-75) 753-7122, Fax: (81-75) 753-7198, e-mail: fujita@kier.kyoto-u.ac.jp

1 Introduction

How do knowledge creation and transfer perpetuate themselves? How do agents change during this process? What are the implications of the knowledge production process for R & D team size and characteristics?

There are a number of empirical regularities that we seek to address. Kaplinsky (1983) finds that in industries with rapid technical change, such as the computer software industry, firms start large but then generate spin-offs, so the average firm size decreases dramatically at a certain point in time. Why and when does this occur? Can we explain why the mean number of team members in the Broadway musical industry increased from 2 to 7 between 1880 and 1930, and has remained constant since then?[1] Why are there a large number of small firms in Higashi Osaka or in Ota ward in Tokyo, each specializing in different but related manufacturing services? Another example is the third Italy, where a large number of small firms produce a great variety of differentiated products. Yet another example is the restaurant industry in Berkeley, California.[2] In each case, the heterogeneity of workers and tacit knowledge accumulated within firms play a central role in the operation of the firms.

To address these empirical questions, we construct a dynamic model of group knowledge creation. As people create and transfer knowledge, they change. Thus, the history of meetings and their content is important. If people meet for a long time, then their base of knowledge in common increases, and their partnership eventually becomes less productive. Similarly, if two persons have very different knowledge bases, they have little common ground for communication, so their partnership will not be very productive.

For these reasons, we attempt to model endogenous agent heterogeneity, or horizontal agent differentiation, to look at the *permanent* effects of knowledge creation and growth.[3] In describing our model, the analogy between partner dancing and working jointly to create and exchange knowledge is useful, so we will use terms from these activities interchangeably. We assume that it is not possible for more than two persons to meet or dance at one time, though more than one couple can dance simultaneously. When agents meet, they create

[1] See Guimerà *et al* (2005) and Barabási (2005) for data and comment.

[2] In Berkeley, the parent restaurant is Chez Panisse.

[3] For simplicity, we employ a deterministic framework. It seems possible to add stochastic elements to the model, but at the cost of complexity. It should also be possible to apply the law of large numbers to a more basic stochastic framework to obtain equivalent results.

new, shared knowledge, thus building up knowledge in common. When agents are not meeting with each other, their knowledge bases grow more different. The fastest rate of knowledge creation occurs when common and differential knowledge are in balance. Knowledge creation and individual production all occur simultaneously at each point in time. The income of an agent at any given time is generated at a rate proportional to the agent's current stock of knowledge, as is new knowledge when an agent dances alone. Agents seek to maximize the current flow of income (the same as production) under certainty about everyone's state of knowledge, so a myopic core concept is used. The dancers can work alone or with a partner. The suitability of dance partners depends on the stock of knowledge they have in common and their respective stocks of exclusive knowledge.

For simplicity, we deal primarily with the case when the agents are symmetric. Our model is analytically tractable, so we do not have to resort to simulations; we find each equilibrium path explicitly. In this paper we consider only knowledge creation, not transfer. In Berliant and Fujita (2006), we work out the two person case with both knowledge creation and transfer, while allowing asymmetries. The results are similar, but the calculations are more complicated.

Our results are summarized as follows. There is a unique sink point that depends discontinuously on initial conditions. Only one of four specified sequences of dance patterns can occur along the equilibrium path. When the initial state features relative homogeneity of knowledge between agents, the sink will be the most productive state, where the population splits into smaller groups of optimal size; close interaction takes place within each group only. This optimal size is larger as the heterogeneity of knowledge is more important in the knowledge production process. The result demonstrating that the sink point is the most productive state is most surprising to us, as we posit a model with myopic agents and no markets, but rather with only externalities in interactions between agents, so one would not expect efficient outcomes. It is also surprising to note that from an initially symmetric situation, the model generates asymmetries in the following sense. Both the size and characteristics of research teams are endogenous; workers in the same group continue to work together, but their knowledge profile drifts away from the profile of workers who are not part of the same group. This creates the boundaries of research teams endogenously.

The model is also at an intermediate level of aggregation. That is, although it is at a more micro level than large aggregate models, we do not work out completely its microfoundations. That is left to future research.

Future applications and extensions of the model are numerous.[4] For instance, to address questions related to how knowledge diversity or patent policy affects long run economic growth, it would be possible to employ our model as the R & D sector of a growth model. To address questions related to R & D firm agglomeration, a spatial dimension could be introduced explicitly, where agents can interact only with others in their region, but migration can lead to recombinations of agents from different regions.

We believe that our model can be tested further by examining the dynamic pattern of teams. Do R & D teams experience employment turnover early in an industry and then settle down? Does firm size in an industry begin large and then suddenly become small? Is team size related to the mutual knowledge required for an industry's R & D process? Do coauthorships in economics or other fields follow the interaction paths predicted by our model?

Differentiation of agents in terms of *quality* (or vertical characteristics) of knowledge is studied in Jovanovic and Rob (1989) in the context of a search model. In contrast, our model examines (endogenous) horizontal heterogeneity of agents and its effect on knowledge creation and consumption.

Our work is related to the literature on teams; see for example Holmstrom (1982) or Aoki (1994). In general, this literature explores how the moral hazard or free rider problem is solved in group production. Smaller or more homogeneous groups will reduce this problem. Our framework abstracts from information asymmetry, instead focusing on the dynamics of the endogenous composition and size of teams due to knowledge heterogeneity. Empirical work must account for both aspects.

Section 2 gives the model and notation, Section 3 analyzes equilibrium in the case of two participants or dancers for expositional purposes, Section 4 extends the model to N persons and analyzes equilibrium, whereas Section 5 explores the efficiency properties of equilibrium. In order to investigate the nature of optimal group size, Section 6 extends the basic model to allow the importance of heterogeneity in knowledge creation to vary exogenously. Section 7 gives our conclusions and suggestions for future dancing. Two appendices provide the proofs of key results.

[4] An important question for future research is whether the surprising result that the sink point is the most productive state is robust in a more general context, for example with knowledge transfer. We conjecture that it is, because when agents determine their group size, they account for knowledge transfer from other groups.

2 The Model - Ideas and Knowledge

In this section, we introduce the basic concepts of our model of ideas and knowledge.

An *idea* is represented by a box. It has a label on it that everyone can read (the label is common knowledge in the game we shall describe). This label describes the contents. Each box contains an idea that is described by its label. Learning the actual contents of the box, as opposed to its label, takes time, so although anyone can read the label on the box, they cannot understand its contents without investing time. This time is used to open the box and to understand fully its contents. An example is a recipe for making "udon noodles as in Takamatsu." It is labelled as such, but would take time to learn. Another example is reading a paper in a journal. Its label or title can be understood quickly, but learning the contents of the paper requires an investment of time. Production of a new paper, which is like opening a new box, either jointly or individually, also takes time.

Suppose we have an infinite number of boxes, each containing a different piece of knowledge, which is what we call an idea. We put them in a row in an arbitrary order.

There are N persons in the economy, where N is a finite integer. People are indexed by i and j. At this point, we assume that there are only two people; general indexing is used so that we can add more people to the model later. We assume that each person has a replica of the infinite row of boxes introduced above, and that each copy of the row has the same order. Our model features continuous time. Fix time $t \in \mathbb{R}_+$ and consider any person i. A box is indexed by $k = 1, 2, ...$ Take any box k. If person i knows the idea inside that box, we put a sticker on it that says 1; otherwise, we put a sticker on it that says 0. That is, let $x_i^k(t) \in \{0, 1\}$ be the sticker on box k for person i at time t. The *state of knowledge*, or just *knowledge*, of person i at time t is thus defined to be $K_i(t) = (x_i^1(t), x_i^2(t), ...) \in \{0, 1\}^\infty$. The reason we use an infinite vector of possible ideas is that we are using an infinite time horizon, and there are always new ideas that might be discovered, even in the preparation of udon noodles. More formally, let $\mathcal{H}$ be the *Hilbert cube*; it consists of all real sequences with values in $[0, 1]$. That is, if $\mathbb{N}$ is the set of natural numbers, then $\mathcal{H} = [0, 1]^{\mathbb{N}}$. So the knowledge of person i at time t, $K_i(t)$, is a vertex of the Hilbert cube $\mathcal{H}$. Notice that given any vertex of $\mathcal{H}$, there exists an infinite number of adjacent vertices. That is, given $K_i(t)$ with only finitely many non-zero components, there is an infinite number of ideas that could be created in the next step.

In this paper, we will treat ideas symmetrically. Extensions to idea hierarchies and knowledge structures will be discussed in the conclusions.

Given $K_i(t) = (x_i^1(t), x_i^2(t), ...)$,

$$n_i(t) = \sum_{k=1}^{\infty} x_i^k(t) \tag{1}$$

represents the number of ideas known by person i at time t. Next, we will define the number of ideas that two persons, i and j, both know. Assume that $j \neq i$. Define $K_j(t) = (x_j^1(t), x_j^2(t), ...)$ and

$$n_{ij}^c(t) = \sum_{k=1}^{\infty} x_i^k(t) \cdot x_j^k(t) \tag{2}$$

So $n_{ij}^c(t)$ represents the number of ideas known by both persons i and j at time t. Notice that i and j are symmetric in this definition, so $n_{ij}^c(t) = n_{ji}^c(t)$. Define

$$n_{ij}^d(t) = n_i(t) - n_{ij}^c(t) \tag{3}$$

to be the number of ideas known by person i but not known by person j at time t. Then, it holds by definition that

$$n_i(t) = n_{ij}^c(t) + n_{ij}^d(t) \tag{4}$$

Define $n^{ij}(t)$ be the total number of ideas possessed by persons i and j together at time t. Then, tautologically

$$n^{ij}(t) = n_{ij}^c(t) + n_{ij}^d(t) + n_{ji}^d(t) \tag{5}$$

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

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

So consider a given time t. In order to explain how knowledge creation and commodity production work, it is useful for intuition (but not technically necessary) to view this time period of fixed length as consisting of subperiods of fixed length. Each individual is endowed with a fixed amount of labor that is supplied inelastically during the period. In the first subperiod, individual production takes place. We shall assume constant returns to scale in physical production, so it is not beneficial for individuals to collaborate in production. Each individual uses their labor during the first subperiod to produce consumption good on their own, whether or not they are meeting. We shall assume below that although there are no increasing returns to scale in production, the productivity of a person's labor depends on their stock of knowledge. Activity in the second subperiod depends on whether or not there is a meeting. If there is no meeting, then each person spends the second subperiod creating new knowledge on their own. Evidently, the new knowledge created during this subperiod differs between the two persons, because they are not communicating. They open different boxes. *Since there is an infinity of different boxes, the probability that the two agents will open the same box (even at different points in time), either working by themselves or in distinct meetings, is assumed to be zero.* If there is a meeting, then they create new knowledge *together*, so they open boxes together.[5] We wish to emphasize that the division of a time period into subperiods is purely an expositional device. Rigorously, whether or not a meeting occurs determines how much attention is devoted to the various activities at a given time.

What do the agents know when they face the decision about whether or not to meet a potential partner j at time t? Each person knows both $K_i(t)$ and $K_j(t)$. In other words, each person is aware of their own knowledge and is also aware of all others' knowledge. Thus, they also know $n_i(t)$, $n_j(t)$, $n^c_{ij}(t) = n^c_{ji}(t)$, $n^d_{ij}(t)$, and $n^d_{ji}(t)$ (for all $j \neq i$) when they decide whether or not to meet at time t. The notation for whether or not a meeting of persons i and j actually occurs at time t is: $\delta_{ij}(t) = \delta_{ji}(t) = 1$ if a meeting occurs and $\delta_{ij}(t) = \delta_{ji}(t) = 0$ if no meeting occurs at time t. For convenience, we define $\delta_{ii}(t) = 1$ when person i works in isolation at time t, and $\delta_{ii}(t) = 0$ when person i meets with another person at time t.

Next, we must specify the dynamics of the knowledge system and the objectives of the people in the model in order to determine whether or not two persons decide to meet at a particular time. In order to accomplish this, it is

[5] Clearly, the creation of this paper is an example of the process described.

easiest to abstract away from the notation for specific boxes, $K_i(t)$, and to focus on the dynamics of the quantity statistics related to knowledge, $n_i(t)$, $n_j(t)$, $n^c_{ij}(t) = n^c_{ji}(t)$, $n^d_{ij}(t)$, and $n^d_{ji}(t)$. Since we are treating ideas symmetrically, in a sense these quantities are sufficient statistics for our analysis.[6]

The simplest piece of the model to specify is what happens if there is no meeting between person i and anyone else, so i works in isolation. Let $a_{ii}(t)$ be the rate of creation of new ideas created by person i in isolation at time t (this means that i meets with itself). Then we assume that the creation of new knowledge during isolation is governed by the following equation:

$$a_{ii}(t) = \alpha \cdot n_i(t) \text{ when } \delta_{ii}(t) = 1. \tag{6}$$

So we assume that if there is no meeting at time t, individual knowledge grows at a rate proportional to the knowledge already acquired by an individual.

If a meeting occurs between i and j at time t ($\delta_{ij}(t) = 1$), then joint knowledge creation occurs, and it is governed by the following dynamics:[7]

$$a_{ij}(t) = \beta \cdot [n^c_{ij}(t) \cdot n^d_{ij}(t) \cdot n^d_{ji}(t)]^{\frac{1}{3}} \text{ when } \delta_{ij}(t) = 1 \text{ for } j \neq i \tag{7}$$

So when two people meet, joint knowledge creation occurs at a rate proportional to the normalized product of their knowledge in common, the differential knowledge of i from j, and the differential knowledge of j from i. The rate of creation of new knowledge is highest when the proportions of ideas in common, ideas exclusive to person i, and ideas exclusive to person j are split evenly. Ideas in common are necessary for communication, whereas ideas exclusive to one person or the other imply more heterogeneity or originality in the collaboration. If one person in the collaboration does not have exclusive ideas, there is no reason for the other person to meet and collaborate. The multiplicative nature of the function in equation (7) drives the relationship between knowledge creation and the relative proportions of ideas in common

[6] In principle, all of these time-dependent quantities are positive integers. However, for simplicity we take them to be continuous (in $\mathbb{R}_+$) throughout the paper. One interpretation is that the creation of an idea occurs at a stochastic time, and the real numbers are taken to be the expected number of jumps (ideas learned) in a Poisson process. The use of an integer instead of a real number seems to add little but complication to the analysis.

[7] We may generalize equation (7) as follows:

$$a_{ij}(t) = \max\left\{(\alpha - \varepsilon)\, n_i(t), (\alpha - \varepsilon)\, n_j(t), \beta \left[n^c_{ij}(t) \cdot n^d_{ij}(t) \cdot n^d_{ji}(t)\right]^{\frac{1}{3}}\right\}$$

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

and ideas exclusive to one or the other agent. Under these circumstances, no knowledge creation in isolation occurs.

Whether a meeting occurs or not, there is production in each period for both persons. *Felicity (or instantaneous utility) in that time period is defined to be the quantity of output.*[8] Define $y_i(t)$ to be production output (or felicity) for person i at time t, that is consumed by person i. The output is taken to be numéraire. Normalizing the coefficient of production to be 1, we take

$$y_i(t) = n_i(t) \tag{8}$$

so output of private good for person i at time t is a function of person i's human capital; in turn, this is assumed to be person i's stock of knowledge. Person i's lifetime utility is given by

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

where the constant γ is the discount rate common to all agents. In the present context, since $y_i(t) = n_i(t)$ is the stock variable that is actually fixed at time t for each consumer i, what the consumer can choose at time t is the rate of increase of income or knowledge, the flow variable:

$$\dot{y}_i(t) = \dot{n}_i(t) \tag{9}$$

At this point, we introduce the assumption that the agents are myopic in choosing their partners (or working in isolation). In particular, they do not foresee the consequences of their choice of action on the future path of consumption, but rather only see the immediate consequences, namely agent i's objective at time t is to maximize $\dot{y}_i(t)$.

By definition,

$$\frac{\dot{y}_i(t)}{y_i(t)} = \frac{\dot{n}_i(t)}{n_i(t)} \tag{10}$$

which represents the rate of growth of income. Since $y_i(t)$ is the stock variable, choosing partners to maximize $\dot{y}_i(t)$ is the same as choosing partners to maximize $\frac{\dot{y}_i(t)}{y_i(t)}$.

We now describe the dynamics of the system, dropping the time argument.

[8] Given that the focus of this paper is on *knowledge creation* rather than production, we use the simplest possible form for the production function.

Let us focus on agent i, as the expressions for the other agents are analogous.

$$\begin{aligned} \dot{y}_i &= \dot{n}_i = \sum_{j=1}^{N} \delta_{ij} \cdot a_{ij} && (11) \\ \dot{n}_{ij}^c &= \delta_{ij} \cdot a_{ij} \text{ for all } j \neq i && (12) \\ \dot{n}_{ij}^d &= \sum_{k \neq j} \delta_{ik} \cdot a_{ik} \text{ for all } j \neq i && (13) \end{aligned}$$

Equation (11) is based on the assumption that once learned, ideas are not forgotten. Thus, the increase in the knowledge of person i is the sum of the knowledge created in isolation and the knowledge created jointly with someone else. Equation (12) means that the increase in the knowledge in common for persons i and j equals the new knowledge created jointly by them. Finally, equation (13) means that all the knowledge created by person i either in isolation or joint with persons other than person j becomes a part of the differential knowledge of person i from person j.

By definition, it is also the case that

$$\sum_{j=1}^{N} \delta_{ij} = 1$$

Furthermore, on the equilibrium path it is necessary that

$$\delta_{ij} = \delta_{ji} \text{ for all } i \text{ and } j$$

Concerning the rule used by an agent to choose their best partner, to keep the model tractable in this first analysis, we assume a myopic rule. At each moment of time t, person i would like a meeting with person j when the rate of growth of income while meeting with j is highest among all potential partners, including himself.[9] As we are attempting to model close interactions within groups, we assume that at each time, the myopic persons interacting choose a core configuration. That is, we restrict attention to configurations such that at any point in time, no coalition of persons can get together and make themselves better off *in that time period.* In essence, our solution concept at a point in time is the myopic core.

In order to analyze our dynamic system, we first divide all of our equations by the total number of ideas possessed by i and j:

$$n^{ij} = n_{ij}^d + n_{ji}^d + n_{ij}^c \quad (14)$$

[9] We introduce later the rule used in the case of ties.

and define new variables

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

By definition, m_{ij}^{d} represents the percentage of ideas exclusive to person i among all the ideas known by person i or person j. Similarly, m_{ij}^{c} represents the ideas known in common by persons i and j among all the ideas known by the pair. From (14), we obtain

$$1 = m_{ij}^{d} + m_{ji}^{d} + m_{ij}^{c} \tag{15}$$

Using these new variables, for each pair of dancers i and j ($i \neq j$), we obtain (see Theorem A1 in Technical Appendix a):

$$\frac{a_{ij}}{n_i} = G(m_{ij}^{d}, m_{ji}^{d}) \quad \text{for } i \neq j \tag{16}$$

where

$$G(m_{ij}^{d}, m_{ji}^{d}) \equiv \frac{\beta\left[\left(1 - m_{ij}^{d} - m_{ji}^{d}\right) \cdot m_{ij}^{d} \cdot m_{ji}^{d}\right]^{\frac{1}{3}}}{1 - m_{ji}^{d}} \tag{17}$$

which represents the growth rate of their knowledge when two persons i and j meet. Then, using (6) and (11), we can rewrite the income growth rate, equation (10), as follows:

$$\frac{\dot{y}_i}{y_i} = \frac{\dot{n}_i}{n_i} = \delta_{ii} \cdot \alpha + \sum_{j \neq i} \delta_{ij} \cdot G(m_{ij}^{d}, m_{ji}^{d}) \tag{18}$$

Furthermore, using (6), (7) and (16), we have (see Theorem A2 in Technical Appendix a):

$$\begin{aligned} \dot{m}_{ij}^{d} &= \alpha \cdot \left(1 - m_{ij}^{d}\right) \cdot \left[\delta_{ii} \cdot \left(1 - m_{ji}^{d}\right) - \delta_{jj} \cdot m_{ij}^{d}\right] - \delta_{ij} \cdot m_{ij}^{d} \cdot \left(1 - m_{ji}^{d}\right) \cdot G(m_{ij}^{d}, m_{ji}^{d}) \\ &\quad + \left(1 - m_{ij}^{d}\right) \cdot \left(1 - m_{ji}^{d}\right) \cdot \sum_{k \neq i,j} \delta_{ik} \cdot G(m_{ik}^{d}, m_{ki}^{d}) \\ &\quad - \left(1 - m_{ij}^{d}\right) \cdot m_{ij}^{d} \cdot \sum_{k \neq i,j} \delta_{jk} \cdot G(m_{jk}^{d}, m_{kj}^{d}) \end{aligned} \tag{19}$$

for $i, j = 1, 2, \cdots, N$. Thus, the dynamics of the system are described in terms of m_{ij}^{d} ($i, j = 1, 2, \cdots, N$) only. Before analyzing the general model with any population, to provide intuition we first examine the two person case. This system, with analogous equations for agent j, represents a partner dance on the vertices of the Hilbert cube.

3 The Two Person Model

3.1 Equilibrium Dynamics

Consider $N = 2$ and we call the two agents i and j. Applying (18) to the present context and setting $\delta_{ii} = 1 - \delta_{ij}$ and $\delta_{jj} = 1 - \delta_{ji}$ yields

$$\begin{aligned} \frac{\dot{y}_i}{y_i} &= \frac{\dot{n}_i}{n_i} = (1 - \delta_{ij}) \cdot \alpha + \delta_{ij} \cdot G(m_{ij}^d, m_{ji}^d) \\ \frac{\dot{y}_j}{y_j} &= \frac{\dot{n}_j}{n_j} = (1 - \delta_{ji}) \cdot \alpha + \delta_{ji} \cdot G(m_{ji}^d, m_{ij}^d) \end{aligned} \tag{20}$$

Likewise, by omitting the last two lines in equation (19) and setting $\delta_{ii} = 1 - \delta_{ij}$ and $\delta_{jj} = 1 - \delta_{ji} = 1 - \delta_{ij}$ (since $\delta_{ij} = \delta_{ji}$ in equilibrium), we have

$$\begin{aligned} \dot{m}_{ij}^d &= (1 - \delta_{ij}) \cdot \alpha \cdot \left(1 - m_{ij}^d\right) \cdot \left(1 - m_{ij}^d - m_{ji}^d\right) \\ &\quad - \delta_{ij} \cdot m_{ij}^d \cdot \left(1 - m_{ji}^d\right) \cdot G(m_{ij}^d, m_{ji}^d) \end{aligned} \tag{21}$$

$$\begin{aligned} \dot{m}_{ji}^d &= (1 - \delta_{ij}) \cdot \alpha \cdot \left(1 - m_{ji}^d\right) \cdot \left(1 - m_{ij}^d - m_{ji}^d\right) \\ &\quad - \delta_{ji} \cdot m_{ji}^d \cdot \left(1 - m_{ij}^d\right) G(m_{ji}^d, m_{ij}^d) \end{aligned}$$

The general two person system, allowing asymmetric situations, is studied in detail in Berliant and Fujita (2006). To provide intuition, here we focus on the special case where the initial state is symmetric, namely $m_{ij}^d(0) = m_{ji}^d(0) = m(0)$. It should be clear from these equations that once the general system attains a symmetric state, say at time 0, then in equilibrium the state remains symmetric forever.[10] Along any symmetric equilibrium path,

$$\begin{aligned} m_{ij}^d &= m_{ji}^d = m \\ m^c &= 1 - m \end{aligned}$$

Hence the state of the system is completely specified by the scalar m, representing the percentage of the total number of ideas exclusive to each person.

To study this system in greater detail, we must study whether each person does better creating new ideas in isolation or together. Setting

$$y_i = y_j = y$$

[10] Berliant and Fujita (2006) show that there is a large set of initial conditions from which the equilibrium process reaches a symmetric state in finite time.

we use (20) to obtain

$$\frac{\dot{y}(t)}{y(t)} = [1-\delta_{ij}(t)]\cdot\alpha + \delta_{ij}(t)\cdot G(m(t),m(t))$$

To simplify notation, we define the growth rate when the two persons meet, $\delta_{ji}=\delta_{ij}=1$, as

$$g(m) \equiv G(m,m) = \beta\cdot\frac{[(1-2m)\cdot m^2]^{\frac{1}{3}}}{1-m} \tag{22}$$

Thus

$$\frac{\dot{y}(t)}{y(t)} = [1-\delta_{ij}]\cdot\alpha + \delta_{ij}\cdot g(m) \tag{23}$$

In order to maximize the income growth rate given by (23), both agents want to meet (i.e., $\delta_{ij}=\delta_{ji}=1$) when

$$g(m) > \alpha$$

Thus, the meeting between agents i and j actually occurs when this inequality holds.

Figure 1 illustrates the graph of the function $g(m)$ as a bold line for $\beta=1$.

FIGURE 1 GOES HERE

Differentiating $g(m)$ yields

$$g'(m) = \frac{\beta}{3}\left[\frac{(1-2m)\cdot m^2}{(1-m)^3}\right]^{-\frac{2}{3}}\cdot\frac{m\cdot(2-5m)}{(1-m)^4} \tag{24}$$

implying that

$$g'(m)\begin{matrix}>\\=\\<\end{matrix}0 \quad\text{as}\quad m\begin{matrix}<\\=\\>\end{matrix}\frac{2}{5} \quad\text{for } m\in(0,\frac{1}{2}) \tag{25}$$

Thus, $g(m)$ is strictly quasi-concave on $[0,\frac{1}{2}]$, achieving its maximal value at $m^B=\frac{2}{5}$; we call the latter the "Bliss Point." Presuming that $g(m^B)>\alpha$, it is the point where the rate of increase in income is maximized for each person. Define the set of states where meetings occur to be

$$M = \{m\in[0,\frac{1}{2}] \mid g(m)>\alpha\}$$

Since g is strictly quasi-concave, M is convex. Let m^J be the greatest lower bound of M and let m^I be the least upper bound of M. Hence $M=(m^J,m^I)$; see Figure 1. Whenever $M\neq\emptyset$, $m^B\in M$, so $m^J<2/5$ (as long as $g(m^B)>\alpha$).

In fact we can describe the properties of the set M in general. As α increases, the productivity of creating ideas alone increases, so people are less likely to want to meet to create new ideas, implying that M shrinks as α increases. If α is a little more than $\beta \cdot (4/9)^{\frac{1}{3}}$, M disappears.

Next we discuss the dynamics of the system, assuming that the equilibrium condition $\delta_{ij} = \delta_{ji} = \delta$ always holds. Consider first the case where there is no meeting, so $\delta = 0$ is fixed exogenously. Then from equations (21), the dynamics are given by the following equation:

$$\dot{m} = \alpha \cdot (1 - m)(1 - 2m)$$

If there is no meeting ($\delta = 0$), then $\dot{m}$ is non-negative, and positive on $(0, 1/2)$. So if there is no meeting, the vector field points to the right, and the system tends to $m = 1/2$.

With a meeting, $\delta = 1$. Then (21) implies:

$$\dot{m} = -m \cdot \beta \cdot [(1 - 2m) \cdot m^2]^{\frac{1}{3}} \tag{26}$$

This expression is negative on $(0, 1/2)$ and the vector field points to the left. The sink is at 0, so the system eventually moves there under the assumption of a meeting.

Next, we combine the case where there is no meeting ($\delta_{ij} = 0$) with the case where there is a meeting ($\delta_{ij} = 1$), and let the agents choose whether or not to meet. The model follows the dynamics for meetings ($\delta_{ij} = 1$) on M and the dynamics for no meetings ($\delta_{ij} = 0$) on the complement of M.

The state $m = 1/2$ is a stable point of the system; the myopic return to no meeting dominates the return to meeting, since the two persons have little in common. This stable point, however, is not very interesting.

We have not completely specified the dynamics. This is especially important on the boundary of M (namely m^I and m^J), where people are indifferent between meeting and not meeting. We take an arbitrarily small unit of time, Δt, and assume that if both people become indifferent between meeting and not meeting, but the two persons are currently meeting, then the meeting must continue for at least Δt units of time. Similarly, if the two persons are not meeting when both people become indifferent between meeting and not meeting, then they cannot meet for at least Δt units of time. So if people become indifferent between meeting or not meeting at time t, the function $\delta_{ij}(t)$ cannot change its value until time $t + \Delta t$. Finally, when at least one person initially happens to be on the boundary of M (that is, at least one person is indifferent between meeting and not meeting), then they cannot meet for at least Δt units of time. Under this set of rules, we can be more specific about the dynamic process near the boundary of M.

In terms of dynamics, if the system does not evolve toward the uninteresting stable point where there are no meetings (and the two people have nothing in common), eventually the system reaches the point m^J. It is the remaining stable point of our model. Small movements around m^J will continue due to our assumption about the dynamics at the boundary of M, namely that meetings or isolation are sticky. As $\Delta t \to 0$, the process converges to the point m^J. The point m^J features symmetry between the two agents with a large degree of homogeneity relative to the remainder of the points in M and the other points in $[0, 1/2]$ generally.

So given various initial compositions of knowledge $m(0)$, where will the system end up? If the initial composition of knowledge is such that the couple has little in common, namely $m(0) \geq m^I$, the sink will be $m = 1/2$. If the initial composition of knowledge is such that the couple has more in common, namely $m(0) < m^I$, then the sink point will be $m = m^J$.

The point $m^J < 2/5$ exists and is unique as long as $M \neq \emptyset$.

Without loss of generality, we can allow δ_{ij} to take values in $[0, 1]$ rather than in $\{0, 1\}$. The interpretation of a fractional δ_{ij} is that at each instant of time, a person divides their time between a meeting δ_{ij} proportion of that instant and isolation $(1-\delta_{ij})$ proportion of that instant.[11] The purpose of this generalization is to capture the limit of a sequence of dance patterns where two partners are alternating between working together and then in isolation for shorter and shorter periods. All of our results concerning the model when δ_{ij} is restricted to $\{0, 1\}$ carry over to the case where $\delta_{ij} \in [0, 1]$. The reason is that except on the boundary of M, persons strictly prefer $\delta_{ij} \in \{0, 1\}$ to fractional values of δ_{ij}, as each person's objective function is linear in δ_{ij}. On the boundary of M, our rule concerning dynamics prevents δ_{ij} from taking on fractional values, as it must retain its value from the previous iteration of the process for at least time $\Delta t > 0$. So if the process pierces the boundary from inside M, it must retain $\delta_{ij} = 1$ for an additional time of at least Δt. If it pierces the boundary from outside M, it must retain $\delta_{ij} = 0$ for an additional time of at least Δt. As $\Delta t \to 0$, the process converges to the point m^J in finite time.

[11] An alternative interpretation is that at each instant of time, they devote their attention to working together for δ_{ij} proportion of that instant and to working in isolation for $(1-\delta_{ij})$ proportion of that instant.

It may seem trivial to allow fractional δ_{ij} when discussing equilibrium behavior with two people, but allowing fractional δ_{ij} is crucial to Section 4, where we consider the general case.

3.2 Efficiency

To construct an analog of Pareto efficiency for this model, we use a social planner who can choose whether or not people should meet in each time period. As noted above, we shall allow the social planner to choose values of δ_{ij} in $[0,1]$, so that persons can be required to meet for a percentage of the total time in a period, and not meet for the remainder of the period. The feasibility condition $\delta_{ij} = \delta_{ji}$ is imposed for all paths considered. To avoid dependence of our notion of efficiency on a discount rate, we employ the following alternative concepts. The first is stronger than the second. A *path of* δ_{ij} is a measurable function of time (on $[0,\infty)$) taking values in $[0,1]$. For each path of δ_{ij}, there corresponds a unique time path of m_{ij}^d determined by equation (21), respecting the initial condition, and thus a unique time path of income $y_i(t;\delta_{ij})$. We say that a path δ'_{ij} *(strictly) dominates* a path δ_{ij} if

$$y_i(t;\delta'_{ij}) \geq y_i(t;\delta_{ij}) \text{ and } y_j(t;\delta'_{ij}) \geq y_j(t;\delta_{ij}) \text{ for all } t \geq 0$$

with strict inequality for at least one person over a positive interval of time. As this concept is quite strong, and thus difficult to use as an efficiency criterion, it will sometimes be necessary to employ a weaker concept, which we discuss next. We say that a path δ_{ij} *is overtaken* by a path δ'_{ij} if there exists a t' such that

$$y_i(t;\delta'_{ij}) \geq y_i(t;\delta_{ij}) \text{ and } y_j(t;\delta'_{ij}) \geq y_j(t;\delta_{ij}) \text{ for all } t > t'$$

with strict inequality for at least one person over a positive interval of time.

We defer discussion of additional efficiency criteria using the intertemporal utilitarian welfare function to Appendix 3.

Once again, for efficiency analysis, we consider only symmetric equilibrium paths (namely with $m_{ij}(t) = m_{ji}(t)$ for all $t \in \mathbb{R}_+$); if the system is started with symmetric initial conditions, they are maintained over all time.

Two sink points were analyzed in the last subsection. First consider equilibrium paths that have m^J as the sink point; they reach m^J in finite time and stay there. Using Figure 1, we will construct a symmetric alternative path δ'_{ij} that dominates the equilibrium path δ_{ij}.

Let t' be the time at which the equilibrium path reaches m^J. Let the planner set $\delta'_{ij}(t) = \delta_{ij}(t)$ for $t \leq t'$, taking the same path as the equilibrium path until t'. At time t', the planner takes $\delta'_{ij}(t) = 0$ until m^I is attained, prohibiting meetings so that the dancers can profit from ideas created in isolation. Then the planner sets $\delta'_{ij}(t) = 1$ until m^J is attained, permitting meetings and the development of more knowledge in common. The last two phases are repeated as necessary.

From Figure 1, the income paths $y_i(t; \delta'_{ij})$ and $y_j(t; \delta'_{ij})$ generated by the path δ'_{ij} clearly dominate the income paths $y_i(t; \delta_{ij})$ and $y_j(t; \delta_{ij})$ generated by the equilibrium path δ_{ij}. Thus, the equilibrium is far from the most productive path in the two person model.

Other dominating paths can be generated by altering this example. For instance, the path in the example above could be replicated until reaching m^B for the first time, followed by alternating rapidly between joint work and work in isolation to stay close to m^B. Notice that when the persons are working in isolation, their productivity is low. Hence, it is not clear which path, the original path that cycles or one that tries to maintain a position near m^B, generates higher welfare in the context of the sum of discounted felicity. To determine which is more efficient, it would be necessary to introduce a discount rate; we avoid that complication here.

Next consider equilibrium paths $\delta_{ij}(t)$ that end in sink point 1/2. Our dominance criterion cannot be used in this situation, since in potentially dominating plans, the planner will need to force the couple to meet outside of region M in Figure 1 in early time periods. During this time interval, the dancers could do better by not meeting, and thus a comparison of the income derived from the paths would rely on the discount rate, something we are trying to avoid. So we will use our weaker criterion here, that of overtaking.

Given an equilibrium path $\delta_{ij}(t)$ with sink point 1/2, the planner can construct an overtaking path $\delta'_{ij}(t)$ in a manner analogous to the construction of a dominating path for paths that end in sink point m^J.

The most productive state m^B is characterized by less homogeneity than the stable point m^J. However, in the present context of two persons, it is not possible to maintain m^B while achieving the highest growth rate $g(m^B)$. For maintaining m^B requires the social planner to force the two persons not to meet some of the time, leading to an income growth rate strictly between α and $g(m^B)$. Thus, it will be surprising to see in the next section that when N is large enough, the equilibrium process will converge to the most productive state and maintain it for a large set of initial conditions.

4 Equilibrium Dynamics

4.1 The General Framework

The model with only two people is very limited. Either two people are meeting or they are each working in isolation. With more people, the dancers can be partitioned into many pairs of dance partners. Within each pair, the two dancers are working together, but pairs of partners are working simultaneously. This creates more possibilities in our model, as the knowledge created within a dance pair is not known to other pairs. Thus, knowledge differentiation can evolve between different pairs of dance partners. Furthermore, the option of switching partners is now available.

We limit ourselves to the case where N is divisible by 4. This is a square dance on the vertices of the Hilbert cube. When the population is not divisible by 4, our most useful tool, symmetry, cannot be used to examine dynamics. Although this may seem restrictive, when N is large, asymmetries apply only to a small fraction of the population, and thus become negligible.[12] In the general case, we impose the assumption of pairwise symmetric initial heterogeneity conditions for all agents.

The initial state of knowledge is symmetric among the dancers, and given by

$$n_{ij}^c(0) = n^c(0) \text{ for all } i \neq j \tag{27}$$

$$n_{ij}^d(0) = n^d(0) \text{ for all } i \neq j \tag{28}$$

At the initial state, each pair of dancers has the same number of ideas, $n^c(0)$, in common. Moreover, for any pair of dancers, the number of ideas that one dancer knows but the other does not know is the same and equal to $n^d(0)$. Given that the initial state of knowledge is symmetric among the four dancers, it turns out that the equilibrium configuration at any time also maintains the basic symmetry among the dancers.[13]

[12] See footnote 18 for further discussion of this point.

[13] It is possible that, beginning from an initial state that is asymmetric, there are asymmetric equilibria in the general case of an arbitrary number of people. (Beginning with an initial state that is symmetric, asymmetric equilibria are impossible.) Our clues about the possibility of asymmetric equilibria come from the two person case, as detailed in Section 3. Analyzing asymmetric equilibria in the general case seems intractable.

When all dancers are pairwise symmetric to each other, that is, when

$$m_{ij}^{d} = m_{ji}^{d} \text{ for all } i \neq j \tag{29}$$

using the function g defined by (22), the income growth rate (18) is simplified as

$$\frac{\dot{y}_i}{y_i} = \frac{\dot{n}_i}{n_i} = \delta_{ii} \cdot \alpha + \sum_{j \neq i} \delta_{ij} \cdot g(m_{ij}^{d}) \tag{30}$$

and the dynamics (19) can be rewritten as

$$\begin{aligned} \frac{\dot{m}_{ij}^{d}}{1 - m_{ij}^{d}} &= \alpha \cdot \left[\delta_{ii} \cdot \left(1 - m_{ij}^{d}\right) - \delta_{jj} \cdot m_{ij}^{d}\right] - \delta_{ij} \cdot m_{ij}^{d} \cdot g(m_{ij}^{d}) \\ &\quad + \left(1 - m_{ij}^{d}\right) \cdot \sum_{k \neq i,j} \delta_{ik} \cdot g(m_{ik}^{d}) - m_{ij}^{d} \cdot \sum_{k \neq i,j} \delta_{jk} \cdot g(m_{jk}^{d}) \end{aligned} \tag{31}$$

Next, taking the case of $N = 4$, we illustrate the possible meetings, noting that the equilibrium path specifies a meeting for every time. Figure 2 gives the possibilities at any fixed time for $N = 4$. Given that the initial state of knowledge is symmetric among the four dancers, as noted above, the equilibrium configuration at any time also maintains the basic symmetry among dancers.

FIGURE 2 GOES HERE

Panel (a) in Figure 2 represents the case in which each of the four dancers is working alone, creating new ideas in isolation. Panels (b-1) to (b-3) represent the three possible configurations of partner dancing, in which two couples each dance separately but simultaneously. In panel (b-1), for example, 1 and 2 dance together. At the same time, 3 and 4 dance together.

Although panels (a) to (b-3) represent the basic forms of dance with four persons, it turns out that the equilibrium path often requires a mixture of these basic forms. That is, on the equilibrium path, people wish to change partners as frequently as possible. The purpose is to balance the number of different and common ideas with partners as best as can be achieved. This suggests a square dance with rapidly changing partners on the equilibrium path.

Please refer to panels (c-1) to (c-3) in Figure 2. Each of these panels represents square dancing where a dancer rotates through two fixed partners as fast as possible in order to maximize the instantaneous increase in their income. In panel (c-1), for example, dancer 1 chooses dancers 2 and 3 as partners, and rotates between the two partners under equilibrium values of δ_{12} and δ_{13} such that $\delta_{12} + \delta_{13} = 1$. In the case where $\delta_{12} = \delta_{13} = 1/2$, for example, this means in practice that person 1 works with person 2 for half a week and with person 3 for half a week. This can be the best of two worlds: it can achieve high knowledge productivity through cooperation while it simultaneously avoids accumulating too much knowledge in common with any particular partner. Dancers 2, 3 and 4 behave analogously. In order for this type of square dance to take place, of course, all four persons must agree to follow this pattern.[14] Finally, panel (d) depicts square dancing in which each dancer rotates though all three possible partners as fast as possible. That is, for all $i \neq j$, $\delta_{ij} \in (0, 1)$, and for all i, $\delta_{ii} = 0$ and $\sum_{j \neq i} \delta_{ij} = 1$.

At this point, it is useful to remind the reader that we are using a myopic core concept to determine equilibrium at each point in time. In fact, it is necessary to sharpen that concept in the model with N persons. *When there is more than one vector of strategies that is in the myopic core at a particular time, namely more than one vector of joint strategies implies the same, highest first derivative of income for all persons, the one with the highest second derivative of income is selected.* The justification for this assumption is that at each point in time, people are attempting to maximize the flow of income. The formal definition of the myopic core and proof that it is nonempty can be found in Appendix 0. Although the theorem in the appendix is general, in the remainder of this paper we shall focus on the symmetric case.

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

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

which is common for all pairs i and j ($i \neq j$). In Figure 1, let m^J and m^I be defined on the horizontal axis at the left intersection and the right intersection between the $g(m)$ curve and the horizontal line at height α, respectively.

4.2 The Main Result

In the remainder of this paper, we assume that

$$\alpha < g(m^B) \tag{32}$$

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

Figure 3 provides a diagram explaining our main result.

[14] In square dancing terminology, this is the "call."

FIGURE 3 GOES HERE

The top horizontal line represents the initial common state $m^d(0)$, while the bottom horizontal line represents the final common state or sink point, $m^d(\infty)$. There are four regions of the initial state that result in four different sink points. Corresponding to each initial region, the associated equilibrium dance forms are illustrated by taking the case with $N = 4$. To be precise:[15]

Proposition 1: *Assume that* N *is a multiple of* 4. *The equilibrium path and sink point depend discontinuously on the initial condition,* $m^d(0)$. *The pattern of interaction between persons and the sink point as a function of the initial condition are given in Figure 3 and as follows.*

(i) For $0 < m^d(0) \leq 2/5 = m^B$*, the equilibrium path consists of an initial time interval (possibly the empty set) in which all* N *persons work independently, followed by an interval in which all persons work with another but trade partners as rapidly as possible (with* $\delta_{ij} = 1/(N-1)$ *for all* i *and for all* $j \neq i$*). When the bliss point,* $2/5$*, is attained, the agents split into groups of* 4*, and they remain at the bliss point.*[16]

(ii) When $m^B < m^d(0) \leq \widehat{m}$*, where* $\widehat{m}$ *is defined by (52), the equilibrium path consists of three phases. First, the* N *persons are paired arbitrarily and work with their partners for a nonempty interval of time. Second, they switch to new partners and work with their new partners for a nonempty interval of time. Finally, each person works alternately with the two partners with whom they worked in the first two phases, but not with a person with whom they have not worked previously. The sink point is* $1/3$.

(iii) For $\widehat{m} < m^d(0) \leq m^I$*, the equilibrium path pairs the* N *persons into*

[15] At this point, it is useful to recall the following notation. In any symmetric situation, the percentage of ideas known by one agent but not another is given by m^J for the lowest value at which meetings are desirable, m^I for the highest value at which meetings are desirable, and m^B for the bliss point or the maximal productivity of a meeting.

[16] The configuration of workers necessary to maintain the bliss point is not unique. Each dancer must have 3 links to other dancers, communicating with each for an equal share of time. For example, groups of 4 may form, where each worker within a group communicates equally with every other worker in that group. However, it is also possible to have, say, groups of six forming. With such groups, each dancer has communication links to only three other dancers within their group. So not all possible links within a group are actually active. If groups at the bliss point are larger, then their communication structure must become more sparse to maintain the bliss point. The minimal size of groups that coalesce at the bliss point is clearly 4. Nevertheless, all of the calculations apply independent of the size of groups that form at the bliss point. The same remarks apply to the various cases detailed below, except when dancers are in isolation.

$N/2$ couples arbitrarily, and each person dances exclusively with the same partner forever. The sink point is m^J.[17]

(iv) For $m^I < m^d(0) \leq 1/2$, each person dances alone forever. The sink point is $1/2$.

4.2.1 Case (i): $0 < m^d(0) \leq 2/5 = m^B$

First suppose that the initial state is such that

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

Then, since $g(m^d_{ij}(0)) = g\left(m^d(0)\right) > \alpha$ for any possible dance pairs consisting of i and j, no person wishes to dance alone at the start. However, since the value of $g(m^d_{ij}(0))$ is the same for all possible pairs, all forms of (b-1) to (d) in Figure 2 are possible equilibrium dance configurations at the start. To determine which one of them will actually take place on the equilibrium path, we must consider the second derivative of income for all persons.

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

$$m^d_{ij} = m^d \text{ for all } i \neq j \tag{33}$$

where

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

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

$$\delta_{ii} = 0 \text{ and } \sum_{j \neq i} \delta_{ij} = 1 \tag{34}$$

Substituting (33) and (34) into (30) yields

$$\frac{\dot{y}_i}{y_i} = g(m^d)$$

Likewise, substituting (33) and (34) into (31) and arranging terms gives

$$\dot{m}^d_{ij} = \left(1 - m^d\right) \cdot g(m^d) \cdot \left[1 - 2m^d - (1 - m^d) \cdot \delta_{ij}\right] \tag{35}$$

Since the income growth rate $\dot{y}/y$ above is independent of the values of δ_{ij} $(j \neq i)$, in order to examine what values of $\delta_{ij} (j \neq i)$ person i wishes to

[17] As in the two person case, once m^J is attained, the couples split and dance alone frequently in order to maintain state m^J.

choose, we must consider the time derivative of $\dot{y}_i/y_i$. In doing so, however, we cannot use equation (30) because the original variables have been replaced. Instead, we must go back to the original equation (18). Then, using equations (33) to (35) and setting $\delta_{ij} = \delta_{ji}$ (which must hold for any feasible meeting), we obtain the following (see Technical Appendix b for proof):

$$\frac{d\left(\dot{y}_i/y_i\right)}{dt} = \left(1 - m^d\right) \cdot g(m^d) \cdot g'(m^d) \cdot \left[1 - 2m^d - (1 - m^d) \cdot \sum_{j \neq i} \delta_{ij}^2\right] \tag{36}$$

Now, suppose that

$$m^d < m^B \equiv 2/5$$

and hence $g'(m^d) > 0$. Then, in order to maximize the time derivative of the income growth rate, person i must solve the following quadratic minimization problem:

$$\min \sum_{j \neq i} \delta_{ij}^2 \quad \text{subject to} \sum_{j \neq i} \delta_{ij} = 1 \tag{37}$$

which yields the solution for person i:

$$\delta_{ij} = \frac{1}{N-1} \quad \text{for all } j \neq i \tag{38}$$

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

The intuition behind this result is as follows. The condition $m^d < 2/5 \equiv m^B$ means that the dancers have relatively too many ideas in common, and thus they wish to acquire ideas that are different from those of each possible partner as fast as possible. That is, when $m^J < m_{ij}^d = m^d < m^B$ in Figure 1, each dancer wishes to move the knowledge composition m_{ij}^d to the right as quickly as possible, thus increasing the growth rate $g(m_{ij}^d)$ as fast as possible. Taking the case of $N = 4$ and using Figure 2, let us consider how this objective can be achieved in a cooperative manner.

Given that $m^J < m_{ij}^d = m^d < m^B$ and thus $\alpha < g(m_{ij}^d)$, dance form (a) in which everyone is dancing alone is out of the question. Dancing alone achieves only the growth rate α less than $g(m_{ij}^d)$; the latter could be achieved by dancing with any other person. Dance in the form (b-1), where $\{1, 2\}$ and $\{3, 4\}$ respectively dance exclusively with one partner, is possible. In this manner, however, each pair just accumulates more knowledge in common, pushing m_{12}^d and m_{34}^d to the left in Figure 1. Indeed, setting $\delta_{12} = 1$ and $\delta_{34} = 1$ in (35) yields

$$\dot{m}_{12}^d = \dot{m}_{34}^d = -(1 - m^d) \cdot g(m^d) \cdot m^d < 0,$$

working against the objective of increasing the myopic growth rate $g(m_{ij}^d)$. Observe that when partners dance in form (b-1), the actual pair $\{1, 2\}$ accumulates more ideas in common. But *from the view point of dancers* 1 *and* 2, dancers 3 and 4 are accumulating new ideas that are different. Consider, for example, the potential partners 1 and 3, who are not dancing together at present and hence $\delta_{13} = 0$. From (35) we have

$$\dot{m}_{13}^d = (1 - m^d) \cdot g(m^d) \cdot (1 - 2m^d) > 0$$

Since g is monotonically increasing on the domain $\left(m^J, 2/5\right)$, *the value* $g(m_{12}^d)$ *of the actual dance partnership* $\{1, 2\}$ *is decreasing with time, while the value* $g(m_{12}^d)$ *of the potential partnership* $\{1, 3\}$ *is increasing with time.* Hence, given the symmetric situation of the four dancers, everyone wants to change partners immediately.

This suggests that when $m^J < m_{ij}^d = m^d < 2/5$ $(= m^B)$ for all $i \neq j$, on the equilibrium path, agents perform a square dance with rapidly changing partners represented by one of panels (c-1) to (d) in Figure 2. Actually, we can show that the square dance configurations (c-1) to (c-3) cannot occur on the equilibrium path. For example, suppose that a dance in the form of panel (c-1) occurs, where $\delta_{12} = \delta_{13} = 1/2$, $\delta_{14} = 0$ and so forth. Then, equation (35) yields

$$\begin{aligned} \dot{m}_{14}^d &= (1 - m^d) \cdot g(m^d) \cdot (1 - 2m^d) \\ &> \dot{m}_{12}^d = \dot{m}_{13}^d = (1 - m^d) \cdot g(m^d) \cdot \frac{1 - 3m^d}{2} \end{aligned}$$

Thus, dancer 1 wants to change partners from 2 and 3 to 4 immediately. Therefore, when $m^J < m_{ij}^d = m^d < 2/5$ $(= m^B)$ for all $i \neq j$, on the equilibrium path, only configuration (d) in Figure 2 can take place, where $\delta_{ij} = 1/3$ for all $i \neq j$.

Returning to the general case with $N \geq 4$, when $m^J < m^d(0) = m_{ji}^d(0) < 2/5$ $(= m^B)$ for all $i \neq j$, on the equilibrium path, the square dance with $\delta_{ij} = 1/(N-1)$ for all $i \neq j$ takes place at the start. Then, since the symmetric condition (33) holds thenceforth, the same square dance will continue as long as $m^J < m^d < 2/5$ $(= m^B)$. The dynamics of this square dance are as follows. The creation of new ideas always takes place in pairs. Pairs are cycling rapidly with $\delta_{ij} = 1/(N-1)$ for all $i \neq j$. Dancer 1, for example, spends $1/(N-1)$ of each period with dancer 2, for example, and $(N-2)/(N-1)$ of the time dancing with other partners. Setting $m^d_{ij} = m^d$ and $\delta_{ij} = 1/(N-1)$ in (35), we obtain

$$\dot{m}^d = (1 - m^d) \cdot g(m^d) \cdot \frac{(N-2) - (2N-3)\, m^d}{N-1} \tag{39}$$

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

$$m^{d*} = \frac{N-2}{2N-3} \tag{40}$$

Surprisingly, when $N = 4$, $m^{d*} = 2/5 = m^B$. The value of $\dot{m}^d$ is positive when $m^d < m^B = 2/5$, and zero if $m^d = 2/5$. Hence, beginning at any point $m^d(0) < 2/5$, the system moves to the right, eventually settling at the bliss point m^B.

Since the right hand side of equation (40) is increasing in N,

$$m^{d*} = \frac{N-2}{2N-3} > 2/5 \equiv m^B \quad \text{when } N > 4. \tag{41}$$

Hence, when $N > 4$ and N is divisible by 4, beginning at any point $m^J < m^d(0) < 2/5$, the system moves to the right and reaches $m^B = 2/5$ in finite time. When N agents reach the bliss point m^B, they break into groups of 4 to maintain heterogeneity at the bliss point.[18]

Next, when $0 \leq m^d(0) < m^J$, it is obvious that the four persons work alone until they reach m^J.[19] Then they follow the path explained above, eventually reaching m^B.

[18] When the number of agents is not divisible by 4, then the bliss point cannot be maintained for the unlucky $N - \tilde{N}$ persons, where $\tilde{N}$ is the largest number divisible by 4 and not exceeding N. Given that our game does not permit any side payments, these unlucky persons have no choice but to do the best by themselves. When $N - \tilde{N} = 3$, the unlucky 3 persons perform a square dance in which they set $\delta_{ij} = 1/3$ for $i \neq j$. Substituting 3 for N in (39) yields

$$\dot{m}^d = \left(1 - m^d\right) \cdot g(m^d) \cdot \frac{1 - 3m^d}{2}$$

Thus, starting from the bliss point, the unlucky 3 persons eventually settle at $m^d = 1/3$. When $N - \tilde{N} = 2$, substituting 2 for N in (39) yields

$$\dot{m}^d = -\left(1 - m^d\right) \cdot g(m^d) \cdot m^d$$

Hence, starting from the bliss point, the unlucky 2 persons gradually move to m^J and stay there. Finally, when $N - \tilde{N} = 1$, this unlucky person dances in solo forever starting from the bliss point. As N becomes larger, however, the fraction of agents for whom the bliss point cannot be maintained becomes small.

[19] Movement to the right beyond m^J requires application of the second order conditions

4.2.2 Case (ii): $m^B < m^d(0) \leq \widehat{m}$ [20]

Next, let us consider the dynamics of the system when it begins to the right of $m^B = 2/5$ but to the left of $\widehat{m} < m^I$ (where $\widehat{m}$ will be defined soon). For example, consider m_0^d in Figure 4, where the $g(m)$ curve from Figure 1 is duplicated in the top part of Figure 4. In other words, the initial state reflects a higher degree of heterogeneity than the bliss point. In this case, the equilibrium process progresses through the following three phases (please refer to the sequence of dance forms (b-1), (b-2) and (c-1), leading to $m^d(\infty) = 1/3$ in the middle of Figure 4).

FIGURE 4 GOES HERE

Phase 1: Since the initial state reflects a higher degree of heterogeneity than the bliss point, the dancers want to increase the knowledge they have in common as fast as possible, leading to fidelity and *couple dances*.

To be precise, since $m_{ij}^d(0) = m^d(0)$ for all $i \neq j$ and $g(m^d(0)) > \alpha$, the situation at time 0 is the same as that in Case (i) except that we now have $m^d(0) > m^B$. Hence, focusing on person i as before, the time derivative of $\dot{y}_i/y_i$ at time 0 is given by (36). However, since $g'(m^d) = g'(m^d(0)) < 0$ at time 0, in order to maximize the right hand side of equation (36), person i now must solve now the following quadratic maximization problem:

$$\max \sum_{j \neq i} \delta_{ij}^2 \text{ subject to } \sum_{j \neq i} \delta_{ij} = 1 \tag{42}$$

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

$$P_1 \equiv \{\{1,2\}, \{3,4\} \{5,6\}, \cdots, \{N-1, N\}\} \tag{43}$$

and initiate pairwise dancing such that[21]

$$\delta_{ij} = \delta_{ji} = 1 \text{ for } \{i,j\} \in P_1, \; \delta_{ij} = \delta_{ji} = 0 \text{ for } \{i,j\} \notin P_1 \tag{44}$$

for equilibrium selection. When the dancers are exactly at m^J, the rate of growth in income from working in isolation and from working with all other dancers with equal intensity are the same, so the derivatives of the respective growth rates of income must be examined to determine which is chosen.

[20] Please note that we have not yet defined $\widehat{m}$. Its definition will appear soon.

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

In order to examine the dynamics for this pairwise dance, let us focus on the partnership $\{1,2\} \in P_1$; the equations for other partnerships are analogous. Since $\delta_{12} = \delta_{21} = 1$ and $\delta_{1k} = \delta_{2k} = 0$ for all $k \neq 1, 2$, setting $i = 1$ and $j = 2$ in (31) yields

$$\dot{m}_{12}^d = -\left(1 - m_{12}^d\right) \cdot m_{12}^d \cdot g(m_{12}^d) < 0 \tag{45}$$

This means, as expected, that the proportion of differential knowledge for each couple decreases with time. Since the dynamics $\dot{m}_{ij}^d$ and the initial point $m_{ij}^d(0) = m^d(0)$ are the same for all $\{i,j\} \in P_1$, as long as the same pairwise dancing continues, we have that

$$m_{12}^d(t) = m_{34}^d(t) = \cdots = m_{N-1,N}^d(t) \equiv m_a^d(t) < m^d(0) \tag{46}$$

where the subscript a in m_a^d means any *actual partnership*.

To study how long the same pairwise dance can continue, let us focus on a *shadow partnership* $\{1,3\} \notin P_1$, which is just a potential partnership for person 1. Since $\delta_{12} = \delta_{34} = 1$ under the present pairwise dance, whereas $\delta_{1k} = 0$ for $k \neq 2$, setting $i = 1$ and $j = 3$ in (31) and using $m_{12}^d = m_{34}^d$ yields

$$\dot{m}_{13}^d = \left(1 - m_{13}^d\right) \cdot \left(1 - 2m_{13}^d\right) \cdot g(m_{12}^d) > 0 \tag{47}$$

implying that the proportion of the differential knowledge increases for any pair of persons who are not dancing together. By symmetry, as long as the same pairwise dancing continues, we have that

$$m^d(0) < m_s^d(t) \equiv m_{13}^d(t) = m_{24}^d(t) = \cdots = m_{ij}^d(t) \quad \text{for all } \{i,j\} \notin P_1 \tag{48}$$

where the subscript s in m_s^d means any *shadow partnership*. Since $g(m)$ is decreasing at $m^d(0) > m^B$, (46) and (48) together mean that the following relationship holds at least initially:

$$g(m_a^d(t)) > g(m^d(0)) > g(m_s^d(t)) \tag{49}$$

Hence, the pairwise dance P_1 will continue at least for a while.

To examine exactly how long the same pairwise dance will continue, let us focus on person 1 again. To see if person 1 continues to dance with person 2 or if person 1 wishes to switch to shadow partner 3, we take the ratio of (45) to (47) at time $t > 0$. Following calculations we obtain:

$$\frac{-\dot{m}_{12}^d(t)}{\dot{m}_{13}^d(t)} > 2 \quad \text{when} \quad m_{12}^d(t) > \frac{2}{5} \equiv m^B$$

The important implication is that $m_{12}^d(t)$ is decreasing at a rate more than twice the speed of increase of $m_{13}^d(t)$, at least initially. Provided that $m^d(0)$ is sufficiently close to $2/5$, eventually there will be a time t' such that $g(m_{12}^d(t')) = g(m_{13}^d(t'))$ and partners change from $\{1,2\}$ and $\{3,4\}$ to, for example, $\{1,3\}$ and $\{2,4\}$.

Indeed, focusing on an actual partnership $\{1,2\}$ and a shadow partnership $\{1,3\}$, we can show the following (see Appendix 1 for a proof of the next result):

Assuming symmetry of initial conditions for N persons, suppose that $2/5 < m^d(0) < m^I$. If initial partnerships are given by P_1 in (43), and the same partnerships are maintained, then there exists a time t' such that for $t > 0$,

$$g(m_{12}^d(t)) \overset{>}{\underset{<}{-}} g(m_{13}^d(t)) \quad \text{as} \quad t \overset{<}{\underset{>}{-}} t' \tag{50}$$

There is a unique the switching time t' as a function of $m^d(0)$, which is denoted by $t^s\left[m^d(0)\right]$. Denoting

$$m_{12}^d\left(t^s\left[m^d(0)\right]\right) \equiv m_{12}^d\left[m^d(0)\right], \; m_{13}^d\left(t^s\left[m^d(0)\right]\right) \equiv m_{13}^d\left[m^d(0)\right]$$

and using the equality in (50) we have the switching position as follows:

$$m_{13}^d\left[m^d(0)\right] = \frac{2}{5} + \frac{\left(m^d(0) - \frac{2}{5}\right)\left(1 - m^d(0)\right)}{m^d(0)^2\left[2 - \left(\frac{1}{m^d(0)} - 2\right)\left(4 - \frac{1}{m^d(0)}\right)\right]} \tag{51}$$

In Figure 4, we draw the $m_{13}^d\left[m^d(0)\right]$ curve in the bottom part (using a bold line). For illustration, we take m_0^d as the initial value of $m^d(0)$ and, using the real lines with arrows, we show in this diagram how to determine the switching positions $m_{13}^d\left[m_0^d\right]$ and $m_{12}^d\left[m_0^d\right]$.

Let $\hat{m}$ be the critical value of $m^d(0)$ such that

$$m_{13}^d\left[\hat{m}\right] = m^I \tag{52}$$

Using Figure 4, we can readily show that $2/5 < \hat{m} < m^I$. Suppose that $2/5 < m^d(0) \leq \hat{m}$. Then, under the partnership $\{1,2\}$ and $\{3,4\}$, it holds that

$$g\left(m_{12}^d(t)\right) > g\left(m_{13}^d(t)\right) > \alpha \quad \text{for} \quad 0 < t < t'$$

and hence partnerships $\{1,2\}$ and $\{3,4\}$ continue until time t'. However, if they maintained the same partnerships longer, then

$$g\left(m_{12}^d(t)\right) < g\left(m_{13}^d(t)\right) \quad \text{for} \quad t > t'$$

This implies that the original partnership cannot be continued beyond time t', and suggests that the dancers switch to the new partnerships. Since all other potential partners are indistinguishable, as shown in Appendix 2, a single dance partner is chosen from those not used in the first phase, and this dance continues for some time.

Two examples of new equilibrium partnerships at time t' are given by

$$P_2 \equiv \{\{1,3\},\{2,4\},\{5,7\},\{6,8\},\cdots,\{N-3,N-1\},\{N-2,N\}\} \tag{53}$$

and

$$P_2' \equiv \{\{N,1\},\{2,3\},\{4,5\},\cdots,\{N-2,N-1\}\} \tag{54}$$

There exist many other possibilities for equilibrium partnerships to be chosen by N dancers at time t'. It turns out, however, that the essential characteristics of equilibrium dynamics are not affected by this choice at time t', as explained at the end of this case.

Phase 2: Hence, let us assume that N persons agree to choose the new partnerships P_2 at time t'. It turns out, however, that these new partnerships last only for a limited time (for details, see Appendix 2). To examine this point, we focus on the dynamics of a four-person group, 1, 2, 3 and 4, where the initial partnerships $\{1,2\}$ and $\{3,4\}$ switch to the new partnerships $\{1,3\}$ and $\{2,4\}$ at time t'. Referring to Figure 4, at the switching time t' the new partnership $\{1,3\}$ is at state $m_{13}^d\left[m_0^d\right]$, whereas the former partnership is at state $m_{12}^d\left[m_0^d\right]$. As the new partnerships mature, partners build up knowledge in common, so m_{13}^d moves to the left from $m_{13}^d\left[m_0^d\right]$, while the former partnership $\{1,2\}$ builds up differential knowledge since they are no longer working together, so m_{12}^d moves to the right from $m_{12}^d\left[m_0^d\right]$. Eventually, the values of the previous partnership $\{1,2\}$ and the current partnership $\{1,3\}$ meet somewhere between $m_{12}^d\left[m_0^d\right]$ and $m_{13}^d\left[m_0^d\right]$. Due to the shape of the function g, the value of partnership $\{1,3\}$ moves up quickly relative to the movement of shadow partnership $\{1,2\}$ up, so equality of the values of the two partnerships is achieved to the left of B at a certain time t''. Let t'' be the time at which $m_{12}^d(t)$ and $m_{13}^d(t)$ become the same:

$$m_{12}^d(t'') = m_{13}^d(t'') \tag{55}$$

Notice that although our focus has been on agent 1, our arguments are applicable to all agents. So, for example, agents 3 and 4 dance in the first phase, whereas agents 2 and 4 dance in the second phase. At the end of the second phase, the values of the four partnerships

$$\{1,2\},\{1,3\},\{3,4\} \text{ and } \{2,4\}$$

are all the same, as are their states.

When equality (55) is achieved, partnerships $\{1,4\}$ and $\{2,3\}$ have never coalesced, so these partners have little in common. Thus, in Figure 4, their state is to the right of the initial state $m_{12}^d\left[m_0^d\right]$, so they will never coalesce. Although we focus on the four person model, the same applies to any potential partnership that never forms in the first two phases. Thus the third phase will involve only those partnerships realized in the first two phases.

Analogous to the notation used for the first switching time and position, there is a unique switching time t'' as a function of $m^d(0)$, which is denoted by $\widetilde{t^s}\left[m^d(0)\right]$. Analogous to the notation used for the first switching position, the second switching position is the unique state where $m_{12}^d(t)$ meets $m_{13}^d(t)$, which is defined by:

$$\tilde{m}^d\left[m^d(0)\right] \equiv m_{12}^d\left(\tilde{t}^s\left[m^d(0)\right]\right) = m_{13}^d\left(\tilde{t}^s\left[m^d(0)\right]\right) = m_{34}^d\left(\tilde{t}^s\left[m^d(0)\right]\right) = m_{24}^d\left(\tilde{t}^s\left[m^d(0)\right]\right) \tag{56}$$

Using the equality in (55), the switching position $\tilde{m}^d\left[m^d(0)\right]$ is:

$$\tilde{m}^d\left[m^d(0)\right] = \frac{2}{5} - \frac{m^d(0) - \frac{2}{5}}{5m^d(0) - 1} \tag{57}$$

In Figure 4, the $\tilde{m}^d\left[m^d(0)\right]$ curve is represented in the bottom part by a bold, broken line. Taking m_0^d as the initial value of $m^d(0)$, and using the broken lines with arrows, we demonstrate how to determine the second switching position $\tilde{m}^d\left[m^d(0)\right]$.

Phase 3: To see what form of dance will take place immediately after the second switching time, observe that if partnerships $\{1,3\}$ and $\{2,4\}$ were maintained beyond time t'', then it would follow that

$$g\left(m_{12}^d(t)\right) > g\left(m_{13}^d(t)\right) \quad \text{for } t > t'' \tag{58}$$

This implies that the same partnerships cannot be continued beyond t''. Furthermore, notice that dancers cannot go back to the previous form of partnerships $\{1,2\}$ and $\{3,4\}$. If they did so, then the proportion of the knowledge in common for the actual partners $\{1,2\}$ would increase, while the proportion of the differential knowledge for the shadow partnership $\{1,3\}$ would increase. This means that the following relationship,

$$m_{12}^d(t) < m^d(t'') < m_{13}^d(t) < m^B$$

holds immediately after t'', and thus

$$g\left(m_{12}^d(t)\right) < g\left(m_{13}^d(t)\right) \tag{59}$$

which contradicts the assumption that $\{1,2\}$ is the actual partnership. Furthermore, since dancers 1 and 4 do not meet before t'', the following inequality

$$g\left(m_{13}^d(t)\right) > g\left(m_{14}^d(t)\right) \tag{60}$$

holds immediately after t''. Thus, immediately after time t'', the equilibrium dance cannot include partnerships $\{1,4\}$ and $\{2,3\}$. Hence, provided that $g(1/3) > \alpha$, we can see from Figure 2 that the only possible equilibrium configuration immediately after t'' is a square dance in the form (c-1), involving a rapid rotation of non-diagonal partnerships, $\{1,2\}$, $\{1,3\}$, $\{2,4\}$ and $\{3,4\}$. That is, for dancer 1, $\delta_{11} = 0$ and $\delta_{1j} = \frac{1}{2}$ if $j = 2$ or 3, $\delta_{14} = 0$.[22] Analogous expressions hold for all other four-person groups, $\{5,6,7,8\}, \cdots, \{N-3, N-2, N-1, N\}$.

The dynamics for this square dance are as follows. We set

$$m_{ij}^d \equiv m^d \text{ for } \{i,j\} \in P_2 \tag{61}$$

Then, since conditions (33) and (34) hold also in the present context, setting $\delta_{ij} = 1/2$ in (35), we get

$$\dot{m}^d = (1 - m^d) \cdot g(m^d) \cdot \frac{1 - 3m^d}{2} \tag{62}$$

which is negative when $m^d > \frac{1}{3}$, and zero if $m^d = \frac{1}{3}$. Thus, beginning at any point $m^d(t'') > \frac{1}{3}$, the system moves to the left, eventually settling at $m^d = \frac{1}{3}$.

We can readily show that, along the path above, relation (60) holds for all $t \geq t''$ where $m_{13}^d(t) \equiv m^d(t)$.

It is interesting to observe that, in the entire equilibrium process starting with the symmetric state of knowledge such that $m_{ij}^d(0) = m^d(0) > m^B$ for all $i \neq j$, partnerships $\{1,4\}$ and $\{2,3\}$, for example, never coalesce. That is, given that the proportion of differential knowledge for all pairs of dancers at the start exceeds the most productive point m^B, they try to increase the proportion of knowledge in common as quickly as possible through partner dancing. These initial stages of building up knowledge in common through partner dancing, however, divide all possible pairs of partners, who were symmetric at the start, into two heterogenous groups: those pairs that developed a sufficient proportion of knowledge in common through actual meetings, and those pairs that increased further the proportion of exclusive knowledge because they did not have a chance to work together. Since the latter group of

[22] Similar to Case (i), this result can also be obtained from the second order condition for equilibrium selection.

potential partners is excluded from the square dance in the last stage, the equilibrium process of the four-person system ends up with a state of knowledge that is less than the most productive state.

Finally, we may note that there exist many different structures of equilibrium partnerships to be chosen at time t'. However, the choice does not affect, in the following two stages, the dynamics of m^d (the proportion of differential knowledge, common to all active partnerships). Indeed, choosing in Phase 2 either partnerships P_2 or P_2', in Phase 3 each person wants to dance with the two partners whom the person met in the previous two stages. Thus, whether P_2 or P_2' is chosen at time t', the dynamics of m^d are the same in the last two stages.

4.2.3 Case (iii): $\widehat{m} < m^d(0) \leq m^I$

Next suppose $m^d(0)$ is such that $\hat{m} < m^d(0) \leq m^I$. As in Case (ii), dancers are more heterogeneous than at the bliss point, so they would like to increase the knowledge they hold in common through couple dancing, for example using configuration (b-1) in Figure 2. The initial phase of Case (iii) is the same as the initial phase of Case (ii). However, using (51), we know that $m_{13}^d[m^d(0)] > m^I$. Thus, $g\left(m_{12}^d(t)\right) > g\left(m_{13}^d(t)\right)$ for all t before $m_{12}^d(t)$ reaches m^J, whereas $g\left(m_{12}^d(t)\right) > \alpha > g\left(m_{13}^d(t)\right)$ when $m_{12}^d(t)$ reaches m^J. So each dancer keeps their original partner as the system climbs up to B and on to J. When the system reaches $m^d(t) = m^J$, each dancer uses fractional δ_{ij} to attain m^J by switching between working in isolation and dancing with their original partner.

4.2.4 Case (iv): $m^I < m^d(0) \leq 1/2$

Finally, suppose $m^d(0) > m^I$. Then, $g\left(m^d(0)\right) < \alpha$, and hence there is no reason for anyone to form a partnership. Thus, each person dances alone forever, and eventually reaches $m^d = 1/2$.

Compiling all four cases, the *Main Result* follows.

There are important remarks to be made about our *Main Result*. First, the sink point changes discontinuously with changes in the initial conditions. Second, from each set of initial conditions, the N persons eventually divide into many separate groups between which no interaction occurs.[23] Thus, from an initial state that is symmetric, we obtain an equilibrium path featuring asymmetry.

[23] Of course, Case (i) is the most interesting of these.

5 Efficiency: The General Case

Next we consider the welfare properties of the equilibrium path. We examine each of the cases enumerated above, beginning with Case (iii). This Case is quite analogous to the two person model with sink point m^J, and essentially the same argument implies that the equilibrium path can be dominated. What distinguishes this case is the fact that at the sink point, meeting and not meeting have the same one period payoff for all persons. Thus, the social planner can change δ_{ij} for a length of time without changing payoffs, but after this length of time, payoffs can be made higher, as illustrated in Section 3.2.

Now consider Case (iv). The equilibrium cannot be dominated. It has each person always working in isolation. Thus, $m^d(0)$ lies in $(m^I, \frac{1}{2}]$ and m^d moves right with time. If there were a dominating path, then the social planner must force some pair to work together over a non-trivial interval of time. The first such interval of time will have values of m^d in $(m^I, \frac{1}{2}]$, so the persons working together will have lower income during this interval, contradicting the assumption of domination.

Consider Case (i). Let $\delta_{ij}(t)$ be the equilibrium path. When $m^d(0) > m^J$, $\delta_{ij}(t) = 1/(N-1)$ for all t and for all pairs i and j, and the payoffs from meeting always exceed not meeting for any person. Examining equation (36) and the implied optimization problem (37), this is the unique path of meetings that maximizes income over each non-negligible interval of time. So the equilibrium path is not dominated by any other feasible path. Furthermore, the equilibrium path either approaches (when $N = 4$) or reaches in finite time (when $N > 4$) the most productive state, m^B. When $m^d(0) \leq m^J$, similar to Case (iv), strict domination cannot occur when $m^d \leq m^J$. The equilibrium path begins at $m^d(0)$ and reaches m^J in finite time. Combining this with what we have determined about the equilibrium path starting at $m^d(0) > m^J$, we obtain that the equilibrium path is not dominated, and approaches the most productive state.

In fact, for case (i) when $m^d(0) > m^J$, there is a much stronger efficiency result. As detailed in Appendix 3, the equilibrium path coincides with a utilitarian socially optimal path where the planner has foresight.

Finally, consider Case (ii), when $m^B < m^d(0) \leq \widehat{m}$. Examining equation (36) and the implied optimization problem (42), this path of meetings maximizes income over each non-negligible interval of time. So the equilibrium path is not dominated by any other feasible path, but unlike Case (i), it approaches $m^d = 1/3$, that is not the most productive state.

Clearly, initial heterogeneity plays an important role in the efficiency properties of the equilibrium path. What distinguishes Case (i), aside from a relatively homogeneous beginning, is that the dancers can switch partners rapidly enough to increase heterogeneity while at the same time maximizing the increase in output. That is because each agent spends $1/(N-1)$ of the time dancing with any particular agent, and $(N-2)/(N-1)$ of the time dancing with others. This is what leads to the most productive state. In other cases, efficiency would require less heterogeneity than in the initial state, which can only be attained by dancing with a restricted set of partners. This builds up an asymmetry in an agent's relationship with others, in that the agent has more in common with those they have danced with previously, and makes the most productive state unattainable without foresight. It also explains how, with a large initial heterogeneity of agents, asymmetry in their relationships is introduced and is built on along the equilibrium path.

6 Why 4?

We have seen that once the agents reach the bliss point (where the growth rate is highest), achieved from large initial homogeneity by cycling through all partners as rapidly as possible, they break into groups of 4 (see Proposition 1, part (i)). This dance pattern allows them to remain at the highest productivity forever. It is natural to ask why 4 is the magic number. In order to see this, we must place the model in a more general context. In particular, we generalize our joint knowledge creation function (7) as follows:

$$a_{ij} = \beta \cdot (n^c_{ij})^{\theta} \cdot (n^d_{ij} \cdot n^d_{ji})^{\frac{1-\theta}{2}} \qquad 0 < \theta < 1$$

The parameter θ represents the weight on knowledge in common as opposed to differential knowledge in the production of new ideas. This parameter is crucial in determining the bliss point. Of course, up to now, we have set $\theta = 1/3$. The remainder of the model is unchanged.

First we calculate the bliss point in this more general setting. Analogous to equation (22), the growth rate function under pairwise symmetry is modified as follows:

$$g(m) \equiv \beta \cdot \frac{(1-2m)^{\theta} \cdot m^{(1-\theta)}}{1-m} \tag{63}$$

Setting $g'(m) = 0$, the bliss point m^B is given by

$$m^B(\theta) = \frac{1-\theta}{2-\theta} \tag{64}$$

As expected, when $\theta = 1/3$, $m^B = 2/5$. For $\theta = 0$, $m^B = 1/2$; m^B decreases monotonically in θ, reaching $m^B = 0$ when $\theta = 1$, which is not surprising.

When all the agents start with initial heterogeneity $m^d(0) < m^B(\theta)$ and are pairwise symmetric, equilibrium dynamics are essentially the same as in Proposition 1 (i). Thus, analogous to the previous derivation of (35), when $m^J < m < m^B(\theta)$, using (33), (34) and (31) we have the following equilibrium dynamics:

$$\dot{m}^d = \{\frac{N-2}{N-1} - m^d \cdot \frac{2N-3}{N-1}\} \cdot \beta \cdot (1-2m^d)^\theta \cdot (m^d)^{1-\theta} \tag{65}$$

Setting $\dot{m}^d = 0$, we obtain the sink point

$$m^{d*} = \frac{N-2}{2N-3} \tag{66}$$

which is independent of θ and is 2/5 when $N = 4$. Since m^{d*} is increasing in N, for sufficiently large N, the sink point heterogeneity exceeds the bliss point heterogeneity, namely $m^{d*} > m^B(\theta)$. So when the equilibrium heterogeneity reaches the bliss point $m^B(\theta)$, the agents must split into smaller groups in order to maintain the optimal level of heterogeneity, $m^B(\theta)$. To analyze the optimal group size, we set the heterogeneity of the sink point of the dynamic process to the heterogeneity at the bliss point:

$$m^{d*} = m^B(\theta)$$

Thus, using (64) and (66), we obtain the optimal group size

$$N^B(\theta) = 1 + \frac{1}{\theta} \tag{67}$$

which is 4 when $\theta = 1/3$, as expected. Assuming that the optimal group size $N^B(\theta)$ and the number of groups $N/N^B(\theta)$ are integers, when the equilibrium dynamics reach $m^B(\theta)$, groups of size $N^B(\theta)$ form and each member of a group dances only with members of the group, spending an equal amount of time dancing with every member of the group with $\delta_{ij}(t) = 1/(N^B(\theta) - 1)$.

Equilibrium dynamics when initial heterogeneity is larger than m^B are essentially unchanged from Proposition 1 (ii)-(iv), but explicit solutions are not readily obtainable. Nevertheless, the equilibrium dance patterns and intuition are robust.

The main implication of this analysis is that if knowledge in common is important (θ is close to 1), the equilibrium and optimal grouping of dancers is rather small. This may explain the large number of small firms in Higashi

Osaka or in Ota ward in Tokyo, each specializing in different but related manufacturing services. Another example is the third Italy, which produces a large variety of differentiated products. In each case, tacit knowledge accumulated within firms plays a central role in the operation of the firms. (An extreme example is marriage, when $N^B(\theta) = 2$.) In contrast, when differentiated knowledge is important (θ is close to 0), then the equilibrium and optimal group size is large, for example in academic departments and research labs.

We may also observe that when the system reaches the bliss point, the dancers break into groups and the system becomes asymmetric, in the following sense. If dancer i belongs to the same group as dancer k, then their differential knowledge remains at the bliss point m^B, maintaining the highest productivity $g(m^B)$. If dancer j belongs to a different group, then the differential knowledge between i and j diverges, namely it moves away from m^B, thus reducing $g(m^d_{ij})$. So once the population splits into groups, dancers i and j will not want to collaborate again. In other words, our model generates boundaries of research teams endogenously.

7 Conjectures and Conclusions

We have considered a model of knowledge creation that is based on individual behavior, allowing myopic agents to decide whether joint or individual production is best for them at any given time. We have allowed them to choose their best partner or to work in isolation. This is a pure externality model of knowledge creation. One would not expect that equilibria would be efficient for two reasons: the agents are myopic, and there are no markets. The emphasis of our model is on endogenous agent heterogeneity, whereas we examine the permanent effects of knowledge creation and accumulation.

In the case of two people, there are two sink points (equilibria) for the knowledge accumulation process. The state where the two agents have a negligible proportion of ideas in common is attainable as an equilibrium from some initial conditions. There is one additional and more interesting sink, involving a large degree of homogeneity in the two agents, and this is attainable from a non-negligible set of initial conditions. Relative to the most productive state, the first sink point has agents that are too heterogeneous, while the second sink point has agents that are too homogeneous.

With N persons, assuming that N is large enough, we find that, surprisingly, for a range of initial conditions that imply a large degree of homogeneity among agents, the sink is the most productive state. The population breaks into optimal size groups when it reaches the most productive state. The size of these groups is inversely related to the weight given to homogeneity in knowledge production.

The sink point depends discontinuously on initial conditions. Moreover, there are only 4 possible equilibrium paths. If agents begin with a large degree of heterogeneity, then the sink is inefficient, and it can be one of several points, including the analog of the relatively homogeneous sink in the two person case. Despite a symmetric set of initial conditions, asymmetries can arise endogenously in our structure. In particular, each agent might communicate pairwise with some, but not all other, agents in equilibrium. The asymmetries that arise can partition the agents endogenously into different groups, giving rise to an asymmetric interaction structure from a situation that is initially symmetric. Bearing in mind its limitations, the model could be tested using data on coauthorships in various academic disciplines or collaborative work in other fields. Returning to a question posed in the introduction, the empirical pattern of team size in Broadway musicals, from smaller teams to larger ones over a 50 year period, is explained by the increasing complexity of producing these musicals, and thus a need for more heterogeneity in the teams.

Many extensions of our work come to mind. It is important and interesting to add knowledge transfer to the model with more than 2 people. Then we can study comparative statics with respect to speeds of knowledge transfer and knowledge creation on the equilibrium outcome and on its efficiency. It would also be interesting to add knowledge transfer without meetings, similar to a public good. For instance, agents might learn from publicly available sources of information, like newspapers or the web.[24] Markets for ideas would also be a nice feature. One set of extensions would allow agents to decide, in addition to the people they choose with whom to work, the intensity of knowledge creation and exchange.

We note that what we have done, in essence, is to open the "black box" of knowledge externalities in more aggregate models to find smaller "black boxes" inside that we use in our model. These "black boxes" are given by the exogenous functions representing knowledge transfer and creation within a meeting of two agents. Our contribution is to solve the matching problem of agents in a dynamic context given this structure. It will be important to

[24]Stability of our equilibria with respect to small amounts of public information or information spillovers is an important topic for future research.

open our “black boxes” as well. That is, the microstructure of knowledge creation and transfer within meetings must be explored. It will be useful to proceed in the opposite direction as well, aggregating our model up to obtain an endogenous growth framework, to see if our equilibrium patterns and efficiency results persist.

Another set of extensions would be to add stochastic elements to the model, so the knowledge creation and transfer process is not deterministic. As remarked in the introduction, probably our framework can be developed from a more primitive stochastic model, where the law of large numbers is applied to obtain our framework as a reduced form.[25]

An important application of our work would be to the literature on intellectual property, to provide microstructure for the idea production process; see Scotchmer (2004) and Boldrin and Levine (2005) for interesting and provocative treatments.

Eventually, we must return to our original motivation for this model, as stated in the introduction. Location seems to be an important feature of knowledge creation and transfer, so regions and migration are important, along with urban economic concepts more generally; for example, see Duranton and Puga (2001) and Helsley and Strange (2004). Using patent data, Agrawal *et al* (2003) find that when an inventor moves, he cites patents from his previous location more than patents at other locations. The reason could be connected with mutual knowledge. It would be very useful to extend the model to more general functional forms. It would be interesting to proceed in the opposite direction by putting more structure on our concept of knowledge, allowing asymmetry or introducing notions of distance, such as a metric, on the set of ideas[26] or on the space of knowledge. Finally, it would be useful to add vertical differentiation of knowledge, as in Jovanovic and Rob (1989), to our model of horizontally differentiated knowledge.

REFERENCES

Agrawal, A.K., Cockburn, I.M., McHale, J., 2003. Gone but not forgotten: Labor flows, knowledge spillovers, and enduring social capital. NBER Working Paper 9950 http://www.nber.org/papers/w9950

[25] We confess that our first attempts to formulate our model of knowledge creation were stochastic in nature, using Markov processes, but we found that they quickly became intractable.

[26] See Berliant *et al* (2006).

Aoki, M., 1994. The contingent governance of teams: Analysis of institutional complementarity. International Economic Review 35, 657-676.

Barabási, A.-L., 2005. Network theory - the emergence of creative enterprise. Science 308, 639-641.

Berliant, M., Fujita, M., 2006. Dynamics of knowledge creation and transfer: The two person case. http://econpapers.hhs.se/paper/wpawuwpga/0401004.htm

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

Boldrin, M., Levine, D.K., 2005. Against Intellectual Monopoly. Mimeo.

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

Guimerà, R., Uzzi, B., Spiro, J., Amaral, L.A.N., 2005, Team assembly mechanisms determine collaboration network structure and team performance. Science 308, 697-702.

Helsley, R.W. and W.C. Strange, 2004. Knowledge barter in cities. Journal of Urban Economics 56, 327-345.

Hildenbrand, W., Kirman, A., 1976. Introduction to Equilibrium Analysis. North Holland/American Elsevier, Amsterdam.

Holmstrom, B., 1982. Moral hazard in teams. The Bell Journal of Economics 13, 324-340.

Jovanovic, B., Rob, R., 1989. The growth and diffusion of knowledge. The Review of Economic Studies 56, 569-582.

Kaplinsky, R., 1983. Firm size and technological change in a dynamic context. The Journal of Industrial Economics 32, 39-59.

Klein, E., Thompson, A.C., 1984. Theory of Correspondences. John Wiley & Sons, New York.

Léonard, D., Van Long, N., 1992. Optimal Control Theory. Cambridge University Press, Cambridge.

Scotchmer, S., 2004. Innovation and Incentives. MIT Press, Cambridge, MA.

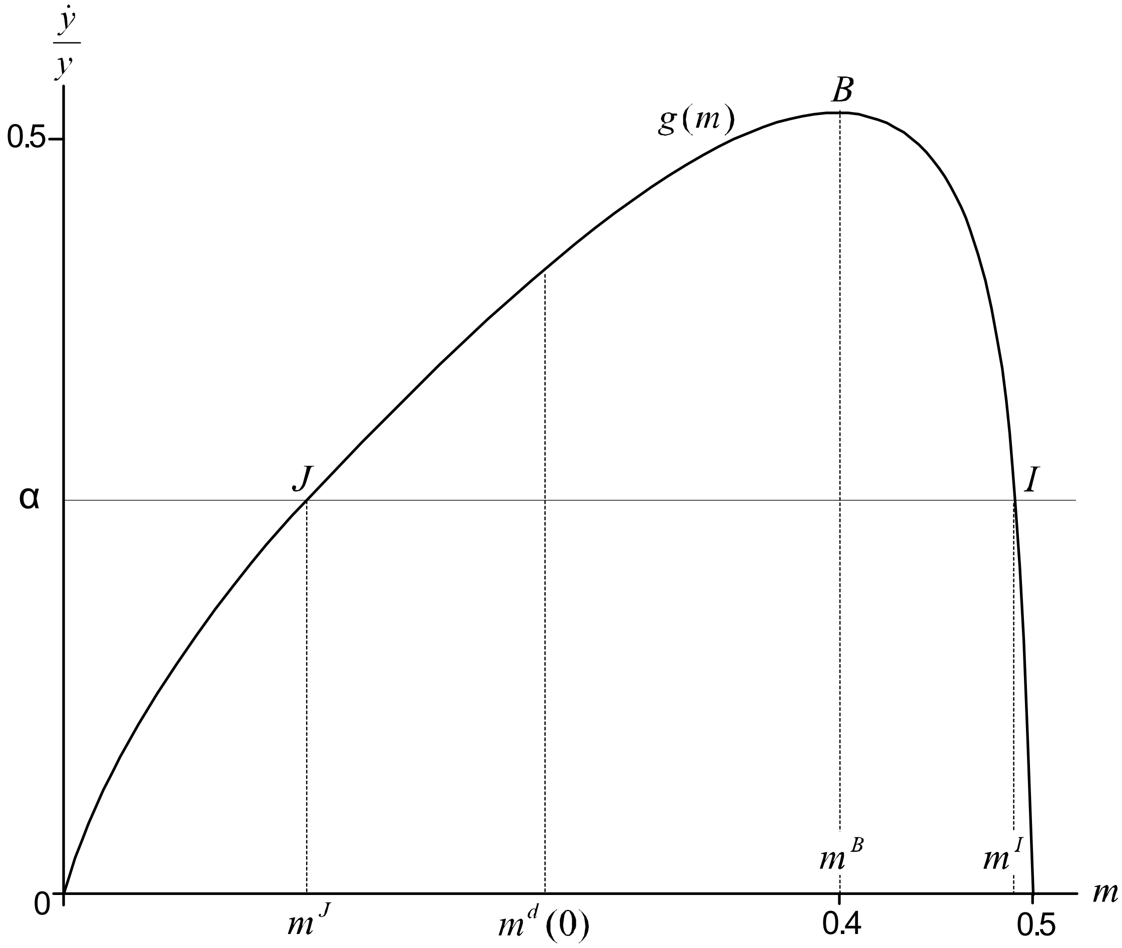

Figure 1: The $g(m)$ curve and the bliss point when $\beta = 1$.

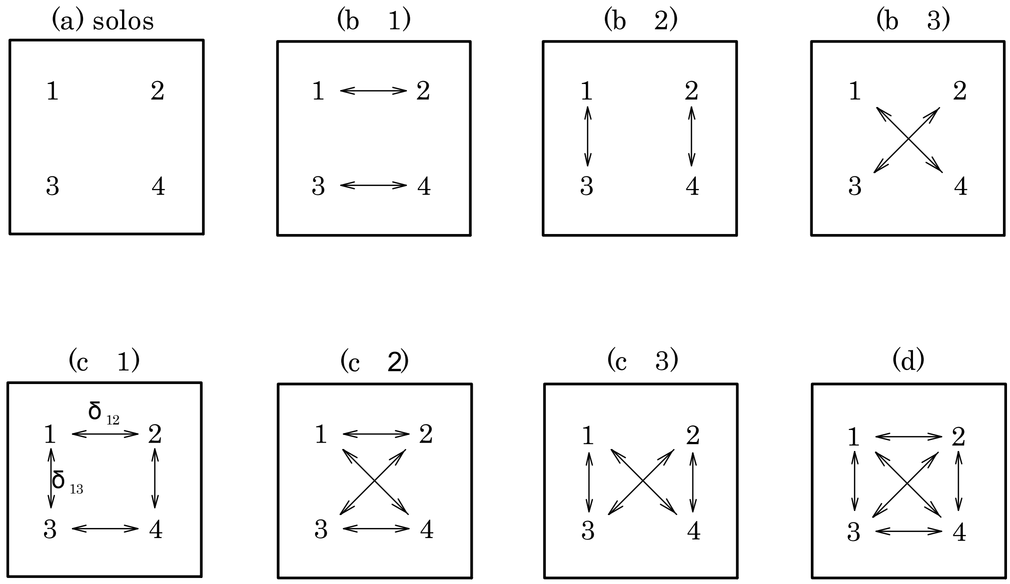

Figure 2: Possible meetings when $N = 4$.

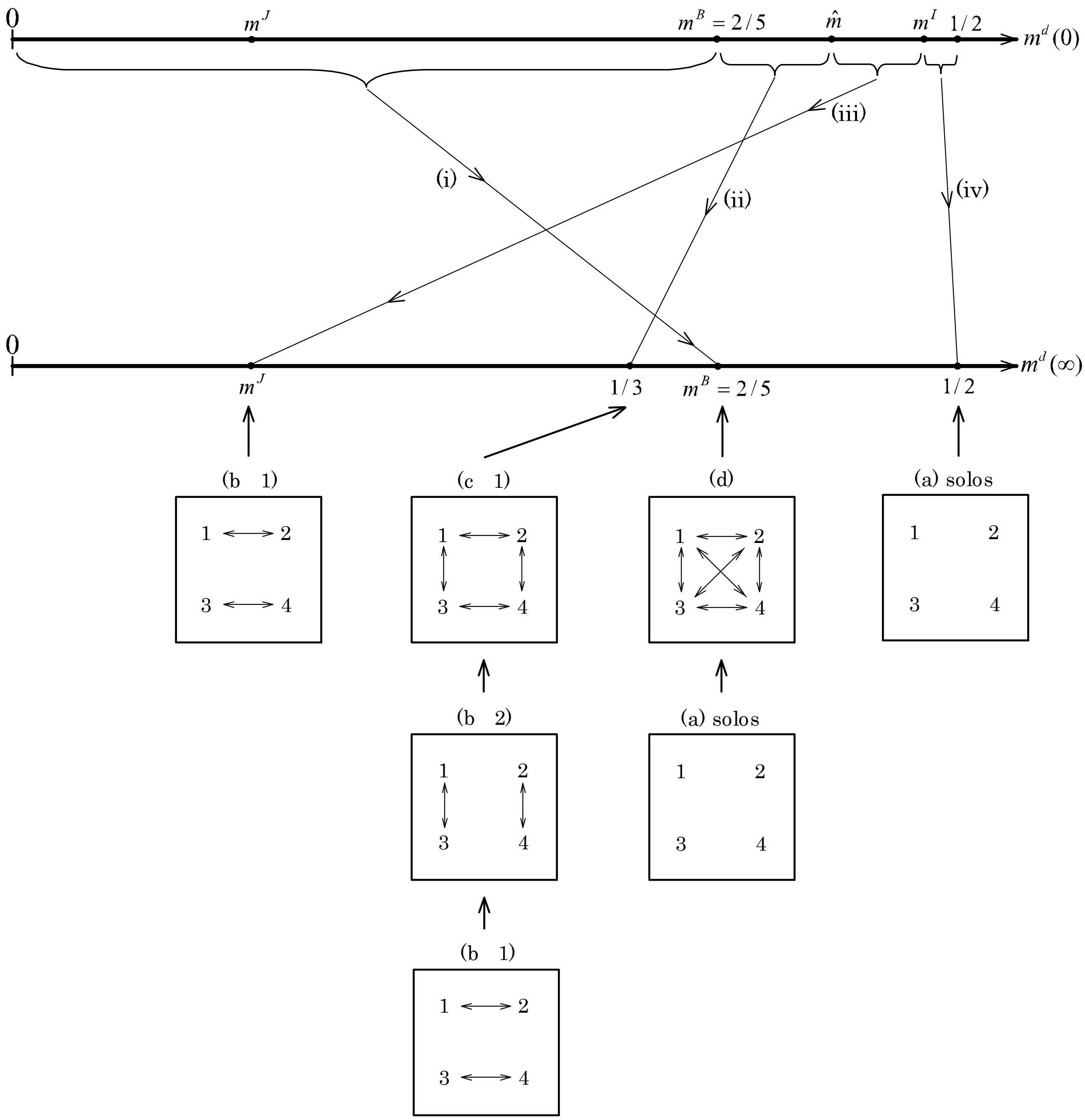

Figure 3: Correspondence between the initial point $m^d(0)$ and the long-run equilibrium point $m^d(\infty)$.

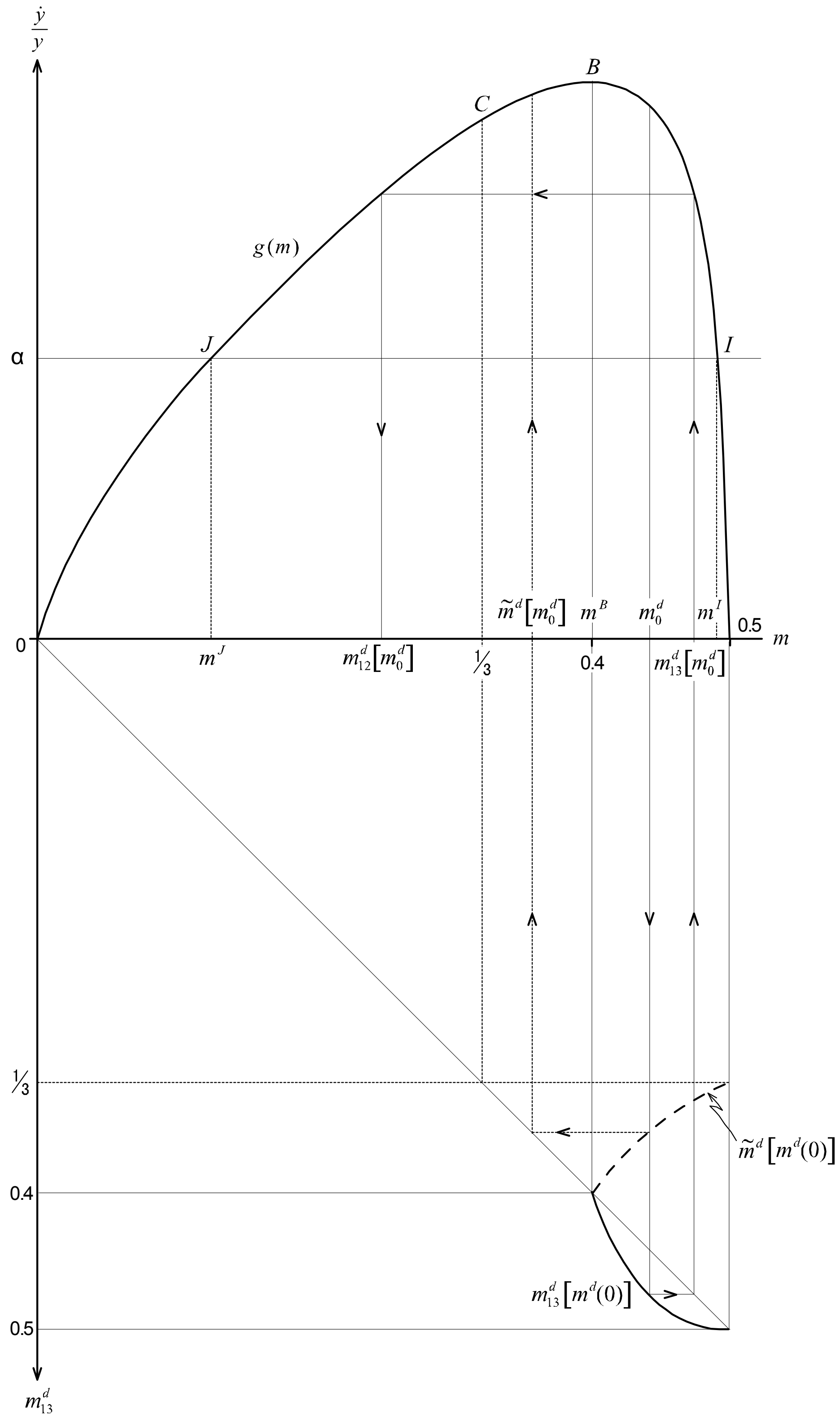

Figure 4: (a) Real lines with arrows: the $m_{13}^d\left[m^d(0)\right]$ curve and the determination of the switching positions $m_{13}^d\left[m_0^d\right]$ and $m_{12}^d\left[m_0^d\right]$. (b) Broken lines with arrows: the $\tilde{m}^d\left[m^d(0)\right]$ curve and the switching position $\tilde{m}^d\left[m_0^d\right]$.

8 Appendix 0: Definition and Nonemptiness of the Myopic Core

Definitions: We say that measurable paths $\delta_{ij} : \mathbb{R}_+ \to [0,1]$ for $i, j = 1, ..., N$ are *feasible* if for all $t \in \mathbb{R}_+$, $\sum_{i=1}^N \delta_{ij} = 1$ for $j = 1, ..., N$; $\sum_{j=1}^N \delta_{ij} = 1$ for $i = 1, ..., N$; $\delta_{ij} = \delta_{ji}$ for $i = 1, ..., N$, $j = 1, ..., N$. We associate with any feasible paths $\{\delta_{ij}\}$ continuous functions $a_{ij} : \mathbb{R}_+ \to \mathbb{R}_+$ for $i = 1, ..., N$, $j = 1, ..., N$, satisfying the equations of motion (6), (7), (12), and (13) where δ_{ij} is permitted to take on fractional values. For notational simplicity, we omit the paths δ_{ij} as arguments in the functions a_{ij}. For any coalition $S \subseteq \{1, ..., N\}$, let $D_S = \{(d_{ij})_{i,j \in S}$ with $d_{ij} \in [0,1]$ for all $i, j \in S$, $\sum_{i \in S} d_{ij} = 1$ for $j \in S$; $\sum_{j \in S} d_{ij} = 1$ for $i \in S$; $d_{ij} = d_{ji}$ for $i = 1, ..., N$, $j = 1, ..N\}$. Paths $\{\delta_{ij}\}$ are in the *myopic core* if they are feasible and at each time $t \in \mathbb{R}_+$, there is no coalition $S \subseteq \{1, ..., N\}$ and $(d_{ij})_{i,j \in S} \in D_S$ such that for all $i \in S$ $\sum_{,j \in S} d_{ij} a_{ij}(t) > \sum_{,j \in S} \delta_{ij}(t) a_{ij}(t)$.

Theorem 0: The myopic core is nonempty. Moreover, if $N = 2$, there is a myopic core path with $\delta_{ij}(t) \in \{0,1\}$.

Proof of Theorem 0: For any fixed time t and any coalition S if we define $V(S) = \{(u_1, ..., u_N) \in \mathbb{R}^N \mid \exists (d_{ij})_{i,j \in S} \in D_S$ such that $\forall i \in S$ $u_i \leq \sum_{j \in S} d_{ij} a_{ij}(t)\}$ then V defines a nontransferable utility game in characteristic function form. Next we show that the myopic core is nonempty. To accomplish this, we show that the game at each period t is balanced and apply Scarf's theorem (see Hildenbrand and Kirman, 1976, p. 71). Let $\mathcal{S}$ be a balanced family of coalitions and let w_S $(S \in \mathcal{S})$ be the balancing weights. So for each i, $\sum_{\{S \in \mathcal{S} | i \in S\}} w_S = 1$. Let $(u_1, ..., u_N) \in \cap_{S \in \mathcal{S}} V(S)$. So for each $S \in \mathcal{S}$, for each $i \in S$, there exists $(d^*_{ij}(S))_{i,j \in S} \in D_S$ with $u_i \leq \sum_{j \in S} d^*_{ij}(S) a_{ij}(t)$. Then for each $i = 1, ..., N$

$$u_i \leq \sum_{S \in \mathcal{S}} w_S \sum_{i,j \in S} d^*_{ij}(S) a_{ij}(t) = \sum_{j=1}^N \sum_{\{S \in \mathcal{S} | i,j \in S\}} w_S \cdot d^*_{ij}(S) a_{ij}(t).$$

Then $(\sum_{\{S \in \mathcal{S} | i,j \in S\}} w_S \cdot d^*_{ij}(S))_{i,j=1}^N \in D_{\{1,...,N\}}$. Hence by definition of $V(\{1, ..., N\})$, $\sum_{j=1}^N \sum_{\{S \in \mathcal{S} | i,j \in S\}} w_S \cdot d^*_{ij}(S) a_{ij}(t) \in V(\{1, ..., N\})$, and the game is balanced. Applying Scarf's theorem, the core at each time t is nonempty. Using a standard selection result (Klein and Thompson, 1984, p. 163), since we know that the correspondence from time t to myopic core at that time is closed valued, we can select from it a measurable myopic core path.

If $N = 2$ and if $\delta_{ij}(t) \in (0,1)$, then it must be the case that $a_{ij}(t) = a_{ii}(t) = a_{jj}(t)$. So without loss of generality, we can take $\delta_{ij}(t) \in \{0,1\}$.$\blacksquare$

Remark: In the case where the derivative of the percentage increase in income is used as a further refinement of myopic core, the same type of result holds. Simply fix a time t, add $\epsilon > 0$ to those $a_{ij}(t)$ with highest second derivatives, apply the proof, and let ϵ tend to 0. We obtain a sequence of $\delta_{ij}(t)$ vectors that are in the core of the modified game at time t. As the vectors of feasible $\delta_{ij}(t)$ lie in a compact set, a convergent subsequence can be drawn that has a limit in the refined core at time t. Again, the refined myopic core is closed valued, so we can select from it a measurable refined myopic core path.

9 Appendix 1: Analysis of Phase 1

Lemma 1: *Assuming symmetry of initial conditions for four persons, suppose that* $2/5 < m^d(0) < 1/2$. *If initial partnerships are given by* P_1 *in* (43)*, and the same partnerships are maintained, then there exists a time* t' *such that for* $t > 0$,

$$g(m^d_{12}(t)) \gtreqless g(m^d_{13}(t)) \quad \text{as} \quad t \lesseqgtr t' \tag{68}$$

and the following relationship holds at time t'*:*

$$m^d_{13}(t') = \frac{2}{5} + \frac{\left(m^d(0) - \frac{2}{5}\right)\left(1 - m^d(0)\right)}{m^d(0)^2\left[2 - \left(\frac{1}{m^d(0)} - 2\right)\left(4 - \frac{1}{m^d(0)}\right)\right]} \tag{69}$$

Proof of Lemma 1: Under the partnerships P_1 in (43), first we show that there exists a unique time $t' > 0$ such that

$$g(m^d_{12}(t')) = g(m^d_{13}(t')) \tag{70}$$

To show this, we make a few preliminary observations. First, for any $i \neq j$ at any time, since $n^d_{ij} = n^d_{ji}$ means $m^d_{ij} = m^d_{ji}$, going backward through the last part of the calculations in the proof of Theorem A1 (Technical Appendix a), and recalling the definition of the function $g(m)$ and a_{ij} for $i \neq j$, we can readily show that

$$g(m^d_{ij}) = \frac{\beta \cdot \left[n^c_{ij} \cdot (n^d_{ij})^2\right]^{\frac{1}{3}}}{n_i} \quad \text{when } n^d_{ij} = n^d_{ji} \tag{71}$$

Next, under the partnerships P_1, since $\delta_{1k} = 0$ for all $k \neq 2$, we have by (12) that $\dot{n}^d_{12} = 0$; by symmetry, $\dot{n}^d_{21} = 0$. That is, when 1 and 2 are dancing together, since there is no creation of differential knowledge between the two, it holds at any time t that

$$n^d_{12}(t) = n^d_{21}(t) = n^d(0) \tag{72}$$

Thus, using (7), the number of ideas created by the partnership $\{1, 2\}$ from time 0 to time t is given by

$$\Delta n_{12}^c(t) = \int_0^t \beta \left[n_{12}^c(s) \cdot n^d(0)^2\right]^{\frac{1}{3}} ds \tag{73}$$

and hence

$$n_{12}^c(t) = n_{21}^c(t) = n^c(0) + \Delta n_{12}^c(t) \tag{74}$$

Concerning the shadow partnership $\{1, 3\}$, since dancers 1 and 3 have not met prior to time t, the number of ideas they have in common is the number they had in common initially:

$$n_{13}^c(t) = n^c(0) \tag{75}$$

Furthermore, setting $i = 1$ and $j = 3$ in (12) where $\delta_{12} = 1$ and $\delta_{1k} = 0$ for all $k \neq 2$ under the partnerships P_1, we have

$$\dot{n}_{13}^d = a_{12} = \beta \left[n_{12}^c \cdot (n_{12}^d)^2\right]^{\frac{1}{3}}$$

Thus, using (72) and (74), and recalling (73),

$$\begin{aligned} n_{13}^d(t) &= n_{13}^d(0) + \int_0^t \beta \cdot \left[n_{12}^c(s) \cdot n_{12}^d(t)^2\right]^{\frac{1}{3}} \\ &= n^d(0) + \int_0^t \beta \cdot \left[n_{12}^c(s) \cdot n^d(0)^2\right]^{\frac{1}{3}} \\ &= n^d(0) + \Delta n_{12}^c(t) \end{aligned}$$

That is, the number of ideas that dancer 1 knows but dancer 3 does not know at time t is the number of ideas that dancer 1 knows but dancer 3 does not know initially, plus the number of ideas that dancers 1 and 2 created during their partnership from time 0 to time t. Similarly,

$$n_{31}^d(t) = n^d(0) + \Delta n_{34}^c(t) = n^d(0) + \Delta n_{12}^c(t) = n_{13}^d(t) \tag{76}$$

where $\Delta n_{34}^c(t) = \Delta n_{12}^c(t)$ by symmetry.

Now, at time $t = t'$, setting $i = 1$ and $j = 2$ in (71), and using (72) and (74), we have

$$g\left(m_{12}^d(t')\right) = \frac{\beta \cdot \left\{[n^c(0) + \Delta n_{12}^c(t')] \cdot n^d(0)^2\right\}^{\frac{1}{3}}}{n_1(t')}$$

Likewise, setting $i = 1$ and $j = 3$ in (71), and using (75) and (76),

$$g(m_{13}^d(t')) = \frac{\beta \cdot \left\{n^c(0) \cdot \left[n^d(0) + \Delta n_{12}^c(t')\right]^2\right\}^{\frac{1}{3}}}{n_1(t')}$$

Hence, the equality (70) holds if and only if

$$[n^c(0) + \Delta n^c_{12}(t')] \cdot n^d(0)^2 = n^c(0) \cdot \left[n^d(0) + \Delta n^c_{12}(t')\right]^2$$

which can be rewritten as follows:

$$\Delta n^c_{12}(t') \cdot n^d(0)^2 \left\{1 - \frac{2n^c(0)}{n^d(0)} - \frac{n^c(0)}{n^d(0)} \frac{\Delta n^c_{12}(t')}{n^d(0)}\right\} = 0$$

Since $\Delta n^c_{12}(t') \cdot n^d(0)^2 > 0$ for any $t' > 0$, this means that the terms inside the braces be zero, or

$$\frac{\Delta n^c_{12}(t')}{n^d(0)} = \frac{n^d(0)}{n^c(0)} - 2 \tag{77}$$

On the other hand, using (75) and (76),

$$m^d_{13}(t') \equiv \frac{n^d_{ij}(t')}{n^{ij}(t')} = \frac{n^d(0) + \Delta n^c_{12}(t')}{n^c(0) + 2\left[n^d(0) + \Delta n^c_{12}(t')\right]}$$

which can be restated as

$$n^c(0) + 2\left[n^d(0) + \Delta n^c_{12}(t')\right] = \frac{n^d(0)}{m^d_{13}(t')} + \frac{\Delta n^c_{12}(t')}{m^d_{13}(t')}$$

or

$$\frac{n^c(0)}{n^d(0)} + 2 - \frac{1}{m^d_{13}(t')} = \frac{\Delta n^c_{12}(t')}{n^d(0)} \left(\frac{1}{m^d_{13}(t')} - 2\right)$$

Substituting (77) into the right hand side of this equation and arranging terms yields

$$\begin{aligned} m^d_{13}(t') &= \frac{\frac{n^d(0)}{n^c(0)} - 1}{\frac{n^c(0)}{n^d(0)} + \frac{2n^d(0)}{n^c(0)} - 2} \\ &= \frac{1 - \frac{n^c(0)}{n^d(0)}}{\left(\frac{n^c(0)}{n^d(0)}\right)^2 + 2 - \frac{2n^c(0)}{n^d(0)}} \end{aligned} \tag{78}$$

Setting $t = 0$ and using $m^d_{13}(0) = m^d(0)$, we have

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

or

$$\frac{n^c(0)}{n^d(0)} = \frac{1}{m^d(0)} - 2 \tag{79}$$

Substituting (79) into (78) yields

$$\begin{aligned} m^d_{13}(t') &= \frac{3 - \frac{1}{m^d(0)}}{\left(\frac{1}{m^d(0)} - 2\right)^2 + 2 - 2\left(\frac{1}{m^d(0)} - 2\right)} \\ &= \frac{3 - \frac{1}{m^d(0)}}{2 - \left(\frac{1}{m^d(0)} - 2\right)\left(4 - \frac{1}{m^d(0)}\right)} \end{aligned}$$

Deducting 2/5 from the both sides of this equation, we can obtain

$$m_{13}^d(t') - \frac{2}{5} = \frac{\left(m^d(0) - \frac{2}{5}\right)\left(1 - m^d(0)\right)}{m^d(0)^2\left[2 - \left(\frac{1}{m^d(0)} - 2\right)\left(4 - \frac{1}{m^d(0)}\right)\right]}$$

which leads to equation (69) in Lemma 1. Hence, relation (68) holds if and only if equation (69) holds. We can readily see that the right hand side of equation (69) increases continuously from 2/5 to 1/2 as $m^d(0)$ moves from 2/5 to 1/2. On the other hand, using (47), we can see that the value of $m_{13}^d(t)$ increases continuously from $m^d(0)$ to 1/2 as t increases from 0 to ∞. Therefore, for any $m^d(0) \in (2/5, 1/2)$, relation (69) defines uniquely the time $t' > 0$ at which the equality (68) holds. Finally, since $m_{12}^d(t)$ decreases and $m_{13}^d(t)$ increases with time t and since the function $g(m)$ is single-peaked at $m = 2/5$, we have relation (50).■

10 Appendix 2: Analysis of Phase 2

Lemma 2: *At the time t' that is defined by (50) each agent switches to a unique, new partner with whom they have not worked previously.*

Proof of Lemma 2: To examine precisely what form of dance begins at time t', first notice by symmetry that the following relationship holds at time t':

$$m_{ij}^d(t') = m_{12}^d\left[m^d(0)\right] \quad \text{for all } \{i,j\} \in P_1 \tag{80}$$

$$m_{ij}^d(t') = m_{13}^d\left[m^d(0)\right] \quad \text{for all } \{i,j\} \notin P_1 \tag{81}$$

Furthermore, assuming that $2/5 < m^d(0) < \hat{m}$, it holds that

$$g\left(m_{12}^d\left[m^d(0)\right]\right) = g\left(m_{13}^d\left[m^d(0)\right]\right) > \alpha \tag{82}$$

and hence dancer i chooses at time t' a strategy under the following condition:

$$\delta_{ii} = 0 \ \text{ and } \ \sum_{j \neq i} \delta_{ij} = 1 \tag{83}$$

Using (80) to (83), at time t' we have

$$\frac{\dot{y}_i}{y_i} = g\left(m_{12}^d\left[m_0^d\right]\right)$$

which is independent of δ_{ij}. Thus, the equilibrium selection at time t' requires the evaluation of the derivative of percent income growth.

Hereafter we focus on person 1, and simplify the notation as follows:

$$m_{12}^d \left[m^d(0)\right] \equiv \bar{m}_{12}^d, \quad m_{13}^d \left[m^d(0)\right] \equiv \bar{m}_{13}^d. \tag{84}$$

Then, the time derivative of the percent income growth rate at time t' (divided by a positive constant) is given as follows (see Technical Appendix c for proof):

$$\begin{aligned} \frac{d\left(\dot{y}_1/y_1\right)/dt}{g(\bar{m}_{12}^d)} &= \left(1-\bar{m}_{12}^d\right)\cdot g'(\bar{m}_{12}^d)\cdot\delta_{12}\cdot\left\{1-2\bar{m}_{12}^d-\left(1-\bar{m}_{12}^d\right)\cdot\delta_{12}\right\} \\ &\quad +\left(1-\bar{m}_{13}^d\right)\cdot g'(\bar{m}_{13}^d)\cdot\left\{\left(1-2\bar{m}_{13}^d\right)\cdot(1-\delta_{12})-\left(1-\bar{m}_{13}^d\right)\sum_{j\geq 3}\delta_{1j}^2\right\} \end{aligned} \tag{85}$$

Since $g'(\bar{m}_{13}^d) < 0$, when we fix δ_{12} at any value between 0 and 1 (in particular, at its optimal value), the maximization of (85) leads to the following problem:

$$\max \sum_{j\geq 3}\delta_{1j}^2 \ \text{ subject to } \sum_{j\geq 3}\delta_{1j} = 1-\delta_{12}$$

which requires choice of a single $k \geq 3$ and setting

$$\delta_{1k} = 1-\delta_{12}, \text{ whereas } \delta_{1j} = 0 \text{ for } j \neq k, \text{ where } k, j \geq 3. \tag{86}$$

Thus, we can rewrite (85) as follows:

$$\begin{aligned} \frac{d\left(\dot{y}_1/y_1\right)/dt}{g(\bar{m}_{12}^d)} &= \left(1-\bar{m}_{12}^d\right)\cdot g'(\bar{m}_{12}^d)\cdot\delta_{12}\cdot\left\{1-2\bar{m}_{12}^d-\left(1-\bar{m}_{12}^d\right)\cdot\delta_{12}\right\} \\ &\quad +\left(1-\bar{m}_{13}^d\right)\cdot g'(\bar{m}_{13}^d)\cdot(1-\delta_{12})\cdot\left\{1-2\bar{m}_{13}^d-\left(1-\bar{m}_{13}^d\right)\cdot(1-\delta_{12})\right\} \end{aligned} \tag{87}$$

Given that $g'(\bar{m}_{12}^d) > 0$ and $g'(\bar{m}_{13}^d) < 0$, we can readily see that the right hand side of (87) is negative when $\delta_{12} = 1$, whereas it is positive when $\delta_{12} = 0$. Indeed, we can show that it achieves its maximum at $\delta_{12} = 0$ (see Technical Appendix c for proof). Thus, setting $\delta_{12} = 0$ in (86), the second order condition for equilibrium selection requires that person 1 chooses at time t' any new partner $k \neq 2$, and set $\delta_{1k} = 1$. Likewise, each dancer switches to a new partner at time t'.■

Lemma 3: *In the context of Lemma 1, suppose that the initial partnerships* $\{1,2\}$ *and* $\{3,4\}$ *switch to the new partnerships* $\{1,3\}$ *and* $\{2,4\}$ *at time* t' *where*

$$g\left(m_{12}^d(t')\right) = g(m_{13}^d(t'))$$

and

$$m_{12}^d(t') = m_{34}^d(t') < m^B < m_{13}^d(t') = m_{14}^d(t')$$

Assuming that the new partnerships are kept after time t', let t'' be the time at which $m^d_{12}(t)$ and $m^d_{13}(t)$ become the same:

$$m^d_{12}(t'') = m^d_{13}(t'')$$

Then, it holds for $t > t'$,

$$g\left(m^d_{12}(t)\right) \lesseqgtr g(m^d_{13}(t)) \quad \text{as} \quad t \lesseqgtr t'' \tag{88}$$

and

$$g\left(m^d_{13}(t)\right) > g(m^d_{14}(t)) \quad \text{for} \quad t' < t \leq t'' \tag{89}$$

Hence, indeed, the new partnerships $\{1,3\}$ and $\{2,4\}$ formed at time t' can be sustained until time t''. This second switching-time, t'', is uniquely determined by solving the following relationship:

$$\Delta n^c_{13}(t', t'') = n^d_{13}(t') - n^d_{12}(t') \equiv \Delta n^c_{12}(t') \tag{90}$$

where $\Delta n^c_{13}(t', t)$ is the number of ideas created under the partnership $\{1,3\}$ from time t' to time $t \geq t'$, which is given by (92). The position where $m^d_{12}(t)$ meets $m^d_{13}(t)$ is given by

$$m^d_{12}(t'') = m^d_{13}(t'') = \frac{2}{5} - \frac{m^d(0) - \frac{2}{5}}{5m^d(0) - 1} \tag{91}$$

Proof of Lemma 3: To examine how long the new partnerships will be maintained, let us focus on the partnership $\{1,3\}$. Let $\Delta n^c_{13}(t', t)$ be the number of ideas created under the partnership $\{1,3\}$ from time t' to time $t \geq t'$, which is given by

$$\Delta n^c_{13}(t', t) = \int_{t'}^{t} \beta \left[n^c_{13}(s) \cdot n^d_{13}(s)^2\right]^{1/3} ds \tag{92}$$

$$\begin{aligned} n^c_{13}(t) \quad &= n^c_{31}(t) \quad = n^c_{13}(t') + \Delta n^c_{13}(t', t) \\ &\qquad\qquad\quad = n^c(0) + \Delta n^c_{13}(t', t) \end{aligned} \tag{93}$$

$$n^d_{13}(t) = n^d_{31}(t) = n^d_{13}(t') = n^d(0) + \Delta n^c_{12}(t') \tag{94}$$

Substituting (93) and (94) into (92) and solving the integral equation yields

$$\Delta n^c_{13}(t', t) = \left[n^c(0)^{2/3} + \frac{2}{3}\beta n^d_{13}(t')^{2/3}(t - t')\right]^{3/2} - n^c(0) \tag{95}$$

Using (93) and (94),

$$\begin{aligned} n^{13}(t) &= n^c_{13}(t) + 2n^d_{13}(t) \\ &= n^c(0) + 2n^d_{13}(t') + \Delta n^c_{13}(t', t) \end{aligned}$$

So,

$$m_{13}^d(t) = \frac{n_{13}^d(t')}{n^c(0) + 2n_{13}^d(t') + \Delta n_{13}^c(t', t)} \tag{96}$$

At any time $t > t'$, dancer 1 could switch from the present partner 3 to the previous partner 2 who has been dancing with person 4 since time t'. Then,

$$n_{12}^c(t) = n_{12}^c(t') \tag{97}$$

$$n_{12}^d(t) = n_{12}^d(t') + \Delta n_{13}^c(t', t) \tag{98}$$

$$n_{21}^d(t) = n_{12}^d(t) \quad \text{by symmetry}$$

so

$$n^{12}(t) = n_{12}^c(t') + 2\left[n_{12}^d(t') + \Delta n_{13}^c(t', t)\right]$$

which leads to

$$m_{12}^d(t) = \frac{n_{12}^d(t') + \Delta n_{13}^c(t', t)}{n_{12}^c(t') + 2\left[n_{12}^d(t') + \Delta n_{13}^c(t', t)\right]} \tag{99}$$

Likewise, at any time $t > t'$, dancer 1 could switch from the present partner 3 to person 4 (instead of person 2). Then, since persons 1 and 4 never danced together previously,

$$n_{14}^c(t) = n^c(0) \tag{100}$$

$$\begin{aligned} n_{14}^d(t) &= n^d(0) + \Delta n_{12}^c(t') + \Delta n_{13}(t', t) \\ &= n_{13}^d(t') + \Delta n_{13}(t', t) \end{aligned} \tag{101}$$

$$n_{41}^d(t) = n_{14}^d(t) \quad \text{by symmetry}$$

so

$$n^{14}(t) = n^c(0) + 2\left[n_{13}^d(t') + \Delta n_{13}(t', t)\right]$$

and hence

$$m_{14}^d(t) = \frac{n_{13}^d(t') + \Delta n_{13}(t', t)}{n^c(0) + 2\left[n_{13}^d(t') + \Delta n_{13}(t', t)\right]} \tag{102}$$

By differentiating (96), (99) and (102), we have

$$\dot{m}_{12}^d(t) = \frac{n_{12}^c(t')}{\left(n_{12}^c(t') + 2\left[n_{12}^d(t') + \Delta n_{13}^c(t', t)\right]\right)^2} \cdot \Delta \dot{n}_{13}^c(t', t) > 0 \tag{103}$$

$$\dot{m}_{13}^d(t) = -\frac{n_{13}^d(t')}{\left[n^c(0) + 2n_{13}^d(t') + \Delta n_{13}^c(t', t)\right]^2} \cdot \Delta \dot{n}_{13}^c(t', t) < 0 \tag{104}$$

$$\dot{m}_{14}^d(t) = \frac{n^c(0)}{\left(n^c(0) + 2\left[n_{13}^d(t') + \Delta n_{13}(t', t)\right]\right)^2} \cdot \Delta \dot{n}_{13}^c(t', t) > 0 \tag{105}$$

where, from (95),

$$\Delta \dot{n}^c_{13}(t',t) = \beta \left[n^c(0)^{2/3} + \frac{2}{3}\beta n^d_{13}(t')^{2/3}(t-t') \right]^{1/2} n^d_{13}(t')^{2/3} > 0$$

Hence, under the partnerships $\{1,3\}$ and $\{1,4\}$, both $m^d_{12}(t)$ and $m^d_{14}(t)$ increase while $m^d_{13}(t)$ decreases with time t. Let t'' be the time at which $m^d_{12}(t)$ becomes equal to $m^d_{13}(t)$:

$$m^d_{12}(t'') = m^d_{13}(t'') \tag{106}$$

Then, since $m^d_{12}(t') < m^B < m^d_{13}(t') = m^d_{14}(t')$ and since $g(m)$ is single-peaked at m^B, it holds that

$$\min\left\{ g(m^d_{12}(t)), g(m^d_{13}(t)) \right\} > g(m^d_{12}(t')) > g(m^d_{14}(t)) \quad \text{for } t' < t \le t'' \tag{107}$$

Hence, in the time interval $(t', t'']$, dancer 1 never desires to switch partners from person 3 to person 4. It is, however, not *a priori* obvious which of $g(m^d_{12}(t))$ and $g(m^d_{13}(t))$ is greater in the interval (t', t''). However, given that function $g(m)$ is steeper on the right of bliss point m^B in Figure 4, we can guess that the value of $g(m^d_{13}(t))$ is increasing faster (initially, at least) than the value of $g(m^d_{12}(t))$, and hence the partnership $\{1,3\}$ will continue until $m^d_{13}(t)$ crosses the bliss point and then becomes the same as $m^d_{12}(t)$. Indeed, we prove this next.

In the context of Lemma 1, suppose that the initial partnerships $\{1,2\}$ and $\{3,4\}$ switch to the new partnerships $\{1,3\}$ and $\{2,4\}$ at time t', when condition (70) holds. And assume that the new partnerships are kept after time t'. Then, since each of $\{1,2\}$ and $\{1,3\}$ is pairwise symmetric, applying (71) in the present context, for $t \ge t'$ we have

$$g\left(m^d_{13}(t)\right) \gtreqless g\left(m^d_{12}(t)\right) \quad \text{as} \quad n^c_{13}(t) n^d_{13}(t)^2 \gtreqless n^c_{12}(t) n^d_{12}(t)^2 \tag{108}$$

Using (93), (94), (97) and (98), it follows that

$$\begin{aligned}
& n^c_{13}(t) n^d_{13}(t)^2 - n^c_{12}(t) n^d_{12}(t)^2 \\
= \; & \left[n^c_{13}(t') + \Delta n^c_{13}(t',t) \right] n^d_{13}(t')^2 - n^c_{12}(t') \left[n^d_{12}(t') + \Delta n^c_{13}(t',t) \right]^2 \\
= \; & \Delta n^c_{13}(t',t) n^d_{13}(t')^2 \left\{ 1 - \frac{2 n^c_{12}(t') n^d_{12}(t')}{n^d_{13}(t')^2} - \frac{n^c_{12}(t')}{n^d_{13}(t')^2} \cdot \Delta n^c_{13}(t',t) \right\}
\end{aligned}$$

Hence, for $t \ge t'$, it holds that

$$g\left(m^d_{13}(t)\right) \gtreqless g\left(m^d_{12}(t)\right) \quad \text{as} \quad \Delta n^c_{13}(t',t) \lesseqgtr \frac{n^d_{13}(t')^2}{n^c_{12}(t')} - 2 n^d_{12}(t') \tag{109}$$

To simplify the expression above, we derive a useful equality. By definition, the following identity holds at any time t:

$$n_1(t) = n_{12}^c(t) + n_{12}^d(t) = n_{13}^c(t) + n_{13}^d(t) \tag{110}$$

Setting $t = t'$ in (110), using the equality in (108) to substitute for $n_{13}^c(t')$, and solving for $n_{12}^c(t')$ yields

$$n_{12}^c(t') = \frac{n_{13}^d(t')^2}{n_{12}^d(t') + n_{13}^d(t')} \tag{111}$$

Similarly,

$$n_{13}^c(t') = \frac{n_{12}^d(t')^2}{n_{12}^d(t') + n_{13}^d(t')} \tag{112}$$

Substituting (111) into the last term in (109) gives

$$\begin{aligned}
\frac{n_{13}^d(t')^2}{n_{12}^c(t')} - 2n_{12}^d(t') &= n_{13}^d(t') - n_{12}^d(t') \\
&= \left(n^d(0) + \Delta n_{12}^c(t')\right) - n^d(0) \\
&= \Delta n_{12}^c(t')
\end{aligned}$$

using (94) at $t = t'$. Thus, we can conclude that

$$g\left(m_{13}^d(t)\right) \gtreqless g\left(m_{12}^d(t)\right) \quad \text{as} \quad \Delta n_{13}^c(t', t) \lesseqgtr \Delta n_{12}^c(t') \tag{113}$$

Let t'' be the time such that

$$\Delta n_{13}^c(t', t'') = \Delta n_{12}^c(t') \tag{114}$$

Since equation (95) implies that $\Delta n_{13}^c(t', t') = 0$ and since $\Delta n_{13}^c(t', t)$ increases continuously to ∞ as t tends to ∞, equation (114) uniquely defines $t'' > t'$. Hence, we can conclude from (113) that for $t \geq t'$,

$$g\left(m_{13}^d(t)\right) \gtreqless g\left(m_{12}^d(t)\right) \quad \text{as} \quad t \lesseqgtr t'' \tag{115}$$

Substituting (111) into (99) and setting $t = t''$ and using $\Delta n_{13}^c(t', t'') = \Delta n_{12}^c(t') = n_{13}^d(t') - n_{12}^d(t')$ yields

$$m_{12}^d(t'') = \frac{n_{13}^d(t')}{\frac{n_{13}^d(t')^2}{n_{12}^d(t') + n_{13}^d(t')} + 2n_{13}^d(t')}$$

Likewise, using (93) to set $n_{13}^c(t') = n^c(0)$ in (96) and using (112) also yields

$$m_{13}^d(t'') = \frac{n_{13}^d(t')}{\frac{n_{13}^d(t')^2}{n_{12}^d(t') + n_{13}^d(t')} + 2n_{13}^d(t')}$$

Hence, rewriting the expression above, and using the relations $n_{13}^d(t') = n^d(0) + \Delta n_{12}^c(t')$ and $n_{12}^d(t') = n^d(0)$, we have

$$\begin{aligned}
m_{12}^d(t'') = m_{13}^d(t'') &= \frac{1}{\frac{n_{13}^d(t')}{n_{12}^d(t')+n_{13}^d(t')}+2} \\
&= \frac{1}{\frac{n^d(0)+\Delta n_{12}^c(t')}{2n^d(0)+\Delta n_{12}^c(t')}+2} \\
&= \frac{1}{\frac{1+\frac{\Delta n_{12}^c(t')}{n^d(0)}}{2+\frac{\Delta n_{12}^c(t')}{n^d(0)}}+2} \\
&= \frac{1}{3-\frac{n^c(0)}{n^d(0)}} \quad \text{(using (77))} \\
&= \frac{1}{5-\frac{1}{m^d(0)}} \quad \text{(using (79))}
\end{aligned}$$

which can be rewritten as (91). Thus,

$$m_{12}^d(t'') = m_{13}^d(t'') < m^B = 2/5 \tag{116}$$

This gives the alternative definition of time t'', which has been introduced in (55). Thus, (113) and (114) imply (88) and (90) in Lemma 3. Finally, relation (89) follows immediately from (107).■

11 Appendix 3: Efficiency of the Equilibrium Path

Here we discuss efficiency in the context of an intertemporal utilitarian social welfare function. We consider the following planner's problem, where the planner chooses $\{\delta_{ij}(\cdot)\}_{i,j=1}^N$ in order to:

$$\max W = \sum_{i=1}^N U_i(0) = \sum_{i=1}^N \int_0^\infty e^{-\gamma t} \cdot y_i(t)dt = \sum_{i=1}^N \int_0^\infty e^{-\gamma t} \cdot n_i(t)dt$$

subject to

$$\begin{aligned}
\dot{n}_i &= \sum_{j=1}^N \delta_{ij} \cdot a_{ij} = n_i \left\{ \delta_{ii} \cdot \alpha + \sum_{j\neq i} \delta_{ij} \cdot G(m_{ij}^d, m_{ji}^d) \right\} \\
&\quad \text{and} \\
\dot{m}_{ij}^d &= \alpha \cdot \left(1-m_{ij}^d\right) \cdot \left[\delta_{ii} \cdot \left(1-m_{ji}^d\right) - \delta_{jj} \cdot m_{ij}^d\right] - \delta_{ij} \cdot m_{ij}^d \cdot \left(1-m_{ji}^d\right) \cdot G(m_{ij}^d, m_{ji}^d) \\
&\quad + \left(1-m_{ij}^d\right) \cdot \left(1-m_{ji}^d\right) \cdot \sum_{k\neq i,j} \delta_{ik} \cdot G(m_{ik}^d, m_{ki}^d) \\
&\quad - \left(1-m_{ij}^d\right) \cdot m_{ij}^d \cdot \sum_{k\neq i,j} \delta_{jk} \cdot G(m_{jk}^d, m_{kj}^d)
\end{aligned}$$

given $n_i\left(0\right) > 0$ and $m^d_{ij}(0) > 0$, for $i, j = 1, ..., N$. We must also account for the obvious constraints:

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

We assume that the discount rate is sufficiently large, $\gamma > g(m^B)$, in order to ensure that the objective is finite. Optimality requires that at each moment of time, the following Hamiltonian is maximized by choosing $\{\delta_{ij}\}_{i,j=1}^{N}$ and taking into account the obvious constraints:

$$H = \sum_{i=1}^{N} e^{-\gamma t} \cdot n_i(t) + \sum_{i=1}^{N} \lambda_i \cdot \dot{n}_i + \sum_{i=1}^{N} \sum_{j \neq i} \chi_{ij} \cdot \dot{m}^d_{ij}$$

where the multipliers follow the dynamics:

$$\begin{aligned}
\dot{\lambda}_i &= -\frac{\partial H}{\partial n_i} \text{ for } i = 1, ..., N \\
\dot{\chi}_{ij} &= -\frac{\partial H}{\partial m^d_{ij}} \text{ for } i, j = 1, ..., N, \ i \neq j
\end{aligned}$$

and satisfy the following transversality condition:[27]

$$\lim_{t \to \infty} H(t) = 0$$

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

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

Recall that the myopic equilibrium path for case (i) when $m^J < m^d_{ij}(0)$ is:

$$\begin{aligned}
\delta_{ij}(t) &= \frac{1}{N-1} \text{ for } t < t^B \text{ for } i, j = 1, ..., N, \ i \neq j \qquad (117) \\
\delta_{ij}(t) &= \frac{1}{N^B - 1} \text{ for } t > t^B \text{ when } i \text{ and } j \text{ belong to the same group}
\end{aligned}$$

[27] This transversality condition comes from Léonard and Van Long (1992), Theorem 9.6.1, p. 299.

where t^B is the first time t such that $m(t) = m^d(t) = m^B$, the bliss point m^B is given by (64) and the group size N^B is given by (67).

Under these initial conditions, it can be verified that if N is sufficiently large, then there exists a set of multipliers such that the myopic equilibrium path detailed in (117) for case (i) satisfies the necessary conditions for optimality.

When $m^d(t) < m^B$ and therefore $t < t^B$, if each person works with every other person with equal intensity, then knowledge productivity is higher than working in isolation and m^d_{ij} moves almost as fast to the right as working in isolation. The intuition for this result follows from a combination of two reasons. Productivity is higher when working with others as opposed to working alone on this part of the path. When N is sufficiently large, working with others is very close to working in isolation when the accumulation of differential knowledge is considered, so cooperation with others will be better on net. Once the bliss point is attained, the system reaches the highest productivity possible, and remains there.

This intuition indicates that, when $m^d(t) < m^B$, working with a smaller group than the other $N - 1$ dancers generates movement to the right that is slower than working with everyone but oneself. So coalitions cannot block this path. Furthermore, once the bliss point is achieved, this is the highest productivity possible, so coalitions cannot block this part of the path either. Thus, the path chosen by myopic agents, that coincides with the utilitarian welfare optimal path, is in the core with rational expectations.

12 Technical Appendix

12.1 Appendix a

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

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

where the function G is defined by (17).
Proof: From (4) and (5),

$$n_i = n^{ij} - n_{ji}^d = n^{ij} \cdot \left(1 - \frac{n_{ji}^d}{n^{ij}}\right) = n^{ij} \cdot \left(1 - m_{ji}^d\right)$$

thus

$$\frac{n_i}{n^{ij}} = 1 - m_{ji}^d \tag{118}$$

Now, from (7),

$$\begin{aligned}
\frac{a_{ij}}{n_i} &= \frac{n^{ij}}{n_i} \cdot \frac{a_{ij}}{n^{ij}} \\
&= \frac{1}{1 - m_{ji}^d} \cdot \beta \left[m_{ij}^c \cdot \cdot m_{ij}^d \cdot m_{ji}^d\right]^{\frac{1}{3}} \\
&= G(m_{ij}^d, m_{ji}^d)
\end{aligned}$$

which leads to (16).
Theorem A2: Knowledge dynamics evolve according to the system:

$$\begin{aligned}
\dot{m}_{ij}^d &= \alpha \cdot \left(1 - m_{ij}^d\right) \cdot \left[\delta_{ii}\left(1 - m_{ji}^d\right) - \delta_{jj} m_{ij}^d\right] - \delta_{ij} \cdot m_{ij}^d \cdot \beta \cdot \left(1 - m_{ji}^d\right) \cdot G(m_{ij}^d, m_{ji}^d) \\
&\quad + \left(1 - m_{ij}^d\right) \cdot \left(1 - m_{ji}^d\right) \cdot \sum_{k \neq i,j} \delta_{ik} \cdot G(m_{ik}^d, m_{ki}^d) \\
&\quad - \left(1 - m_{ij}^d\right) \cdot m_{ij}^d \cdot \sum_{k \neq i,j} \delta_{jk} \cdot G(m_{jk}^d, m_{kj}^d)
\end{aligned}$$

for $i, j = 1, 2, \cdots, N$.

Proof: By definition,

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

where, using (13) and (118), we have

$$\begin{aligned}
\frac{\dot{n}_{ij}^{d}}{n^{ij}} &= \frac{\sum\limits_{k \neq j} \delta_{ik} \cdot a_{ik}}{n^{ij}} \\
&= \frac{\delta_{ii} \cdot \alpha \cdot n_i}{n^{ij}} + \sum_{k \neq i,j} \delta_{ik} \cdot \frac{a_{ik}}{n^{ij}} \\
&= \frac{\delta_{ii} \cdot \alpha \cdot n_i}{n^{ij}} + \sum_{k \neq i,j} \delta_{ik} \cdot \frac{n_i}{n^{ij}} \cdot \frac{n^{ik}}{n_i} \cdot \frac{a_{ik}}{n^{ik}} \\
&= \frac{n_i}{n^{ij}} \cdot \left\{\delta_{ii} \cdot \alpha + \sum_{k \neq i,j} \delta_{ik} \cdot \frac{n^{ik}}{n_i} \cdot \frac{a_{ik}}{n^{ik}}\right\} \\
&= \left(1 - m_{ji}^{d}\right) \cdot \left\{\delta_{ii} \cdot \alpha + \sum_{k \neq i,j} \delta_{ik} \cdot \frac{1}{1 - m_{ki}^{d}} \cdot \beta \left[m_{ik}^{c} \cdot m_{ik}^{d} \cdot m_{ki}^{d}\right]^{\frac{1}{3}}\right\} \\
&= \left(1 - m_{ji}^{d}\right) \cdot \left\{\delta_{ii} \cdot \alpha + \sum_{k \neq i,j} \delta_{ik} \cdot \frac{\beta\left[\left(1 - m_{ik}^{d} - m_{ki}^{d}\right) \cdot m_{ik}^{d} \cdot m_{ki}^{d}\right]^{\frac{1}{3}}}{1 - m_{ki}^{d}}\right\} \\
&= \left(1 - m_{ji}^{d}\right) \cdot \left\{\delta_{ii} \cdot \alpha + \sum_{k \neq i,j} \delta_{ik} \cdot G(m_{ik}^{d}, m_{ki}^{d})\right\}
\end{aligned}$$

Similarly,

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

while using (12) yields

$$\begin{aligned}\frac{\dot{n}^c_{ij}}{n^{ij}} &= \delta_{ij}\cdot\frac{a_{ij}}{n^{ij}}\\ &= \delta_{ij}\cdot\beta\left[m^c_{ij}\cdot m^d_{ij}\cdot m^d_{ji}\right]^{\frac{1}{3}}\\ &= \delta_{ij}\cdot\beta\left[\left(1-m^d_{ij}-m^d_{ji}\right)\cdot m^d_{ij}\cdot m^d_{ji}\right]^{\frac{1}{3}}\\ &= \delta_{ij}\cdot(1-m^d_{ij})\cdot G(m^d_{ij},m^d_{ji})\end{aligned}$$

Thus,

$$\begin{aligned}\dot{m}^d_{ij} &= \left(1-m^d_{ij}\right)\cdot\left(1-m^d_{ji}\right)\left\{\delta_{ii}\cdot\alpha+\sum_{k\neq i,j}\delta_{ik}\cdot G(m^d_{ik},m^d_{ki})\right\}\\ &\quad -\delta_{ij}\cdot m^d_{ij}\cdot\left(1-m^d_{ij}\right)\cdot G(m^d_{ij},m^d_{ji})\\ &\quad -m^d_{ij}\cdot\left(1-m^d_{ij}\right)\cdot\left\{\delta_{jj}\cdot\alpha+\sum_{k\neq i,j}\delta_{jk}\cdot G(m^d_{jk},m^d_{kj})\right\}\\ &= \alpha\cdot\left(1-m^d_{ij}\right)\cdot\left[\delta_{ii}\cdot\left(1-m^d_{ji}\right)-\delta_{jj}\cdot m^d_{ij}\right]-\delta_{ij}\cdot m^d_{ij}\cdot\left(1-m^d_{ij}\right)\cdot G(m^d_{ij},m^d_{ji})\\ &\quad +\left(1-m^d_{ij}\right)\cdot\left(1-m^d_{ji}\right)\cdot\sum_{k\neq i,j}\delta_{ik}\cdot G(m^d_{ik},m^d_{ki})\\ &\quad -\left(1-m^d_{ij}\right)\cdot m^d_{ij}\cdot\sum_{k\neq i,j}\delta_{jk}\cdot G(m^d_{jk},m^d_{kj})\end{aligned}$$

12.2 Appendix b

Lemma A1: When $m^d_{ij}=m^d$ for all $i\neq j$ and $g(m^d)>\alpha$, the time derivative of $\dot{y}_i/y_i$ is given by

$$\frac{d\left(\dot{y}_i/y_i\right)}{dt}=\left(1-m^d\right)\cdot g(m^d)\cdot g'(m^d)\cdot\left[1-2m^d-(1-m^d)\cdot\sum_{j\neq i}\delta^2_{ij}\right]$$

Proof Lemma A1: From (18) we have

$$\frac{\dot{y}_i}{y_i}=\delta_{ii}\cdot\alpha+\sum_{j\neq i}\delta_{ij}\cdot G\left(m^d_{ij},m^d_{ji}\right)$$

When $m^d_{ij}=m^d$ for all $i\neq j$ and $g(m^d)>\alpha$, person i never wishes to dance alone, and hence

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

Thus, we can set

$$\frac{\dot{y}_i}{y_i}=\sum_{j\neq i}\delta_{ij}\cdot G\left(m^d_{ij},m^d_{ji}\right)$$

Let

$$\frac{\partial G\left(m_{ij}^d, m_{ji}^d\right)}{\partial m_{ij}^d} \equiv G_1\left(m_{ij}^d, m_{ji}^d\right)$$
$$\frac{\partial G\left(m_{ij}^d, m_{ji}^d\right)}{\partial m_{ji}^d} \equiv G_2\left(m_{ij}^d, m_{ji}^d\right)$$

Then, taking the time derivative of $(\dot{y}_i/y_i)$ when $m_{ij}^d = m^d$ for all $i \neq j$, we claim that

$$\frac{d\left(\dot{y}_i/y_i\right)}{dt} = \sum_{j\neq i} \delta_{ij} \cdot \left[G_1\left(m^d, m^d\right) \cdot \dot{m}_{ij}^d + G_2\left(m^d, m^d\right) \cdot \dot{m}_{ji}^d\right] \tag{119}$$

To see this, we use the definition of the derivative (an explanation follows immediately):

$$\begin{aligned}
\frac{d\left(\dot{y}_i/y_i\right)}{dt} &= \lim_{t'\to t}\left[\frac{\sum_{j\neq i}\delta_{ij}(t')\cdot G\left(m_{ij}^d(t'), m_{ji}^d(t')\right) - \sum_{j\neq i}\delta_{ij}(t)\cdot G\left(m_{ij}^d(t), m_{ji}^d(t)\right)}{t'-t}\right] \\
&= \lim_{t'\to t}\left[\frac{\sum_{j\neq i}\delta_{ij}(t)\cdot \{G\left(m_{ij}^d(t'), m_{ji}^d(t')\right) - G\left(m_{ij}^d(t), m_{ji}^d(t)\right)\}}{t'-t}\right] \\
&\quad + \lim_{t'\to t}\left[\frac{\sum_{j\neq i}\{\delta_{ij}(t') - \delta_{ij}(t)\}\cdot G\left(m_{ij}^d(t), m_{ji}^d(t)\right)}{t'-t}\right] \\
&\quad + \lim_{t'\to t}\left[\frac{\sum_{j\neq i}\{\delta_{ij}(t') - \delta_{ij}(t)\}\cdot \{G\left(m_{ij}^d(t'), m_{ji}^d(t')\right) - G\left(m_{ij}^d(t), m_{ji}^d(t)\right)\}}{t'-t}\right] \\
&= \sum_{j\neq i} \delta_{ij} \cdot \left[G_1\left(m^d, m^d\right) \cdot \dot{m}_{ij}^d + G_2\left(m^d, m^d\right) \cdot \dot{m}_{ji}^d\right]
\end{aligned}$$

The first term of the expression becomes the right hand side of equation (119). The second term is zero because $m_{ij}^d = m^d$ for all $i \neq j$, $g(m^d) > \alpha$ and g is continuous so $\sum_{j\neq i}\delta_{ij}(t') = \sum_{j\neq i}\delta_{ij}(t) = 1$ for t' close to t. Noting that G, $m_{ij}^d(t)$, and $m_{ji}^d(t)$ are all differentiable, the last term is $\lim_{t'\to t}\sum_{j\neq i}\{\delta_{ij}(t') - \delta_{ij}(t)\}\cdot\{G_1\left(m_{ij}^d(t), m_{ji}^d(t)\right)\cdot\dot{m}_{ij}^d(t) + G_2\left(m_{ij}^d(t), m_{ji}^d(t)\right)\cdot\dot{m}_{ji}^d(t)\}$. Again, symmetry and the fact that no agent is dancing alone imply that it is zero.

When $m_{ij}^d = m^d$ for all $i \neq j$, $\dot{m}_{ij}^d$ is given by (35). Furthermore, on any feasible path, $\delta_{ij} = \delta_{ji}$. Thus, using (35),

$$\dot{m}_{ij}^d = \dot{m}_{ji}^d = \left(1 - m^d\right)\cdot g(m^d)\left[1 - 2m^d - \left(1 - m^d\right)\cdot\delta_{ij}\right] \tag{120}$$

Hence,

$$\frac{d\left(\dot{y}_i/y_i\right)}{dt} = \sum_{j\neq i} \delta_{ij} \cdot \left[G_1\left(m^d, m^d\right) + G_2\left(m^d, m^d\right)\right] \cdot \dot{m}_{ij}^d$$

Straightforward calculations yield that

$$G_1\left(m^d, m^d\right) = \frac{\frac{1}{3}\cdot\beta\cdot\left[\left(1 - 2m^d\right)\cdot\left(m^d\right)^2\right]^{-\frac{2}{3}}}{1 - m^d}\cdot\left(1 - 3m^d\right)\cdot m^d$$

$$
\begin{aligned}
G_2\left(m^d, m^d\right) &= \frac{\beta \cdot \left[\left(1-2m^d\right)\cdot\left(m^d\right)^2\right]^{\frac{1}{3}}}{\left(1-m^d\right)^2} \\
&\quad + \frac{\frac{1}{3}\cdot\beta\cdot\left[\left(1-2m^d\right)\cdot\left(m^d\right)^2\right]^{-\frac{2}{3}}}{1-m^d}\cdot\left(1-3m^d\right)\cdot m^d
\end{aligned}
$$

Adding together and arranging terms give

$$
\begin{aligned}
& G_1\left(m^d, m^d\right) + G_2\left(m^d, m^d\right) \\
=\ & \frac{\beta}{3}\cdot\left[\frac{\left(1-2m^d\right)\cdot\left(m^d\right)^2}{\left(1-m^d\right)^3}\right]^{-\frac{2}{3}}\cdot\frac{m^d\cdot\left(2-5m^d\right)}{\left(1-m^d\right)^4} \\
=\ & g'(m^d)
\end{aligned}
$$

which follows from (24). Thus,

$$
\begin{aligned}
\frac{d\left(\dot{y}_i/y_i\right)}{dt} &= g'(m^d)\cdot\sum_{j\neq i}\delta_{ij}\cdot\dot{m}^d_{ij} \\
&= g'(m^d)\cdot\sum_{j\neq i}\delta_{ij}\cdot\left(1-m^d\right)\cdot g(m^d)\cdot\left[1-2m^d-\left(1-m^d\right)\cdot\delta_{ij}\right] \\
&= \left(1-m^d\right)\cdot g(m^d)\cdot g'(m^d)\cdot\left[\left(1-2m^d\right)\cdot\sum_{j\neq i}\delta_{ij}-\left(1-m^d\right)\cdot\sum_{j\neq i}\delta^2_{ij}\right] \\
&= \left(1-m^d\right)\cdot g(m^d)\cdot g'(m^d)\cdot\left[1-2m^d-\left(1-m^d\right)\cdot\sum_{j\neq i}\delta^2_{ij}\right]
\end{aligned}
$$

as was to be shown.■

12.3 Appendix c

Lemma A2: In the context of Lemma 1, the time derivative of the percent income growth rate at time t' (divided by a positive constant $g(\bar{m}^d_{12})$) is given by

$$
\begin{aligned}
\frac{d\left(\dot{y}_1/y_1\right)/dt}{g(\bar{m}^d_{12})} &= \left(1-\bar{m}^d_{12}\right)\cdot g'(\bar{m}^d_{12})\cdot\delta_{12}\cdot\left\{1-2\bar{m}^d_{12}-\left(1-\bar{m}^d_{12}\right)\cdot\delta_{12}\right\} \qquad (121) \\
&\quad + \left(1-\bar{m}^d_{13}\right)\cdot g'(\bar{m}^d_{13})\cdot\left\{\left(1-2\bar{m}^d_{13}\right)\cdot\left(1-\delta_{12}\right)-\left(1-\bar{m}^d_{13}\right)\sum_{j\geq 3}\delta^2_{1j}\right\}
\end{aligned}
$$

which achieves its maximum value when $\delta_{1k} = 1$ for any single $k \neq 2$ whereas $\delta_{1j} = 0$ for all $j \neq k$.

Proof of Lemma A2: Using (80) to (84) and since $\delta_{ij} = \delta_{ji}$ for any feasible path, from (31) we have at time t'

$$\begin{aligned} \dot{m}_{ij}^d &= \dot{m}_{ji}^d = \left(1-\bar{m}_{12}^d\right)\cdot g(\bar{m}_{12}^d)\cdot\left\{1-2\bar{m}_{12}^d-\left(1-\bar{m}_{12}^d\right)\cdot\delta_{ij}\right\} \text{ for } \{i,j\}\in P_1 \\ \dot{m}_{ij}^d &= \dot{m}_{ji}^d = \left(1-\bar{m}_{13}^d\right)\cdot g(\bar{m}_{13}^d)\cdot\left\{1-2\bar{m}_{13}^d-\left(1-\bar{m}_{13}^d\right)\cdot\delta_{ij}\right\} \text{ for } \{i,j\}\notin P_1 \end{aligned} \tag{122}$$

Next, focusing on person $i = 1$, similar to the derivation of (119) in the proof of Lemma A1 in Technical Appendix b, at time t' we can obtain that

$$\frac{d\left(\dot{y}_1/y_1\right)}{dt} = \sum_{j\neq 1}\delta_{1j}\cdot\left[G_1(m_{1j}^d, m_{1j}^d)\cdot\dot{m}_{1j}^d + G_2(m_{1j}^d, m_{1j}^d)\cdot\dot{m}_{j1}^d\right]$$

where the functions G_1 and G_2 have been defined in Technical Appendix b. Again, in the same manner as in Technical Appendix b, we can show that

$$\begin{aligned} G_1(m_{12}^d, m_{12}^d) + G_2(m_{12}^d, m_{12}^d) &= g'(\bar{m}_{12}^d) \\ G_1(m_{1j}^d, m_{1j}^d) + G_2(m_{1j}^d, m_{1j}^d) &= g'(\bar{m}_{13}^d) \text{ for } j\geq 3 \end{aligned}$$

Thus, since $\dot{m}_{1j}^d = \dot{m}_{j1}^d$, it follows that

$$\frac{d\left(\dot{y}_1/y_1\right)}{dt} = g'(m_{12}^d)\cdot\delta_{12}\cdot\dot{m}_{12}^d + g'(m_{13}^d)\cdot\sum_{j\geq 3}\delta_{1j}\cdot\dot{m}_{1j}^d$$

Substituting (122) into the right hand side above, and using the relation that $\sum_{j\geq 3}\delta_{1j} = 1-\delta_{12}$, we have equation (121) or (85) in the text. As explained in the text, since condition (86) must hold at the equilibrium selection, we can rewrite the equation (85) as equation (87).

Next, dividing both sides of equation (87) by a positive constant,

$$\left(1-\bar{m}_{12}^d\right)\cdot g'(\bar{m}_{12}^d)\cdot\left(1-2\bar{m}_{12}^d\right)$$

we have that

$$\begin{aligned} V(\delta_{12}) &\equiv \frac{d\left(\dot{y}_1/y_1\right)/dt}{g(\bar{m}_{12}^d)\cdot\left(1-\bar{m}_{12}^d\right)\cdot g'(\bar{m}_{12}^d)\cdot\left(1-2\bar{m}_{12}^d\right)} \\ &= \delta_{12}\cdot\left\{1-\frac{1-\bar{m}_{12}^d}{1-2\bar{m}_{12}^d}\cdot\delta_{12}\right\} \\ &\quad +C\cdot(1-\delta_{12})\cdot\left\{1-\frac{1-\bar{m}_{13}^d}{1-2\bar{m}_{13}^d}\cdot(1-\delta_{12})\right\} \end{aligned}$$

where

$$\begin{aligned} C &\equiv \frac{\left(1-\bar{m}_{13}^d\right)\cdot g'(\bar{m}_{13}^d)\cdot\left(1-2\bar{m}_{13}^d\right)}{\left(1-\bar{m}_{12}^d\right)\cdot g'(\bar{m}_{12}^d)\cdot\left(1-2\bar{m}_{12}^d\right)} \\ &= \frac{\left(1-\bar{m}_{13}^d\right)\cdot g'(\bar{m}_{13}^d)\cdot\left(1-2\bar{m}_{13}^d\right)}{\left(1-\bar{m}_{12}^d\right)\cdot g'(\bar{m}_{12}^d)\cdot\left(1-2\bar{m}_{12}^d\right)}\cdot\frac{g(\bar{m}_{12}^d)}{g(\bar{m}_{13}^d)} \end{aligned}$$

since $g(\bar{m}^d_{12}) = g(\bar{m}^d_{13})$. Using (22) and (24) yields

$$C = -\frac{\bar{m}^d_{13} - \frac{2}{5}}{\frac{2}{5} - \bar{m}^d_{12}} \cdot \frac{\bar{m}^d_{12}}{\bar{m}^d_{13}}$$

Thus,

$$\begin{aligned} V(\delta_{12}) &= \delta_{12} \cdot \left\{1 - \frac{1 - \bar{m}^d_{12}}{1 - 2\bar{m}^d_{12}} \cdot \delta_{12}\right\} \\ &\quad -\frac{\bar{m}^d_{13} - \frac{2}{5}}{\frac{2}{5} - \bar{m}^d_{12}} \cdot \frac{\bar{m}^d_{12}}{\bar{m}^d_{13}} \cdot (1 - \delta_{12}) \cdot \left\{1 - \frac{1 - \bar{m}^d_{13}}{1 - 2\bar{m}^d_{13}} \cdot (1 - \delta_{12})\right\} \end{aligned} \tag{123}$$

Since we have

$$0 < \bar{m}^d_{12} < \frac{2}{5} < \bar{m}^d_{13} < \frac{1}{2}$$

it follows that

$$V(0) = \frac{\bar{m}^d_{13} - \frac{2}{5}}{\frac{2}{5} - \bar{m}^d_{12}} \cdot \frac{\bar{m}^d_{12}}{\bar{m}^d_{13}} \cdot \frac{\bar{m}^d_{13}}{1 - 2\bar{m}^d_{13}} > 0 \tag{124}$$

$$V(1) = \frac{-\bar{m}^d_{12}}{1 - 2\bar{m}^d_{12}} < 0 \tag{125}$$

Next, taking the derivative of V at $\delta_{12} = 0$ yields

$$V'(0) = 1 - D$$

where

$$D \equiv \frac{\bar{m}^d_{13} - \frac{2}{5}}{\frac{2}{5} - \bar{m}^d_{12}} \cdot \frac{\bar{m}^d_{12}}{\bar{m}^d_{13}} \cdot \frac{1}{1 - 2\bar{m}^d_{13}} \tag{126}$$

To investigate whether D exceeds 1 or not, denoting $m^d(0) \equiv m^d_0$, we have from (33) that

$$\bar{m}^d_{13} = \frac{2}{5} + y(m^d_0) \tag{127}$$

where

$$y(m^d_0) \equiv \frac{\left(m^d_0 - \frac{2}{5}\right) \cdot \left(1 - m^d_0\right)}{2(m^d_0)^2 - (1 - 2m^d_0) \cdot (4m^d_0 - 1)} = \frac{\left(m^d_0 - \frac{2}{5}\right) \cdot \left(1 - m^d_0\right)}{10(m^d_0)^2 - 6m^d_0 + 1}$$

We can readily see that

$$0 < y(m^d_0) < \frac{1}{10} \text{ for } \frac{2}{5} < m^d_0 < \frac{1}{2}$$

On the other hand, using the equality $g(\bar{m}^d_{12}) = g(\bar{m}^d_{13})$ and following the steps in the proof of Lemma 1 (with $m^d_{12}(t') = \frac{n^d(0)}{n^c(0) + \Delta n^c_{12}(t') + 2n^d(0)}$), we can obtain that

$$\bar{m}^d_{12} = \frac{2}{5} - x(m^d_0) \tag{128}$$

where

$$x(m_0^d) \equiv \frac{\left(m_0^d - \frac{2}{5}\right) \cdot \left(4m_0^d - 1\right)}{5(m_0^d)^2 - 4m_0^d + 1}$$

We can also show that

$$0 < x(m_0^d) < \frac{2}{5} \text{ for } \frac{2}{5} < m_0^d < \frac{1}{2}$$

Substituting (127) and (128) into (126) gives

$$\begin{aligned} D &= \frac{y(m_0^d)}{x(m_0^d)} \cdot \frac{\frac{2}{5} - x(m_0^d)}{\frac{2}{5} + y(m_0^d)} \cdot \frac{1}{\frac{1}{5} - 2y(m_0^d)} \\ &= \frac{y(m_0^d)}{x(m_0^d)} \cdot \frac{\frac{2}{5} - x(m_0^d)}{1 - 10y(m_0^d)} \cdot \frac{5}{\frac{2}{5} + y(m_0^d)} \end{aligned}$$

Recalling that $\frac{2}{5} < m_0^d < \frac{1}{2}$, let us evaluate each component above. First, since $10(m_0^d)^2 - 6m_0^d + 1 > 0$ for $\frac{2}{5} < m_0^d < \frac{1}{2}$,

$$\begin{aligned} \frac{y(m_0^d)}{x(m_0^d)} &= \frac{5(m_0^d)^2 - 4m_0^d + 1}{10(m_0^d)^2 - 6m_0^d + 1} \cdot \frac{1 - m_0^d}{4m_0^d - 1} \\ &= \frac{1}{2} \cdot \frac{5(m_0^d)^2 - 4m_0^d + 1}{\left[5(m_0^d)^2 - 4m_0^d + 1\right] + m_0^d - \frac{1}{2}} \cdot \frac{1 - m_0^d}{4m_0^d - 1} \\ &> \frac{1}{2} \cdot \frac{1 - m_0^d}{4m_0^d - 1} > \frac{1}{2} \cdot \frac{1 - \frac{1}{2}}{4 \cdot \frac{1}{2} - 1} = \frac{1}{4} \end{aligned}$$

thus

$$\frac{y(m_0^d)}{x(m_0^d)} > \frac{1}{4}$$

Next,

$$\begin{aligned}
\frac{\frac{2}{5}-x(m_0^d)}{1-10y(m_0^d)} &= \frac{\frac{2}{5}-\frac{\left(m_0^d-\frac{2}{5}\right)\cdot\left(4m_0^d-1\right)}{5(m_0^d)^2-4m_0^d+1}}{1-\frac{10\left(m_0^d-\frac{2}{5}\right)\cdot\left(1-m_0^d\right)}{10(m_0^d)^2-6m_0^d+1}} \\
&= \frac{2}{5}\cdot\frac{1-\frac{5\left(m_0^d-\frac{2}{5}\right)\cdot\left(4m_0^d-1\right)}{10(m_0^d)^2-8m_0^d+2}}{1-\frac{10\left(m_0^d-\frac{2}{5}\right)\cdot\left(1-m_0^d\right)}{10(m_0^d)^2-6m_0^d+1}} \\
&= \frac{2}{5}\cdot\left\{1-1+\frac{1-\frac{5\left(m_0^d-\frac{2}{5}\right)\cdot\left(4m_0^d-1\right)}{10(m_0^d)^2-8m_0^d+2}}{1-\frac{10\left(m_0^d-\frac{2}{5}\right)\cdot\left(1-m_0^d\right)}{10(m_0^d)^2-6m_0^d+1}}\right\} \\
&= \frac{2}{5}\cdot\left\{1+\frac{\frac{10\left(m_0^d-\frac{2}{5}\right)\cdot\left(1-m_0^d\right)}{10(m_0^d)^2-6m_0^d+1}-\frac{5\left(m_0^d-\frac{2}{5}\right)\cdot\left(4m_0^d-1\right)}{10(m_0^d)^2-8m_0^d+2}}{1-\frac{10\left(m_0^d-\frac{2}{5}\right)\cdot\left(1-m_0^d\right)}{10(m_0^d)^2-6m_0^d+1}}\right\} \\
&= \frac{2}{5}\cdot\left\{1+\frac{5\left(m_0^d-\frac{2}{5}\right)\cdot\frac{4m_0^d-1}{10(m_0^d)^2-6m_0^d+1}\left[\frac{2(1-m_0^d)}{4m_0^d-1}-\frac{10(m_0^d)^2-6m_0^d+1}{\left[10(m_0^d)^2-6m_0^d+1\right]+(1-2m_0^d)}\right]}{1-\frac{10\left(m_0^d-\frac{2}{5}\right)\cdot\left(1-m_0^d\right)}{10(m_0^d)^2-6m_0^d+1}}\right\}
\end{aligned}$$

Since $\frac{2}{5}<m_0^d<\frac{1}{2}$ and $10(m_0^d)^2-6m_0^d+1>0$ for $\frac{2}{5}<m_0^d<\frac{1}{2}$,

$$\frac{2(1-m_0^d)}{4m_0^d-1}>1$$

$$\frac{10(m_0^d)^2-6m_0^d+1}{\left[10(m_0^d)^2-6m_0^d+1\right]+(1-2m_0^d)}<1$$

and

$$1-\frac{10\left(m_0^d-\frac{2}{5}\right)\cdot\left(1-m_0^d\right)}{10(m_0^d)^2-6m_0^d+1}=\frac{5(2m_0^d-1)^2}{10(m_0^d)^2-6m_0^d+1}>0$$

Thus,

$$\frac{\frac{2}{5}-x(m_0^d)}{1-10y(m_0^d)}>\frac{2}{5}$$

Finally, since $y(m_0^d)<\frac{1}{10}$ for $\frac{2}{5}<m_0^d<\frac{1}{2}$,

$$\frac{5}{\frac{2}{5}+y(m_0^d)}>\frac{5}{\frac{2}{5}+\frac{1}{10}}=10$$

so

$$\frac{5}{\frac{2}{5}+y(m_0^d)}>10$$

Therefore

$$D>\frac{1}{4}\cdot\frac{2}{5}\cdot 10=1$$

Hence

$$V'(0) < 0$$

Since the function V is quadratic in δ_{12}, from the fact that $V(0) > 0$, $V(1) < 0$ and $V'(0) < 0$, we can see that $V(\delta_{12})$ achieves its maximum at $\delta_{12} = 0$. Therefore, we can conclude from (86) that the right hand side of (121) achieves the maximum value when $\delta_{1k} = 1$ for any one $k \neq 2$ and $\delta_{1j} = 0$ for all $j \neq k$, as was to be shown.$\blacksquare$