MERIT-Infonomics Research Memorandum series

Information Technology and the Dynamics of Joint Innovation

Robin Cowan & Nicolas Jonard

2004-014

Information Technology and the Dynamics of Joint Innovation*

Robin Cowan♮, Nicolas Jonard‡

♮MERIT, Maastricht University, P.O. Box 616,
6200 MD Maastricht, The Netherlands.
‡CNRS, CREA, Ecole Polytechnique,
1 Rue Descartes, 75005 Paris, France.

July 1, 2004

Abstract

We develop a model of an innovative industry to examine how information technology, by both enhancing matching efficiency and knowledge sharing, can have an ambiguous effect on the total amount of innovation. We consider a population of firms holding different knowledge expertise, and forming partnerships to conduct joint R&D. We assume that bringing together different expertise has positive value for innovating but also that joint innovation implies a partial convergence of the partners' expertise. We study how the distribution of firms changes and thus how the innovative potential of the economy evolves. We show that as heterogeneity is used as an input by the innovative process, the industry must eventually collapse to a unique expertise, but how fast this takes place depends on the quality of IT. As a result of falling dispersion, a tension arises between static and dynamic efficiency.

JEL Classification Numbers: C78,O33,O38.

1   Introduction

This paper develops a dynamic model of innovation to examine the net effect of two features of the information technology revolution: improvements in IT both make it easier for firms to find better partners for innovation, and make partnerships more efficient in sharing knowledge. These two properties of new information technology can create a tension between static and dynamic efficiency in the innovation system.

It is now a commonplace that innovation is a key driver of economic growth. Further, in terms of economic value the vast majority of innovation consists in recombination of existing ideas.[1] This implies that improved access not only to more knowledge or information but also to a wider variety of knowledge should improve efficiency in R&D (see Weitzman, 1998 or Olsson, 2000). As competitive pressures increase, firms feel the need to improve their innovation levels and consequently are increasingly using knowledge that lies outside their main domains of competence (see Powell et al., 1996 for a discussion). A very common way to acquire this knowledge is to form an alliance with another firm. We have seen evidence over the last two decades of rapidly increasing inter-firm R&D alliances and industry-university cooperation (Hagedoorn, 2002; Poyago-Theotoky et al., 2002), particularly in industries that are technologically complex or which perform considerable R&D (see for example Freeman, 1991; Powell et al., 1996).

Information technology contributes to this process in several ways. Kaufman et al. (2003) propose the following grouping of the effects of IT on the innovation process: distribution of information; collection of information; and interactive communication.[2] As any of these three activities becomes more efficient, the value of jointly conducted R&D will increase. This is one of the conclusions of Gaspar and Glaeser (1998) regarding the effects of IT on urbanization. In their model improvements in IT imply a substitution between face-to-face and electronic communication, but also a general increase in the frequency of communication because the expected returns from new relationships have increased.

Joint, as opposed to individual, innovation will only be valuable if the partners' knowledge stocks complement each other. For an individual agent considering joint R&D then, one issue is finding the right partner, Narula (1999), but finding a partner will typically involve search, which in turn involves having information about other firms. The IT revolution notwithstanding, information remains costly, but

---

[1]This idea is present in the work of Schumpeter (1934). Koestler (1964) provides an interesting discussion of it. For several examples of innovation by recombination see van Alstyne and Brynjolfsson (1997, pp. 22-23).

[2]Similarly Dudley (1999), formalizing the ideas of Innis (1951), and North (1981), argue that output will grow, or equivalently innovation will increase, when the amount of stored knowledge increases; the cost of transmitting knowledge over distance (as a means of coordinating activities) falls; and the cost of decoding information falls. All these effects follow from advances in IT.

1

reductions in the cost of obtaining information will in general improve the global efficiency of the matching of innovation partners. Falling costs of IT imply that the quantity of public information has increased, as firms' activities are broadcast on the World Wide Web, and patent databases are easier to search for example. The information needed to locate the best available partner is easier to find, which should imply that in general the matching between firms will be more efficient.[3] Van Alstyne and Brynjolfsson argue that improvements in IT provide agents with greater control of "interaction and content filtering" and so improvements in IT make it easier for innovating agents to find better partners (1996, p.19).[4] This general trend may be part of the explanation for the rapid increase in inter-continental R&D alliances for European firms in the past decades — their ability to find the right partner has increased, and when technology leaders are abroad, this will imply an increase in international partnerships.[5]

Information technology serves a second role in inter-firm R&D alliances. It is an input into the management of them. Successful R&D partnering involves creating a stable environment between the partners, in which trust can develop. It also involves management of knowledge flows. IT plays a role in both. While face to face contacts are still necessary, IT can be used to enhance the exchange of information, and thereby support the development of the necessary relationships. Improvements to IT mean that responses to queries are faster, more information can be transferred, and so on. The survey results of Kaufman et al. (2004) show "unambiguously that the Internet is a technology which helps to increase the efficiency of communication" in existing relationships (p. 414). We can interpret this as an improved knowledge flow enhancing the ability of firms to learn from each other.

In this paper we develop a model of the joint R&D process within an industry to capture these two effects of improvements in IT. The industry consists of a population of firms which are heterogeneous in the type of knowledge they hold. Firms form alliances and jointly innovate. A key assumption is that the amount of knowledge created increases as partners' knowledge is more diverse, because in this

---

[3]This is an idea that policy makers have accepted, as they are supporting websites specifically designed to help firms find the best partner for innovation. See for example www.kompetenze.de or www.tekes.fi.

[4]Kaufman et al. (2003), in a survey of 200 Austrian firms, find that more than half of firms having existing external alliances have found new partners using the internet (p. 415). But they also find that partners tend to be found in regions where the firms already have alliances.

[5]Curiously, as a proportion of all alliances, international alliances have been decreasing over the past 2 decades. Disaggregating, however, Hagedoorn (2002) shows that this is driven by a steep rise in intra-North American alliances, and a decrease in intra-European alliances. This trend is strong "particularly in high-tech industries such as pharmaceuticals (biotechnology) and information technology", (p. 489), where the US is a technology leader. Part of the interpretation may be that European firms are more successful in finding "better" partners, which implicitly involves finding partners in North America.

case there is more room for recombination.[6] The ability of firms to find the best available partner depends on the quality of IT. Finally, the process of innovation changes a firm's expertise. During the joint innovation process partners learn from each other, and so move closer together in the underlying knowledge space.[7] After innovation, alliances dissolve and the process repeats.

Aggregate knowledge production is determined by IT through two channels: In the short run, the efficiency of pair formation affects immediate production. Over time, though, the extent to which firms learn from each other determines how fast they converge in terms of expertise, which affects future innovation potential. In the limit the industry collapses to a single, common expertise, but how fast this happens depends on the IT regime. Thus an inevitable tension arises between static and dynamic efficiency: high quality ITs get the best out of a heterogenous population of firms during early periods, but they exhaust the innovative potential quickly.

2 The model

In this section we set forth the characteristics of joint innovation and describe the matching technology before providing the first results.

2.1 Joint Innovation

One contentious issue in conceptualizing joint innovation is the value of heterogeneity in the expertise of the two firms. Clearly if the two knowledge types are too disparate communication is impossible and joint innovation fruitless. At the other extreme, if firms have identical expertise, there is little to be gained in terms of the re-combination aspect of innovation, on which this paper focuses (see Grant, 1996). Conceiving of heterogeneity as an input to knowledge production, it seems its marginal productivity should be an inverted U. Indeed, this is what Mowery et al. (1998) find empirically, suggesting a positive optimal distance in terms of the partner's knowledge overlap. Noteboom (1999) makes similar conclusions in terms of cognitive distance for jointly performed tasks. In this paper we assume that our agents are relatively similar to each other, and so focus attention on the positively sloped part of this curve. In other words, we assume that in choosing a partner with

---

[6]Clearly too much diversity would mean that partners could not understand each other (see the discussion in Section 2.1 below); the focus here is on an industry in which a certain degree of commonality of knowledge exists.

[7]This effect is empirically demonstrated by Mowery et al. (1998). See the discussion at the beginning of Section 4.

whom to innovate, diversity is a good thing.[8]

To make this concrete, consider a large population of firms indexed by their knowledge expertise over the possibly unbounded interval $S \subseteq \mathbf{R}$. The expertise of a firm is denoted x and is distributed according to an atomless, absolutely continuous distribution function $F(x)$. The density function of firms' expertise is denoted $f(x)$. Assume the measure of firms is normalized to unity, so that it is possible to refer to the absolutely continuous random variable X with distribution function F. X will be used to compact exposition; it is assumed to have finite mean m and variance v.

When participating in a partnership, each firm invests one unit of human capital whose type is the firm's expertise. The output of the collaboration between two participating firms having expertise x and y is denoted $\gamma(x, y)$. We assume that output is positive unless the two firms' expertise coincide: $\gamma(x, y) \geq 0$ and $\gamma(x, y) = 0$ iff $x = y$. The output function is symmetric: $\gamma(x, y) = \gamma(y, x)$. Finally, output increases with firms' dissimilarity, i.e. $\gamma(x, y)$ increases in $|x - y|$.

2.2   Alliance formation

Out of the set of all possible partnerships, which ones will form? To answer that question a matching mechanism needs to be specified, and initially we consider two extreme cases. The first is a barebones version of decentralized search: firms meet randomly and form pairs. Random assignment has minimal information requirements and we use it to represent the situation wherein the quality of IT is very bad. Firms have no way of learning the expertise of the others, and thus are a priori indifferent between any two partners. The second matching mechanism we consider is a process for which there is a well defined equilibrium concept: stability as defined by the non-existence of blocking pairs. In the context of a single population, this is known as a one-sided or roommate matching problem (Gale and Shapley, 1962). Roommate matching, contrasting with random matching, has high information demands and so can only take place when ITs are extremely effective.

2.2.1   Random matching

Under the random matching regime, in any period $t = 0, 1, 2, \ldots$ firms are matched randomly so a firm can have any partner, and the probability that a firm is matched with a partner of expertise less than y is simply $F(y)$. This, again, results from firms' complete ignorance of the expertise of other firms.

---

[8]van Allstyne and Brynjolfssen (1996, 1997) present models in which agents prefer to match with similar agents. Their concern is with the Balkanization of science, and so implicitly they are considering a much more diverse set of agents than we are here. Peretto and Smulders (2002) assume that technological spillovers are among similar agents, but their concern is with economy-wide spillovers thus also a very diverse population.

2.2.2 Roommate matching

When firms have perfect information they know the type of all potential partners, and thus the value of every potential alliance. As a result a firm can form a complete preference ordering over all other firms as potential partners. This results in a one-sided matching problem in any period. In its original form a roommate matching is a one-to-one mapping μ from a discrete finite population into itself such that there is no self-matching and the composition of the mapping μ with itself is the identity function. Stability of a matching μ is then characterized by the absence of blocking pairs: no two firms would prefer each other to the partners they have under μ. On a continuous distribution the logic is unchanged but the formulation differs slightly. Work on the continuous two-sided matching problem includes Sattinger (1975, 1995) for example, where matching is between workers and machines, and the more recent papers by Burdett and Coles (1997) and Bloch and Ryder (2000) who address the marriage market. Our approach differs from these first in that we treat a single population of firms looking for innovation partners and second in that the characteristics and thus the preferences of our population change over time.

A matching function is a one-to-one, measure-preserving mapping from S onto itself, specifying for each expertise x an assignment $\mu(x)$. First a complete preference ordering held by each type $x \in S$ over all possible types must be specified. Given the nature of the innovative process, the utility obtained by firm x from a match with y is written

$$u_x(y) = \gamma(x, y). \tag{1}$$

Expertise y is preferred to z by an x-type firm iff $\gamma(x, y) > \gamma(x, z)$. Denote $u = \{u_x, x \in S\}$ the profile of these preferences: (S, u) is a continuous roommate matching problem. Given a matching function μ, a blocking pair is a pair (x, y) such that $\gamma(x, y) > \gamma(x, \mu(x))$ and $\gamma(y, x) > \gamma(y, \mu(y))$. A matching μ^* is stable in (S, u) if the set of blocking pairs for μ^* is of measure zero.

Under the assumptions on knowledge production, equilibrium can now be fully characterized.

Proposition 1 *The matching problem (S, u) possesses a unique stable matching function $\mu^*(x) = F^{-1}(1 - F(x))$.*

Proof The matching function μ^* is clearly measure preserving as $\mu^*(S) = S$. To show that μ^* is stable, consider any blocking pair (x, y). If x is such that $x < \mu^*(x)$ then any $y > \mu^*(x)$ is preferred to $\mu^*(x)$ by x, but then necessarily $\mu^*(y) < x$ and thus y prefers his current partner to x. By contrast, if x is such that $x > \mu^*(x)$ any $y < \mu^*(x)$ is preferred to $\mu^*(x)$ by x, but then necessarily again $\mu^*(y) > x$ and thus y prefers his current partner to x. Hence the set of blocking couples is empty: μ^* is a stable matching function.

To show that μ^* is the only stable matching function, consider another stable matching $\mu : S \to S$, $\mu \neq \mu^*$. Define $\kappa \in (0, 1/2)$, $\underline{x} = F^{-1}(\kappa)$ and $\overline{x} = F^{-1}(1 - \kappa)$. Observe that in a stable matching agents $x < \underline{x}$ must be matched with agents $\mu(x) > \overline{x}$, as the measures of these two sets are equal, and any agent in the set matching an agent outside the opposing set is made worse off. Define a real-valued sequence $\{k_0, \ldots, k_n\}$ with $\kappa = k_0 < k_1 < k_2 < \cdots < k_n = 1/2$ and let $x_i = F^{-1}(k_i)$ and $x^i = F^{-1}(1 - k_i)$. Define $I(a, b)$ as the open, closed, or semi-closed interval of real values between a and b. In a stable matching it must be the case that $\mu(I(x_0, x_1)) = I(x^1, x^0)$. Indeed if some positive measure of agents in $I(x_0, x_1)$ are matched outside $I(x^1, x^0)$, then the same measure of agents in $I(x^1, x^0)$ are matched outside $I(x_0, x_1)$. Both prefer each other to their matches outside the two intervals, so these agents would block μ. Iterating this argument gives that in a stable matching $\mu(I(x_i, x_{i+1})) = I(x^{i+1}, x^i)$ for all $i = 0, \ldots, n - 1$. Now simply remark that $x^i = \mu^*(x_i)$. Then the previous condition is written

$$\mu(I(x_i, x_{i+1})) = I(\mu^*(x_i), \mu^*(x_{i+1})),$$

for all $i = 0, \ldots, n - 1$. Letting $n \to +\infty$ such that $\max_i |k_i - k_{i+1}| \to 0$ gives $\mu(x_i) = \mu^*(x_i)$ as defined in the proposition for the sub-population $\underline{x} < x < \overline{x}$. Now letting $\kappa \to 0$ completes the proof. $\square$

The intuition can be seen by considering the problem faced by the firm whose expertise lies at the furthest left end of the support. That firm strictly prefers the most distant firm (at the opposite furthest right end) to any other potential partner given the preferences we have specified. But the firm at the furthest right end faces the same problem. Thus the two firms whose types are the most distant must be matched, as they could block any matching not involving them. These two firms are effectively removed from consideration of the rest of the population, and the argument repeats over and over.

As the distribution function is monotonically increasing it is invertible and the partner of an x-type firm at equilibrium is

$$\mu^*(x) = F^{-1}(1 - F(x)). \tag{2}$$

It will be useful to compute the derivative of μ^*, which is equal to $(\mu^*(x))' = -f(x)/f(\mu^*(x))$. So $\mu^*(x)$ is a strictly decreasing function of x.

3 Knowledge creation

In this section we enquire whether one regime is more effective than the other in creating knowledge. To make things tractable the production function will be

specified as $\gamma(x, y) = (x - y)^2$. The aggregate output associated with the roommate matching μ^* is then written

$$\chi(\mu^*) = \int_S (x - \mu^*(x))^2 f(x) \, dx. \tag{3}$$

It is the sum, over all population types, of the output produced by the firms distributed according to $f(x)$. By contrast under the random assignment μ' total expected output is written

$$\chi(\mu') = \int_S \int_S \{x - y\}^2 f(y) f(x) \, dy \, dx. \tag{4}$$

For an x-type firm the conditional expected output under random matching is $\int_S \{x - y\}^2 f(v) \, dy$, and this is aggregated over the possible values for x, yielding total output $\chi(\mu')$. Then the following result obtains.

Proposition 2 *Total knowledge creation is larger under roommate matching than under random matching for any distribution of firms' expertise $F(x)$.*

Proof. Begin by noting that because $\mu^*(x)$ is a measure-preserving, strictly decreasing continuous function of its argument, $\mu^*(X)$ is also a random variable.

Remark now that since X and $\mu^*(X)$ describe the same population, the distribution function of $\mu^*(X)$ is $1 - F(\mu^*(x)) = F(x)$. As a result $\mathbf{E}[\mu^*(X)] = m$ and $\mathbf{V}[\mu^*(X)] = v$. Formula (3) can be expressed in a more synthetic manner as $\chi(\mu^*) = \mathbf{E}[(X - \mu^*(X))]^2$, which can then be expanded into

$$
\begin{aligned}
\chi(\mu^*) &= \mathbf{E}[X^2] + \mathbf{E}[\mu^*(X)^2] - 2 \int_S x \mu^*(x) f(x) \, dx \\
&= 2(\sigma^2 + m^2) - 2\mathbf{E}[X \mu^*(X)] \\
&= 2\{\sigma^2 - \mathrm{cov}(X, \mu^*(X))\}.
\end{aligned}
$$

Similarly, writing (4) as $\chi(\mu') = \mathbf{E}[(X - Y)^2]$, with X and Y i.i.d. with distribution function F yields $\chi(\mu') = 2\sigma^2$, and thus

$$\chi(\mu^*) - \chi(\mu') = -2\mathrm{cov}(X, \mu^*(X)) > 0$$

as $\mu^*(X)$ is a strictly decreasing function of X $((\mu^*(x))' < 0)$. $\square$

With a given initial distribution the roommate matching is more efficient than random assignment. Put another way, on the basis of this first result, high quality ITs are preferable as they generate more total knowledge than does pure random matching. However it is not guaranteed that the greater efficiency of roommate matching is preserved once the second effect of IT is introduced.

4 Dynamics of knowledge types

As discussed above, one role of ITs is to provide information about the activities of firms in the economy. In the previous section we explored the way this affects knowledge production: better ITs imply better information, which implies a more efficient matching of firms, and finally more knowledge creation. The second aspect of IT introduced in the initial discussion plays a different role. When IT is good, partners can share more information more effectively, and so will learn more extensively from each. They rapidly come to know the same things, and thus the loci of their expertise rapidly converge. By contrast, when IT is poor, learning is more difficult and there is much slower convergence between the partners' expertise. Generally speaking joint innovation tends to imply some form of convergence. Indeed, Mowery et al. (1998) find that "technological overlap between joint venture partners after alliance formation is greater than their pre-alliance overlap (p. 517)". (See also Dyer and Nobeoka, 2000.) IT quality simply scales the magnitude of this effect, which is captured in the model by a linear absorption rule. If an x-firm and a y-firm are matched the post-matching knowledge type of the first firm becomes $x(1-\alpha)+\alpha y$ and similarly the second firm is $y(1-\alpha)+x\alpha$. Parameter $\alpha \in (0,1/2)$ measures the quality of IT: the rate at which partners' expertise converges increases in α.

To explore the dynamics of partnerships, consider again the polar cases corresponding to high and low quality IT. The low quality IT context has absorption characterized by α_l. Starting from an initial distribution of firms' expertise $X \equiv X_0$, the next period distribution is described by $X_1 = (1-\alpha_l)X_0 + \alpha_l Y_0$, where Y_0 is identically independently distributed with X_0. By contrast, the high quality IT is a regime in which absorption is characterized by α_h. One period transforms the initial distribution of firms' expertise X_0 into $X_1 = (1-\alpha_h)X_0 + \alpha_h \mu^*(X_0)$.

Under both regimes heterogeneity declines from one period to the next: variance falls while the initial mean expertise is preserved. However heterogeneity declines at different rates. This is stated more precisely in the following proposition.

Proposition 3 *Roommate and random matching are mean preserving and variance reducing transformations. Under both regimes the probability density of knowledge types converges to a Dirac function $\delta(m)$. For any initial distribution of firms' expertise dispersion is smaller in any period under roommate than under random matching.*

Proof. Under the random technology the post matching distribution is described by $X_1 = (1-\alpha_l)X_0 + \alpha_l Y_0$. As X_0 and Y_0 are identically independently distributed $\mathbf{E}[X_1] = m$ and $\mathbf{V}[X_1] = v\{(1-\alpha_l)^2 + \alpha_l^2\}$. After one period the mean is preserved and the variance is reduced by a factor $(1-\alpha_l)^2 + \alpha_l^2 \in (1/2,1)$. Thus after t

periods $\mathbf{V}\left(X_t\right) = v\{(1 - \alpha_l)^2 + \alpha_l^2\}^t$, which gets arbitrarily close to 0 as t grows large. Eventually the distribution concentrates its mass on a unique expertise: the initial average expertise m.

Move now to high quality IT. Linearity of the expectation operator yields $\mathbf{E}\left[X_1\right] = (1 - \alpha_h)m + \alpha_h \mathbf{E}\left[\mu^*\left(X_0\right)\right]$, which from the previous proposition we know to be equal to m. Thus market matching also preserves the mean and generates a variance

$$\mathbf{V}\left(X_1\right) = v\{(1 - \alpha_h)^2 + \alpha_h^2\} + 2\alpha_h\left(1 - \alpha_h\right)\operatorname{cov}\left(X_0, \mu^*(X_0)\right).$$

As $(\mu^*(x))' < 0$ the covariance of X_0 and $\mu^*(X_0)$ is strictly negative, and thus $\mathbf{V}\left(X_1\right) < v\{(1 - \alpha_h)^2 + \alpha_h^2\} < v\{(1 - \alpha_l)^2 + \alpha_l^2\}$ because $(1 - \alpha)^2 + \alpha^2$ is decreasing in α. More generally the variance is described by the recursive equation

$$\mathbf{V}\left(X_{t+1}\right) = \{(1 - \alpha_h)^2 + \alpha_h^2\}\mathbf{V}\left(X_t\right) + 2\alpha_h\left(1 - \alpha_h\right)\operatorname{cov}\left(X_t, \mu^*(X_t)\right).$$

Thus $\mathbf{V}\left(X_{t+1}\right) < \{(1 - \alpha_h)^2 + \alpha_h^2\}\mathbf{V}\left(X_t\right) < \mathbf{V}\left(X_t\right)$, which gets arbitrarily close to 0 as t grows large. Eventually the distribution collapses to a Dirac $\delta\left(m\right)$. Invoking a recurrence argument, suppose now $\mathbf{V}\left(X_t\right) < v\{(1 - \alpha_l)^2 + \alpha_l^2\}^t$. Then

$$
\begin{aligned}
\mathbf{V}\left(X_{t+1}\right) &= \{(1 - \alpha_h)^2 + \alpha_h^2\}\mathbf{V}\left(X_t\right) + 2\alpha_h\left(1 - \alpha_h\right)\operatorname{cov}\left(X_t, \mu^*(X_t)\right) \\
&< \{(1 - \alpha_h)^2 + \alpha_h^2\}\mathbf{V}\left(X_t\right) \\
&< \{(1 - \alpha_h)^2 + \alpha_h^2\}v\{(1 - \alpha_l)^2 + \alpha_l^2\}^t. \\
&< v\{(1 - \alpha_l)^2 + \alpha_l^2\}^{t+1}.
\end{aligned}
$$

Thus in any period dispersion is always smaller under high quality IT than under low quality IT. $\square$

After a round of roommate matching the population distribution has a lower variance than after random matching, implying that typically production in subsequent rounds will also be lower. This suggests a tension between static and dynamic efficiency: early aggregate efficiency is larger, but production possibilities are exhausted more quickly under roommate matching than under random assignment.

5   Efficiency properties of normal distributions

To generate concrete results on knowledge production, we now assume that the initial distribution of expertise is normal $\mathcal{N}\left(0, 1\right)$. The assumption of zero mean is without loss of generality as only symmetry around the mean matters. Similarly setting $\sigma = 1$ has no consequences as the variance is a multiplicative factor in all the calculations. Normal distribution also have interesting convolution properties. Last the assumption of normality is fairly natural as the knowledge dynamics are in

essence a repeated averaging of random variables, so by the central limit theorem a normal distribution should be a good approximation of the true distribution after some time, regardless of the initial distribution. In the following subsection we again examine the two extreme cases of high and low quality ITs and identify the tension between static and dynamic efficiency, following which a more general case is treated.

5.1 Static versus dynamic efficiency

First consider the case of random alliance formation μ'. As X_0 and Y_0 are identically independently $\mathcal{N}(0, 1)$ distributed, it follows that $(1 - \alpha_l)X_0$ and $\alpha_l Y_0$ are independently distributed as $\mathcal{N}(0, (1 - \alpha_l)^2)$ and $\mathcal{N}(0, \alpha_l^2)$. The probability density of $X_1 = (1 - \alpha_l)X_0 + \alpha_l Y_0$ obtains as the convolution of these two densities, which again is normal with variance $(1 - \alpha_l)^2 + \alpha_l^2$. The convolution iterates so that under random matching $X_t \sim \mathcal{N}(0, \{(1 - \alpha_l)^2 + \alpha_l^2\}^t)$. Making use of the results in the previous section relating aggregate output and variance, we get

$$\chi'_t = 2\{(1 - \alpha_l)^2 + \alpha_l^2\}^t. \tag{5}$$

Assuming a discount factor of δ, total knowledge produced over the lifetime of the economy is simply the discounted sum of the χ'_t, which is written

$$\chi' = \sum_{t \geq 0} \delta^t \chi'_t = \frac{2}{1 - \delta\{(1 - \alpha_l)^2 + \alpha_l^2\}}. \tag{6}$$

Consider now roommate matching. As we consider a symmetric distribution with mean zero we have $\mu^*(X) = -X$. The post matching distribution is $X_1 = (1 - \alpha_h)X - \alpha_h X = (1 - 2\alpha_h)X_0$. As X_0 is normally distributed, so is X_1, according to $\mathcal{N}(0, 1 - 2\alpha_h)$. Iterating this result yields $X_t \sim \mathcal{N}(0, (1 - 2\alpha_h)^t)$. Using the fact that $\mathrm{cov}(X, -X) = -\mathbf{V}[X]$, aggregate output is $\chi_0^* = 2\{1 - \mathrm{cov}(X, -X)\} = 4$ and

$$\chi_t^* = 4(1 - 2\alpha_h)^t \tag{7}$$

Again assuming discount factor δ total discounted knowledge is written as

$$\chi^* = \sum_{t \geq 0} \delta^t \chi_t^* = \frac{4}{1 - \delta(1 - 2\alpha_h)}. \tag{8}$$

Knowledge production under both matching regimes decreases over time. We first ask whether there is a time at which production levels are equal under the two regimes. Solving for t^* such that $4(1 - 2\alpha_h)^{t^*} = 2\{(1 - \alpha_l)^2 + \alpha_l^2\}^{t^*}$ yields

$$t^* = \frac{\ln 2}{\ln\left\{\frac{(1 - \alpha_l)^2 + \alpha_l^2}{1 - 2\alpha_h}\right\}}. \tag{9}$$

10

From t^* on, a regime of low quality ITs generates more output than a high quality one. Solving for the critical δ such that $\chi^* = \chi'$ yields

$$\bar{\delta} = \frac{1}{1 + 2\alpha_h - 4\alpha_l + 4\alpha_l^2}. \tag{10}$$

Differentiating with respect to α_l yields $\partial\bar{\delta}/\partial\alpha_l > 0$ and $\partial\bar{\delta}/\partial\alpha_h < 0$, i.e. better absorption in the low quality IT case means a higher critical discount factor, while higher absorption in the high quality IT regime means a lower critical value of the discount factor. The following proposition readily obtains.

Proposition 4 *For a normal distribution of firms' expertise and linear absorption characterized by α_l and α_h, $0 < \alpha_l < \alpha_h < 1/2$, there exists a critical discount factor $\bar{\delta}$ given by Equation 10 such that when $\delta < \bar{\delta}$ (discounting is strong) high quality ITs are desirable as the roommate matching outperforms random assignment, and when $\delta > \bar{\delta}$ (discounting is low) low quality ITs are preferable.*

The tension between static and dynamic efficiency is real in the case of normally distributed knowledge types. Equilibrium matching is very effective in producing knowledge in the early periods because high quality ITs permit firms to find the best available partner. However, diversity is rapidly exhausted and innovation rates decline. By contrast, low quality ITs prevent firms from finding optimal partners, reducing immediate innovation but at the same time restricting the convergence in knowledge types, which means that innovation is more persistent.

5.2 Heterogeneous populations

Above we discussed two effects of improving the quality of information technologies. Better IT implies a better matching among innovation partners, and it improves inter-firm learning. In this section we present a simple way of integrating these two effects into the model. The model already includes the parameter α, which permits a simple scaling between high learning α_h and low learning α_l, which captures the second effect. To capture the first effect we must describe a way to scale between random matching and roommate matching. A simple way to do this is to assume that some measure of the population, β has access to perfect IT, and so is matched within itself by the roommate algorithm, whereas the remainder of the population, of measure $1 - \beta$, has no access to this form of IT, and so is matched randomly. We assume for tractability that the two sub-populations have the same distribution over knowledge types as the entire population. Improvements in IT are captured by an increase in the measure of firms who have access to good information about potential partners, and who also provide that information about themselves.

11

First take α and β as independent parameters. Equations 5 and 7 give us the value of knowledge produced in each of the two sub-populations in a single period. The knowledge production of the entire economy is simply the weighted sum

$$\chi_t = \beta 4(1 - 2\alpha_h)^t + (1 - \beta)\, 2\{(1 - \alpha_l)^2 + \alpha_l^2\}^t \tag{11}$$

in each round. The present value of knowledge produced as a function of the parameters α and β, with a discount factor δ is therefore

$$\chi(\alpha, \beta) = \sum_{t \geq 0} \delta^t \chi_t = \beta \frac{4}{1 - \delta(1 - 2\alpha_h)} + (1 - \beta)\, \frac{2}{1 - \delta\{(1 - \alpha_l)^2 + \alpha_l^2\}}. \tag{12}$$

Proposition 5 *In general, both high quality IT or low quality IT can be optimal from the point of view of knowledge creation, depending on the learning possibilities of the firms in the two regimes.*

To demonstrate the proposition, we show an example in which Equation 12 gives rises to the two situations.
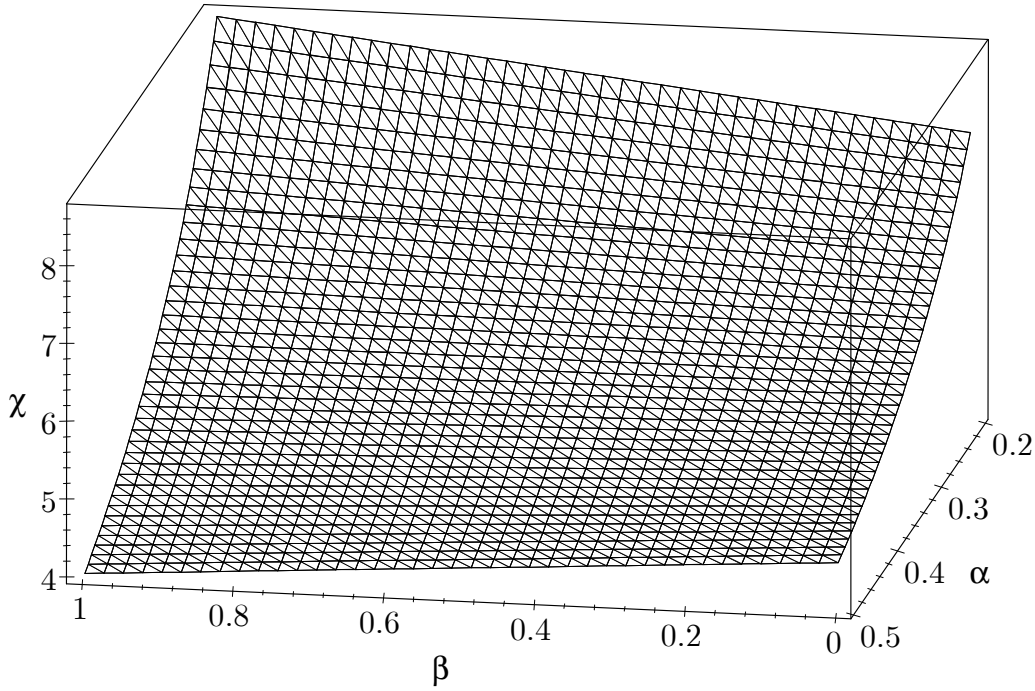

Figure 1: Discounted knowledge production.

Figure 1 shows the relationship between discounted knowledge and the two parameters α and β representing the two effects of IT. It is obtained for relatively

weak discounting ($\delta = 0.9$), with the relation $\alpha_l = \alpha_h/2$ to characterize different learning abilities for the two subpopulations. For any value of α, there is clearly a linear effect of β on knowledge levels. Thus for a given learning capacity, maximal knowledge is produced at extreme values of firms' ability to find good partners. However the slope changes sign with values of α_l and α_h, and so for weak learning, roommate matching is optimal, whereas for strong learning, the random match dominates. Looking along the other axis, for a given value of β, increasing α always decreases knowledge production. That is, if the effectiveness of matching is unchanged, increasing learning increases the rate of convergence in knowledge types, and so reduces the amount of knowledge produced each period.

So far, this discussion ignores the joint aspect of these effects, and has assumed that α and β could change independently. But improvements in IT will in general both increase the ability of firms to find optimal partners, and increase the extent to which firms learn from each other when they collaborate. That is, changes in the quality of IT imply positively correlated movements in α and β. In terms in Figure 1 this implies movement, roughly speaking, in parallel with the diagonal running from $(0,0)$ to $(1/2, 1)$ in the (α, β)-plane. This can result in a variety of shapes for the effect of changing IT quality on knowledge production.

Proposition 6 *An intermediate quality of IT can be optimal from the point of view of knowledge creation.*

Again, to demonstrate the proposition we show the relationship for three values of δ, holding other parameters constant.

In Figure 2 we postulate a linear relationship $\beta = 2\alpha_h$ that captures the correlation described above. We maintain the relationship between α_l and α_h as $\alpha_l = \alpha_h/2$. So in general an increase in IT implies more efficient average search and more efficient knowledge transfer (perhaps equivalently a greater use of codified knowledge). Proposition 2 states that in the initial period roommate matching generates more knowledge than does random matching. This is consistent with the hatched line in Figure 2 : good IT (represented by high values of α and consequently high values of β) are optimal when discounting is severe. Consistent with Proposition 4, the thin solid line shows that when discounting is weak, knowledge production is dominated by long-term performance, and so is highest when convergence in knowledge types takes place slowly (low α and β). For moderate levels of discounting, however, the trade-off between good matching and slow convergence yields an interior optimum, whereby moderately good IT performance produces reasonably good matching coupled with long-term prospects for diversity which eventually maximize discounted knowledge creation.
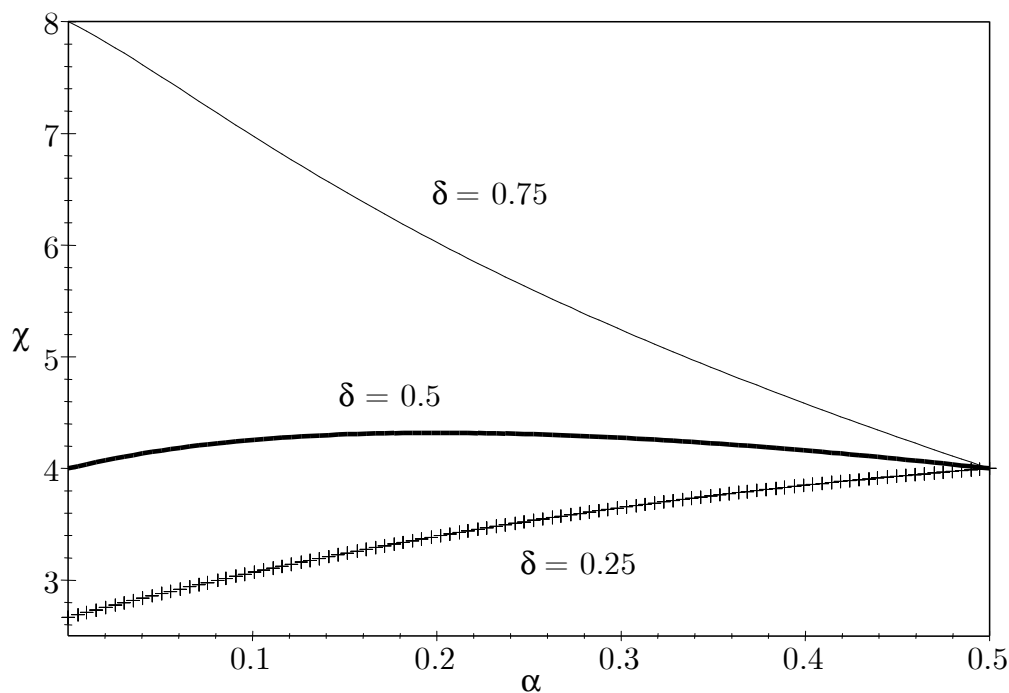
Figure 2: Effect of IT quality on discounted knowledge production.

6 Conclusion

The argument we have made in this paper is that in a changing industry in which joint R&D is common, improvements in information technology are both a good thing and a bad thing. They are good in a private, static sense, in that they assist firms in finding the best possible partner and in assimilating knowledge generated during the collaborative episode. They are bad in a public, dynamic sense in that they generate a convergence in knowledge types, which reduces innovation possibilities for the industry as a whole. The tradeoff between these two effects depends on the strength of discounting. Low discounting, putting weight on future knowledge gains makes it desirable to have imperfect ITs. When ITs are too good, firms learn too much from each other, converge quickly, and rapidly exhaust innovation opportunities. Maintaining variety among the population of firms is the key to long term growth in this case. High discounting, on the other hand, presents a situation in which high quality IT has immediate payoffs through the ability of firms to find good partners and so make big innovations in the short run. It is important to note that discounting is about impatience vis-à-vis the future, an implicit cost of waiting as perceived by the policy maker. In the context of innovation, what the policy maker is concerned about waiting for is new knowledge. The cost of waiting for new

knowledge to be created can be quite high when innovation has a significant effect on competitiveness. This can be the case both at the industry and the national level. So the tension between the short run and the long run depends on how technical change affects competition for the economic system as a whole.

In the model firm heterogeneity exists initially and is used as an innovative input, but eventually vanishes. Because the collapse of heterogeneity is a natural feature of the model it should be seen as capturing a single inventive episode. Regenerating diversity would be necessary for modeling a situation in which an industry presents ongoing innovation. In this case, the economic system as a whole would consist of various industries at different stages in their innovative cycles, with diversity being systematically introduced into one part of the economy by the innovations made in others.

The tension arises in this model because the two effects of IT are not independent. If the policy-maker is able to uncouple these two effects, the policy landscape is richer. If, for example it is not possible to affect the processes by which firms find partners, then the if increased knowledge production is the goal, the policy advice is unequivocal: act to preserve the diversity of knowledge within the industry, either by reducing inter-firm learning (perhaps by strengthening intellectual property rights) or by fostering exogenous variety generation. On the other hand, if it is not possible to affect the degree of inter-firm learning, and so the rate of convergence in expertise, policy can act on matching and the consequent formation of innovation networks. If expertise convergence is slow, improving the technology of matching will yield higher knowledge levels, whereas if convergence between partners is rapid, reducing the efficiency of the matching will increase knowledge levels. The latter may be a difficult policy to pursue, which leaves outside heterogeneity as the only source of ongoing innovation.

References

[1] Bloch, F. and Ryder, H. (2000), "Two-Sided Search, Marriages and Matchmakers", *International Economic Review*, 41, 93-115.

[2] Burdett, Ken, and Melvyn G. Coles (1997) "Marriage and Class", *Quarterly Journal of Economics*, 112, 141-168.

[3] Dudley, L. (1991), "Communications and Economic Growth", *European Economic Review*, 43, 595-619.

[4] Dyer, J.H. and K. Nobeoka (2000), "Creating and Managing a High Performance Knowledge Sharing Network: The Toyota Case", *Strategic Management Journal*, 21, 345-67.

[5] Freeman, C. (1991) "Networks of Innovators: A synthesis", *Research Policy*, 20, 499-514.

[6] Gale D. and Shapley L. (1962) "College Admissions and the Stability of Marriage", *American Mathematical Monthly*, 69, 9-15.

[7] Gaspar, J. and E. Glaeser (1998) "Information Technology and the Future of Cities", *Journal of Urban Economics*, 43, 136-156.

[8] Grant, R.M. (1996) "Towards a Knowledge-Based Theory of the Firm", *Strategic Management Journal*, 17 (Winter Special Issue), 109-122.

[9] Hagedoorn J. (1995) "Strategic Technology Partnering during the 1980s: Trends, Networks and Corporate Patterns in non-Core Technologies", *Research Policy*, 24, 207-231.

[10] Hagedoorn J. (2002), "Inter-Firm R&D Partnerships: an Overview of Major Trends and Patterns since 1960". *Research Policy*, 31, 477-492

[11] Kaufmann, Q., P. Lehner and F. Tödtling (2003) "Effects of the Internet on the Spatial Structure of Innovation Networks", *Information Economics and Policy*, 15, 402-424.

[12] Koestler, A. (1964) *The Act of Creation*, Macmillan, New York.

[13] Mowery, D.C., Oxley, J.E. and B.S. Silverman, (1998) "Technological Overlap and Interfirm Cooperation: Implications for the Resource-based View of the Firm", *Research Policy*, 27, 507-523.

[14] Noteboom B. (1999) *Inter-firm Alliances. Analysis and Design*, Routledge, London.

16

[15] Olsson, O. (2000), "Knowledge as a Set in Idea Space: An Epistemological View on Growth", *Journal of Economic Growth*, 5, 253-276.

[16] Peretto, P. and S. Smulders (2002) "Technological Distance, Growth and Scale Effects", *Economic Journal*, 112, 603-624.

[17] Powell, W.P., K. Koput, and L. Smith-Doerr, (1996) "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology", *Administrative Science Quarterly*, 41, 116-145.

[18] Poyago-Theotoky, J., Beath, J. and D. Siegel (2002) "Universities and Fundamental Research: Policy Implications of the Growth of University-Industry Partnerships", *Oxford Review of Economic Policy*, 18, 10-12.

[19] Sattinger, M.J. (1975) "Comparative Advantage and the Distributions of Earnings and Abilities", *Econometrica,* 455-468.

[20] Sattinger, M.J. (1995) "Search and the Efficient Assignment of Workers to Jobs", *International Economic Review*, 36, 283-300.

[21] Schumpeter, J.A. (1934) *The Theory of Economic Development.* Harvard University Press.

[22] van Alstyne, W. and E. Brynjolfsson (1996) "Widening Access and Narrowing Focus: Could the Internet Balkanize Science?", *Science*, 271, 1479-1480.

[23] van Alstyne, W. and E. Brynjolfsson (1997) "Electronic Communities: Global Village or Cyberbalkans?" Working Paper MIT Sloan School.

[24] Weitzman, M. (1998) "Recombinant Growth", *Quarterly Journal of Economics*, 113, 331-360.