

INFORMATION GATHERING THROUGH ALLIANCES

Wilko Letterie*,

John Hagedoorn,

Hans van Kranenburg and

Franz Palm

October 18, 2004

Abstract

The effect of a firm's uncertainty regarding technological development on the formation of alliances is examined and it is shown that this uncertainty is positively related to the number of alliances. However, higher uncertainty also makes the firm less likely to ally with others if the collected information is redundant. Yet, the higher the similarity between the technological regimes of potential alliance partners the lower the incentive to ally. Also direct ties are preferred to indirect ones if uncertainty is high since direct ties yield more accurate information than indirect ties.

Keywords:

Technology Alliances, Bayesian Learning, Redundant Information, Direct and Indirect Ties, Social and Economic Networks

JEL Codes:

D83, L10, L20

Acknowledgement

An earlier version of this paper was presented at the MERIT workshop on "Strategic Management, Innovation and Econometrics". The authors thank Geert Duysters, their discussants Reinhilde Veugelers and Pierre Mohnen, other workshop participants and two anonymous referees of this journal for their helpful and stimulating comments.

* Corresponding Author, Department of Organization and Strategy, Faculty of Economics and Business Administration, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands - telephone (31) 43-3883645 - fax: (31) 43-3884893, Email: W.Letterie@os.unimaas.nl. The other authors are also affiliated with Maastricht University, Faculty of Economics and Business Administration.

INFORMATION GATHERING THROUGH ALLIANCES

Introduction

In this paper we will consider some theoretical issues regarding the formation of optimal networks through inter-firm alliances from the perspective of an individual company. These inter-firm alliances are defined as collaborative agreements between independent companies. We will concentrate on technology alliances where companies share R&D and other innovative activities through a range of different collaborative agreements (Hagedoorn and Duysters, 2002). Given the context of these alliances, the main purpose of these collaborative activities is to learn about a particular new technology and to reduce the uncertainty surrounding this new technology.

We adopt a Bayesian learning framework (cf. Raiffa and Schlaifer, 1961; DeGroot, 1970). The results of these models have been interpreted in the context of alliance formation by Mody (1993) who argues that firms counter uncertainty by learning through alliances. Mody (1993) stresses that during each period an alliance exists, an experiment is conducted which yields an observation on the desirability of breaking up, continuing or merging. Our contribution is to determine the optimal number of alliances for a particular firm, when each alliance yields an observation on an uncertain technology. In fact, we presume that information has to be acquired from different sources. We find that a higher degree of technological uncertainty necessitates a higher number of alliances. Furthermore, we assume that contacts or ties are partly redundant. In that case, technological information obtained from various sources is to some extent similar. In other words, the firms that may potentially ally with each other share a common technological regime (cf. Dosi, 1982; Nelson and Winter, 1977). We study how this technological commonality affects the incentive to form alliances and we find that a higher degree of commonality reduces the incentive to form alliances.

Our model allows us to study two additional issues. First, according to the standard efficiency approach in network analysis (Burt 1992 a and b) it is the number of non-redundant contacts that counts. Companies should limit getting involved with networks where contacts are duplicated. We consider how uncertainty affects the incentive to form either redundant or non-redundant ties. We argue that non-redundant ties yield better information to the firm though they might be costlier to obtain. In fact, firms that already belong to the network in which a firm participates are usually less costly allies to develop because trust between the members facilitates further alliances (Gulati, 1995b; Nooteboom et al, 1997; Saxton, 1997). The costs of searching for new alliances within the existing network is lower as well (Gomes-Cassares, 1996; Gulati, 1995a; Uzzi, 1997). However, the disadvantage of such ties is that members of an existing network progressively become less attractive candidates to form an alliance with, because the information they yield becomes more and more similar. We show that under higher uncertainty the choice for redundant contacts becomes less likely.

Second, standard network theory argues that 'weak ties' or 'bridge ties' are beneficial in transmitting information from one group of social players to another (Granovetter, 1973; Burt, 1992 a and b). Hence, these ties allow a firm to obtain information from a source to which it is not connected directly. The main argument is that firms should aim at forming efficient networks by means of a limited number of direct ties that act as bridges to indirectly connected companies. The presumption here is that communication between firms that are not directly connected is accurate. We challenge this assumption and study the choice between direct and indirect contacts when information transmission through a bridge tie may be inaccurate. We show, in line with the 'closure' argument by Coleman (1988), that under higher uncertainty the choice for direct contacts becomes more likely in order to improve the communication between the ties.

The paper is structured as follows. In the first section we provide a general outline of the optimisation problem of a firm. In the second section, we investigate under which conditions it

becomes more important to form non-redundant ties. In the third section, we discuss the conditions determining when information obtained from indirect ties is dominated by information from direct ties. We draw some conclusions and provide some directions for future research in the final section.

1 A Bayesian model of information collection through forming alliances

This section presents an outline of the optimisation problem from the perspective of the management of a single, representative firm that is interested in information gathering about product or process innovation through alliances (see also Arora and Gambardella, 1994). The basic idea behind the model is that a firm is uncertain about the future direction in which its main technology will develop. To reduce this uncertainty surrounding technological innovation, the firm can form alliances with other firms to acquire useful information. In the following we adopt a Bayesian two-stage learning model as developed by among others Raiffa and Schlaifer (1961) and DeGroot (1970).¹ In the first stage, the firm decides on the number of its alliances. In the second stage, it decides about the direction of the technology it implements.

To develop some notation, consider a firm that is uncertain about the future technology it needs in order to survive. Suppose that the desired future direction of the firm's technology is reflected by a parameter T of which the value is unknown and outside the control of the firm. The uncertainty about the value of T is represented by a probability density function (pdf). The technological direction ultimately chosen by the firm is given by the decision parameter d . The firm has some prior expectations about the variable T , which are derived from information already

¹ Parts of this section are closely related to Cukierman (1980). In our model the parameter n reflects the number of alliances the firm selects. Cukierman interprets n as the optimal number of periods a firm may collect observations of x .

available. This means that the firm has some notion of the expected direction and degree of uncertainty surrounding technological developments. We formalize this idea by assuming that before the firm forms any alliance, the prior distribution of T is normal with mean μ and variance σ_T^2 :

$$(1) T \sim N(\mu, \sigma_T^2)$$

It is costly to choose a technology deviating from the desired one. Hence, the firm incurs a cost if the direction d does not match the desired direction given by T . In this study the costs when choosing direction d are given by (as an alternative a quadratic cost function could be used):

$$(2) C(T,d) = a|T-d|, \quad a > 0.$$

The idea is that the costs are influenced by the extent to which the direction d chosen by the firm fits with the desired technology represented by the variable T . The firm minimizes the expected value of $C(T,d)$ by choosing d^{opt} . This means that d^{opt} is determined by:

$$(3) \min_d E_T C(T,d),$$

where E denotes that expectations are taken with respect to T . If T is known with certainty, it is obviously optimal to select $d=T$. However, T is assumed to be unknown and the uncertainty about its value is represented by a prior probability distribution function as given by equation (1).

The firm does not need to fix d immediately. In fact, it has the opportunity to collect information about the properties of T by forming alliances with other firms. Before developing the

model in detail we provide an informal description of the information acquisition process by means of alliance formation. Each alliance yields information that is based on two components. First, since each potential alliance partner is different due to firm specific resources (Barney, 1991 and 2001; Eisenhardt and Schoonhoven, 1996; Madhok and Tallman, 1998) each alliance yields some unique information. This information component does not resemble information obtained from other alliances. In other words, this means that this piece of information is uncorrelated and hence is not comparable with that obtained from other alliance partners. Second, the potential partners belong to a group of firms in the sense that they share a common base of knowledge, as found in common technological regimes (Dosi, 1982; Nelson and Winter, 1977). A striking example of this is found in the personal computer industry. After the dominance of the DOS-Windows regime was established in the early 1990s, even companies with competing operating systems such as IBM and Apple became major collaborators on the integration of multimedia with Windows applications, the development of object-oriented software and RISC processing technologies in which both companies had established different capabilities (Hagedoorn, Carayannis and Alexander, 2001).

The implication of shared technological regimes is that information derived from different firms will show some overlap. In statistical terms this means that observations from these firms are correlated and hence are comparable to a certain extent. To analyze these aspects of alliance formation, we consider the following formal model: each time the firm sets up an alliance with another firm it receives information by observing the realization of a random variable x_i where $x_i = y_i + \varepsilon$. This means that each time information is acquired it is based on two different independent sources. In fact, the first variable y_i varies with each alliance formed by the firm. It is independently and identically normally distributed with $y \sim N(E_y, \sigma_y^2)$. This part of the model reflects that each alliance provides some new information due to partner specific resources. However, all observations x_i share a common component. This is captured by the second stochastic

term ε , which does not change with the alliances formed. It has a normal distribution $\varepsilon \sim N(E_\varepsilon, \sigma_\varepsilon^2)$. This assumption means, that information obtained by an alliance is based on some knowledge that is shared among the potential partners. We assume that $E_y + E_\varepsilon = T$.² Hence, the desired future direction of the technology T , is the mean of the normally distributed variable x . As a consequence, if the firm obtains an observation x , this yields some information about T . Note that the firm can only observe the sum of y_i and ε . It does not identify these two terms separately. The variances of the two components y_i and ε are assumed to be given and known.

If the firm obtains various observations by forming alliances they contain partly similar information. There is some overlap because the potential alliances share a technological regime. This comparability is reflected by the degree of correlation of the observations. To see this consider the n by n covariance matrix Σ_x^n of n observations x_1, x_2, \dots, x_n :

$$(4) \quad \Sigma_x^n = \begin{bmatrix} \sigma_y^2 + \sigma_\varepsilon^2 & \sigma_\varepsilon^2 & \cdot & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_y^2 + \sigma_\varepsilon^2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \cdot & \sigma_\varepsilon^2 & \sigma_y^2 + \sigma_\varepsilon^2 \end{bmatrix}$$

The correlation coefficient of two different observations x_i and x_j equals $\rho_{i,j} = \frac{\sigma_\varepsilon^2}{\sigma_y^2 + \sigma_\varepsilon^2} > 0$. It

appears that the correlation between x_i and x_j increases with the variance σ_ε^2 and decreases with the variance σ_y^2 . Hence, if the variance of ε is high relative to that of y , the information carried by the variable x will largely be influenced by the informational content of ε . Therefore, if the variance

² This assumption is not crucial to our main findings. One may also assume that $E_y = T$ (0) and $E_\varepsilon = 0$ (T). It is important that the sum of $E_y + E_\varepsilon$ equals T , since the mean of the sum of normally distributed variables is the sum of the respective means.

σ_ε^2 is high compared to the variance σ_y^2 , the two observations x_i and x_j will both be highly affected by the presence of the technological regime and much less by the alliance specific information (i.e. y_i), implying a high coefficient of correlation.

Information becomes available at a cost, which represents the cost of setting up an alliance and the cost of collecting and evaluating the information. The total cost per observation of x is given by c . If the firm decides to observe n values of the variable x before it chooses its direction d , the firm uses the acquired information to update its knowledge concerning the variable T using Bayesian learning.³ This knowledge is given by the posterior distribution of T which incorporates all information that is contained in the observations x_1, x_2, \dots, x_n . The parameters of the prior pdf (1) are assumed to be given and known. The variance of the random variable x in (4) is also assumed to be given and known. This latter assumption could be dropped easily. When the variance of x is unknown, the uncertainty surrounding its value could be represented by a prior pdf, e.g. using an inverted gamma pdf (see e.g. Raiffa and Schlaifer, 1961).

The optimal number of information-gathering alliances before the firm determines its optimal direction d^{opt} can be determined as follows. The firm chooses n to minimize the expected value of the cost function $C(T, d)$ presented in equation (2) given the posterior beliefs concerning T and the cost of collecting information given by nc . Hence, n is found by:

$$(5) C = \min_n \left(\min_d E C(T, d) + nc \right)$$

The n under the expectations operator in equation (6) denotes that d is chosen, using the posterior distribution of T after n observations of x . The minimized expected cost the firm incurs equals C .

³ Note that n refers to the number of alliances the firm selects.

In order to derive the optimal number of information-gathering alliances, we first need to determine the firm's optimal direction d^{opt} using equations (2) and (3). We find that:

$$(6) \min_d E C(T, d) = a \min_d E |T - d|$$

The optimal direction d^{opt} is equal to the median of the distribution of T (see DeGroot, 1970).

Since the posterior distribution of T is normal and therefore symmetrically distributed around its mean, the optimal strategy of the firm is given by $d^{opt} = \mu_n$, i.e. the expected value of T given its posterior distribution. As shown by DeGroot the minimized value of equation (6) is given by:

$$(6') \min_d E C(T, d) = a \sqrt{\frac{2\sigma_n^2}{\pi}}$$

It appears that the expected cost the firm incurs depends on the degree of uncertainty remaining after n observations as given by σ_n^2 . The higher the uncertainty as measured by σ_n^2 , the higher the cost. Therefore, the firm has an incentive to reduce the uncertainty it faces. Using equation (5) the firm's optimisation problem becomes:

$$(5') \min_n \left(a \sqrt{\frac{2\sigma_n^2}{\pi}} + nc \right)$$

The variance of T given its posterior distribution which incorporates all information x_1, x_2, \dots, x_n is given by:

$$(7) \sigma_n^2 = \left(\frac{1}{\sigma_T^2} + \frac{n}{\sigma_y^2 + n\sigma_\varepsilon^2} \right)^{-1}.$$

The derivation is given in the appendix. Before proceeding with the optimization problem of the firm let us consider two extreme cases. First, consider the situation where each partnership provides some alliance specific and commonly available information. Suppose now that commonly available information is very inaccurate: σ_ε^2 approaches infinity. Then the posterior variance σ_n^2 equals the variance of the prior distribution σ_T^2 . This means that in this case forming alliances does not provide any useful information since the degree of uncertainty faced by the firm remains equal to its prior degree of uncertainty given by σ_T^2 . The explanation is as follows. The correlation coefficient between two observations x_i and x_j , i.e. $\rho_{i,j} = \frac{\sigma_\varepsilon^2}{\sigma_y^2 + \sigma_\varepsilon^2}$, approaches one in this case. This means that two alliances provide exactly the same information. Hence, the maximum number of alliances that possibly provides relevant information is one. However, an observation x_i is drawn out of a distribution with a variance which approaches infinity. Therefore the estimate of T based on observation x_i is very inaccurate and does not lead to a reduction of uncertainty as measured by σ_n^2 . Hence, it is optimal to build no alliances.

Secondly, suppose now that the technological regime is strong. This means that information derived from an alliance is compelled by the technological regime and not by alliance specific information (i.e. y_i). In terms of our model, this happens when σ_y^2 is small relative to σ_ε^2 , because then as the correlation coefficient $\rho_{i,j} = \frac{\sigma_\varepsilon^2}{\sigma_y^2 + \sigma_\varepsilon^2}$ indicates two observations x_i and x_j , are highly correlated and comparable. Consider the extreme case where the variance of the idiosyncratic term is zero, i.e. $\sigma_y^2 = 0$. This implies that information derived from the alliance is

completely based on the shared technological regime. Hence, within the set of potential partners, information about the technology is public knowledge. Again the correlation coefficient $\rho_{i,j}$ equals one. Therefore one alliance is enough to acquire all information available. In fact the posterior variance of T after observing x is equal to $\sigma_n^2 = \left(\frac{1}{\sigma_T^2} + \frac{1}{\sigma_\varepsilon^2} \right)^{-1}$ which is independent of the number of observations: n . If every participant in the industry has a pretty good idea of what the technological regime looks like, i.e. when σ_ε^2 is very small, then the posterior variance will be small as well. Since σ_y^2 is small relative to σ_ε^2 the assumption that σ_ε^2 approaches zero means that σ_y^2 will be zero as well. Furthermore, if σ_ε^2 approaches zero, then σ_n^2 becomes zero as well. In this instance the firm knows the value of T precisely.

Let us now return to the optimization problem faced by the firm. The first and second order conditions for the optimization of equation (5) become:

$$\text{FOC : } -\frac{1}{2}a\sqrt{\frac{2}{\pi}}\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2}\right)^{-3/2}\left(\frac{\frac{\sigma_y^2}{n^2}}{\left(\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2\right)^2}\right) + c = 0$$

(8)

$$\text{SOC : } \frac{1}{2}a\sqrt{\frac{2}{\pi}}\frac{\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2}\right)^{-3/2}}{\left(\sigma_y^2 + n\sigma_\varepsilon^2\right)^3}\left(\frac{3}{2}\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2}\right)^{-1}\left(\frac{\sigma_y^4}{\sigma_y^2 + n\sigma_\varepsilon^2}\right) + 2\sigma_y^2\sigma_\varepsilon^2\right) > 0$$

As all components of the SOC are positive, the second order condition for a minimum is always satisfied. The FOC depicted above implicitly identifies the optimal number of alliances.⁴

Unfortunately, it is not possible to derive an explicit expression for the optimal n (taking n as a continuous variable, in practice one will take the integer value of the optimal value for n resulting from the optimisation problem). However, in the appendix we derive a number of comparative static results that we summarize in the following proposition:

Proposition 1

The optimal number of information gathering alliances, denoted by n ,

- decreases with the cost c of setting up an alliance
- increases with the cost attached to making mistakes as measured by the parameter a
- increases with the degree of uncertainty the firm faces as measured by σ_T^2
- decreases with the variance σ_ε^2 of the common random term ε .

The proposition indicates that the optimal number of alliances decreases with the cost, i.e. c , of setting up an alliance.⁵ If the cost associated with making mistakes increases, as measured by the parameter a , the optimal value of n also increases. It can be seen straightforwardly that the higher the initial uncertainty as measured by the variance σ_T^2 , the larger the number of alliances the firm

⁴ Note that if we substitute $\sigma_\varepsilon^2=0$ in the FOC depicted in equation (11) and solve for n we obtain the well known optimisation problem as studied by Raiffa and Schlaiffer (1961), DeGroot (1970) and Cukierman (1980). The SOC guarantees optimality of this case as well.

⁵ While the first result of proposition 1 is in line with a finding of Mody (1993), we do not account for uncertainty to also increase the possibility of opportunistic behaviour and hence the perceived costs of an alliance, which tend to reduce alliance activity.

is willing to form. Hence, there exists a relationship between a firm's incentive to collect information through alliances and the degree of uncertainty it faces. In classical social network theory this is referred to as 'gregariousness' (Erbe, 1962) where uncertainty is countered by increasing the number of contacts (i.e. alliances) that will lead to increasing flows of information. Some recent empirical studies found that under conditions of increasing uncertainty, companies use large numbers of alliances to improve their learning and information collection capability. These empirical findings refer to a variety of high-tech industries such as semiconductors (Gomes-Casseres, 1996), data processing (Hagedoorn and Duysters, 2002) and biotechnology (Powell et al, 1996; Walker et al, 1997).

The fourth result should be interpreted as follows. We noted before that $\rho_{i,j} = \frac{\sigma_{\varepsilon}^2}{\sigma_y^2 + \sigma_{\varepsilon}^2}$.

Hence, the correlation between x_i and x_j increases with the variance σ_{ε}^2 . This means that if the informational content of the various observations gathered by forming alliances becomes more similar, the optimal number of alliances decreases. In other words, the firm is less likely to form a multitude of alliances if these alliances yield more similar information and therefore have a higher degree of redundancy. Hence, in case the firms share a common technological regime the incentive to form alliances is reduced. This line of thought refers to the current efficiency approach in network analysis (Burt, 1992 a and b) which stresses that the size of a specific network of alliances of a company is not that important for the adequate transfer of information. What really counts is the number of non-redundant contacts, because it is assumed that redundant contacts carry the same information. By definition, dense networks involve a considerable degree of interaction between companies and many of these interactions are expected to be redundant and inefficient. In standard network analysis terminology this implies that the structural equivalence in a network (the degree of interaction with the same group of companies) and the cohesion in networks (the connectivity of companies) should be limited to benefit from its contacts (see also Knoke and

Kuklinski, 1982). Therefore, a company should avoid duplication of contacts, it should create well-informed and selective linkages that generate so-called structural autonomy and that exercise control over rewarding opportunities (i.e. the structural equivalence in its network should be small). According to e.g. Burt (1992a), the lower the number of structural equivalent partners that a firm faces, the more effective a firm's portfolio of alliances (see also Lorrain and White, 1971). The next section provides a more detailed illustration of this issue.

2 When does non-redundancy become important?

In the above, we considered the case where all observations have part of the information in common as these observations are correlated. In other words they share a common knowledge base. Now, we introduce a simple framework in which the firm may choose to acquire observations that are either correlated or uncorrelated with previous ones. Hence, we discuss under which conditions firms should try to avoid duplication of information through alliances.

Suppose now that two networks of firms exist, which we denote by network A and B. Observations obtained from alliances with companies that belong to network A are given by $x_i = y_i + \varepsilon_A$. Like in the previous section the variable y_i varies with each alliance formed by the firm and is independent across different firms. It is identically normally distributed with $y \sim N(E_y, \sigma_y^2)$ with variance assumed to be given and known. The observations x_i share a common term ε_A , which does not change with the alliances formed within network A. It has a normal distribution $\varepsilon_A \sim N(E_\varepsilon, \sigma_\varepsilon^2)$ with variance assumed to be given and known. This assumption implies that observations within a certain network are correlated and hence show some overlap. To put this differently, this means that firms that constitute a particular network are likely to provide comparable information. Support for this assumption is found in Gomes-Casseres (1996) and Uzzi (1997) who found that within a network of interacting firms knowledge across the constituents will become similar. To

keep the model tractable we do not formally derive how information becomes similar within a network. Network B is very similar in the sense that observations provided by firms that make up the group are given by $x_i = y_i + \varepsilon_B$ and the variable y_i varies independently with each alliance formed by the firm. For the sake of convenience we assume it is also identically normally distributed with $y \sim N(E_y, \sigma_y^2)$. The common component of the observations x_i in network B is ε_B which has a normal distribution $N(E_\varepsilon, \sigma_\varepsilon^2)$ as well to facilitate the discussion. We assume again that $E_y + E_\varepsilon = T$. The stochastic terms ε_A and ε_B are assumed to be independent. Hence, if observations are obtained from firms that belong to different networks they do not show any overlap. In statistical terms, these observations are not correlated. Furthermore, we presume that the variances of the stochastic terms ε_A and ε_B are equal. This assumption could be dropped. In fact the degree of precision of the information available within network A and B may depend on the size and structure of the network. However, for convenience and without affecting the main conclusions of our study we abstain from this possibility.

Suppose now that the firm has already formed an alliance with a firm in network A. The firm now belongs to network A and has access to information derived from the technological regime that is present in network A. Information derived from its alliance with a firm in network A is optimally included in its knowledge base by Bayesian learning. The firm we consider and the firms in network A have become similar to some extent in terms of their knowledge bases. In fact, an additional alliance with a firm from network A will yield information that partly overlaps with the information it already possesses (i.e. a new observation is correlated with an observation it already obtained). However, an alliance with a firm from network B will not show any overlap with its knowledge base (i.e. an observation from network B is not correlated with its observation from network A). Hence, our firm may build an alliance with a firm that is similar (from network A) or with a firm that is dissimilar (from network B) in terms of its knowledge base.

Consider now the firm's decision of setting up an alliance with another firm from either network A or B. Using the result depicted in equation (7) the variance of T after two observations from network A is given by:

$$(9) \quad \sigma_{AA}^2 = \left(\frac{1}{\sigma_T^2} + \frac{2}{\sigma_y^2 + 2\sigma_\varepsilon^2} \right)^{-1}$$

If the firm chooses to cooperate with a firm from network B the variance of each observation will be equal to $\sigma_x^2 = \sigma_y^2 + \sigma_\varepsilon^2$. Furthermore, the observations from the two different networks A and B are independent. Therefore, we find, in line with section 1, that after two observations the posterior variance of the variable T becomes:⁶

$$(10) \quad \sigma_{AB}^2 = \left(\frac{1}{\sigma_T^2} + \frac{2}{\sigma_y^2 + \sigma_\varepsilon^2} \right)^{-1}$$

The cost of forming an alliance with an additional company from network A is denoted c_A . If the firm chooses a partner from network B the cost is given by c_B . One may argue that the cost c_B is larger than c_A . For instance, once alliances have been established within a network, trust between the members may facilitate further alliances as found by Gulati (1995b), Nooteboom, Berger and Noorderhaven (1997) and Saxton (1997). Also, the cost of searching for new useful partners may be lower within the existing network (Gomes-Casseres, 1996; Gulati, 1995a; Uzzi, 1997).

However, as indicated in the above, companies might have an advantage in forming alliances with companies from network B. Let σ_A^2 denote the variance of the predictor of T after

⁶ To see this use equation (7) and fill in $n=2$, $\sigma_\varepsilon^2=0$ and $\sigma_y^2=\sigma_y^2+\sigma_\varepsilon^2$

one observation of a firm out of network A. The firm will choose a firm from network B if the marginal benefit of adding a network B observation is higher than the marginal benefit from additional information from network A:

$$(11) \quad -a \left(\sqrt{\frac{2\sigma_{AB}^2}{\pi}} - \sqrt{\frac{2\sigma_A^2}{\pi}} \right) - c_B > -a \left(\sqrt{\frac{2\sigma_{AA}^2}{\pi}} - \sqrt{\frac{2\sigma_A^2}{\pi}} \right) - c_A$$

Since, $\sigma_{AA}^2 > \sigma_{AB}^2$ the inequality of equation (11) will hold if the cost of setting up an alliance with a firm from B is relatively low. Furthermore, if the degree of uncertainty, i.e. σ_T^2 , faced by the firm is larger, equation (11) is also more likely to hold, because

$$(12) \quad \frac{\partial \left(-\sqrt{\frac{2\sigma_{AB}^2}{\pi}} + \sqrt{\frac{2\sigma_{AA}^2}{\pi}} \right)}{\partial \sigma_T^2} > 0.$$

This leads to our second proposition:

Proposition 2

If the environment faced by the firm is more uncertain as measured by σ_T^2 , the firm is more inclined to avoid duplication of information through alliances within its existing network.

Within the context of our model, assuming higher uncertainty, firms that are member of the existing network (here network A) gradually become less attractive allies, because they actually yield similar information. Hence, members of other networks to which the firm has no contacts yet become more appealing candidates to ally with. In fact existing networks may break down when

searching for new information is crucial. The Advanced Computing Environment (ACE) network, established during the 1990s, illustrates this point. This network broke down within a couple of years after its key members began to develop new technologies with other companies outside the ACE network (see Gomes-Casseres, 1996). Another interesting illustration of this point is found in the network of R&D alliances established by IBM during the 1990s.⁷ During the first half of this period, IBM entered into close-knit R&D cooperation through a series of multiple alliances with companies such as Apple, Siemens, Toshiba and Motorola. The technological emphasis in this inter-firm network was on R&D alliances in computer hardware and related activities such as computer-based telecommunication systems and supporting software. However, during the second half of the 1990s, most of these R&D alliances were terminated or only continued at a lower level of partnering intensity with a few short-term alliances. Then, given their focus on somewhat similar interest, many companies such as IBM started to establish R&D alliances in related information technology fields and in sectors outside information technology. These new inter-sectoral R&D alliances concentrated on fields such as microelectronics, software, various internet-related products and services, and a host of multimedia technologies. IBM began to collaborate extensively on joint R&D with a different set of companies with which it had no or very few prior R&D alliances. During that period IBM established multiple R&D alliances with Intel, Netscape, Novell, Oracle, Philips and Sun. Apparently, opportunities for further R&D cooperation with individual companies from the first group of partners, in which IBM was well-embedded through multiple dyadic ties, had dried up in a relatively short period time and other companies became attractive partners for further R&D collaboration.

⁷ This information is drawn from the MERIT-CATI database on inter-firm R&D alliances.

3 When do direct ties dominate indirect ties?

Standard social network theory assumes that it is beneficial for companies to access existing information through a limited number of direct contacts while avoiding direct links to dense inefficient networks. This line of reasoning is based on classical arguments such as for instance found in Granovetter's (1973) 'weak ties' that serve as bridges that can help to transfer information from one group of social players to another. A similar argument is made by Burt (1992 a, b) where 'structural holes' within networks are overarched by bridge ties with as little redundancy as possible.

In this section we present an alternative approach that determines when a direct tie, that carries information directly from another company, is preferred to an indirect tie where information is based on transfer from a company in a network through a single bridge tie. For the sake of convenience we abstract from the existence of technological regimes and assume that $\sigma_\varepsilon^2=0$, implying that different alliances do not yield comparable information. We consider three firms and in order to facilitate the discussion we label them A, B, and C. Suppose now that firm A considers the two following strategies. The first strategy involves the formation of an alliance with firm B. As before this yields an observation denoted x_B , which is normally distributed with mean T and variance σ_x^2 . We assume that firm B has already established an alliance with firm C. Therefore, firm B knows observation x_C , which is also normally distributed with mean T and variance σ_x^2 . As a consequence, an indirect observation of x_C is obtained as well by firm A if it allies with firm B. Hence, firm C is an indirect tie of firm A. However, the transmission of information from firm B to firm A is ambiguous. We assume that it is impossible to observe x_C directly without an error. Instead it is possible to receive a message $z_C = x_C + \eta$, where η is a transmission error which is normally distributed with zero mean and variance σ_η^2 . Hence, message z_C is normally distributed with mean T and variance $\sigma_x^2 + \sigma_\eta^2$. As a result, by only incurring cost c

the firm obtains two observations: x_B and z_C . In the sequel we assume that x_B , x_C , and η are independently distributed.

It is worth noting that the variable x_B does not represent firm B's best guess of the parameter T . If x_B would be its best estimate of T we should derive how observation x_C is incorporated into x_B . We prefer our approach because it is not very likely that firm A receives firm B's best guess of T . Often, firms protect important parts of their knowledge and they attempt to control the information they wish to release to their partners (Arora, 1995; Poppo and Zenger, 2002; Ring, 2002). However, this control of information transfer may be imperfect and some spill over of knowledge may occur (see also Arora, 1995; Goyal and Moraga-González, 2001; Poppo and Zenger, 2002). As a consequence it is possible that some of its information obtained from B's partner C is transmitted to firm A. According to our model the information received by A consists of two parts. One part is independent of firm B's observation of x_C . This is represented by the variable x_B . The other part does depend on observation x_C . However, firm A can only observe x_C imprecisely. This notion is captured by the variable $z_C = x_C + \eta$. We believe our approach comes closer to reality than assuming that firm A observes firm B's best guess.

It is straightforward to show within the present model that alliances yielding both direct and indirect ties are preferred to those that only yield direct ties. In terms of network analysis this implies that a company searches for alliances with other firms that are also well connected to their own specific networks. As a consequence, this company is characterized by both a high network centrality and a high betweenness centrality. This network centrality refers to the number of direct links of a particular company with other companies. In an information-network the possibility to control the flow of information between other companies is also dependent on a company's degree of betweenness centrality that refers to the number of times a company is located on the shortest geodesic path between other companies (Freeman, 1979; Hagedoorn and Duysters, 2002; Knoke and Kuklinski, 1982).

The second strategy the firm may pursue is to form direct alliances with both firm B and C (or forming a direct alliance with C). The advantage of this strategy is that in this case both x_B and x_C are observed without any error. The properties of these observations are the same as before. The cost of observing x_B and x_C equals $2c$. The first strategy is preferred to the second one if ⁸

$$(13) \quad a \cdot \sqrt{\frac{2}{\pi} \left(\frac{1}{\sigma_T^2} + \frac{1}{\sigma_x^2} + \frac{1}{\sigma_x^2 + \sigma_\eta^2} \right)^{-1}} - a \cdot \sqrt{\frac{2}{\pi} \left(\frac{1}{\sigma_T^2} + \frac{2}{\sigma_x^2} \right)^{-1}} - c < 0$$

It is straightforward to show that the left hand side of this expression increases with a higher degree of initial uncertainty, i.e. σ_T^2 , and increases with a higher degree of transmission ambiguity, i.e. σ_η^2 . These properties lead to our third proposition:

⁸ The first term on the right hand side of (13) can be found as follows. Take equation (A2) from the appendix where Σ_x^n is a two by two matrix with the variances of x_b and z_c on its diagonal and zeroes elsewhere. The vector i contains two numbers 1. The expression in brackets of (A2) is equal to the variance of T , which has to be substituted into equation (6) to yield the first element of (13). We assume that x_b and x_c are uncorrelated. This yields the second term on the right hand side of (13) if we apply equation (7).

Proposition 3

Direct ties are preferred to indirect ties if

- the cost of forming alliances, i.e. c , is low
- the cost of making mistakes as measured by a is high
- if uncertainty as measured by σ_T^2 is high
- the degree of transmission ambiguity as measured by σ_η^2 is high.

This implies that with increasing uncertainty a company will expand its number of direct links to a variety of other companies to improve the accuracy of communication between the ties, leading to dense networks. This result contradicts the efficiency approach within current social network theory, which states that under these conditions companies should aim at establishing efficient networks by means of a limited number of ties that act as bridges to other indirectly connected companies (Burt, 1992a; Rowley, Beherens and Krackhardt, 2000). However, our result is more in line with the ‘closure’ argument of Coleman (1988) who states that information rich dense networks of direct ties are more beneficial to network-actors than sparse networks. An illustration of this phenomenon is found in the biotechnology sector where many companies establish a dense and partly overlapping network with a multitude of well-connected partners that enables them to learn about interesting opportunities from a wide variety of sources (see Powell et al, 1996; Walker et al, 1997).

4 Conclusions

We employ a Bayesian learning model in which a firm obtains information about the desired future technology by forming alliances with other firms. We find that the number of firms a company will ally with increases with the amount of uncertainty it faces (cf. Mody, 1993). In classical social network theory this is referred to as ‘gregariousness’ (Erbe, 1962). However, we also argue that the incentive to form alliances decreases once the information of possible partners has become somewhat similar. In fact, we find that a firm is less likely to form many alliances if these different alliances share a common knowledge base. Such alliances have a higher degree of redundancy. This finding is in line with arguments advanced by Burt (1992 a and b) who argues that the number of non-redundant contacts matters. Furthermore, our paper indicates that under higher uncertainty redundant contacts become less attractive partners. If the information these contacts yield is very similar, even if such alliances are less costly to form, a firm will prefer non-redundant ties. Members of networks to which the firm has no ties yet and possess non-redundant information become appealing partners. Standard social network theory predicts that companies should establish efficient networks characterised by a limited number of contacts, that form ties to other indirectly connected companies. However, our model indicates that with higher uncertainty direct contacts, yielding more accurate information than indirect ties, become more attractive. This ultimately leads to dense networks.

Our model employs a number of specific assumptions. Hence, several extensions of the analysis are worth investigating, but they are beyond the scope of the present paper. First, our model discusses the optimisation problem from the point of view of a single firm. A next step is to extend our model to account for the fact that stable links require approval of all participants involved (cf. Jackson and Wolinsky, 1996). Second, we assume key parameters to be given and known to the management of the firm. This can be replaced by an assumption that the uncertainty regarding these parameters can be represented by some prior probability density function. Third, in

section 1 we assume that information obtained from alliances is similar. This degree of similarity is determined exogenously. An extension of the model could account for endogeneity, different technologies and strategic (learning) behaviour of companies. Such an extended model could use other functional forms for the objective function, for instance by allowing for increasing costs of managing a network and it could include other motives than collecting information for network formation (e.g. joint projects). Fourth, we assume that firms learn about one single technology as given by a variable T . The model can be easily extended to allow for learning about various aspects of technology and other matters of interest to the firm, by assuming T and x (i.e. the observation on the technology) to be vectors of values. A formal analysis of such extensions can be found in Raiffa and Schlaifer (1961) or in Zellner (1971). Fifth, one could allow for heterogeneity among potential participants in an alliance. For instance, if the most valued information provider is a rival firm, which has the lowest incentive to share its information, the costs of forming an alliance will be high. This case could be dealt with by differentiating the cost structure on the basis of prior expectations about the value of the information obtained.

It is also interesting to extend our line of modelling into some empirical models, testing major elements of the above in alternative empirical settings such as industries characterised by different levels of technological development. In that context the use of a combination of existing network indicators, derived from current social network analysis, and the development of new network measures seems both appropriate and necessary to further expand this line of work in an empirical direction.

Appendix

Derivation of expression (10)

$$\Sigma_x^n = \begin{bmatrix} \sigma_y^2 + \sigma_\varepsilon^2 & \sigma_\varepsilon^2 & \cdot & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_y^2 + \sigma_\varepsilon^2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \cdot & \sigma_\varepsilon^2 & \sigma_y^2 + \sigma_\varepsilon^2 \end{bmatrix} = (\sigma_y^2 + n\sigma_\varepsilon^2) \cdot \overline{J}_n + \sigma_y^2 \cdot E_n$$

where $\overline{J}_n = \frac{1}{n} J_n$ and J_n is an n by n matrix with all elements being equal to 1. The matrix

$E_n = I_n - \overline{J}_n$ where I_n is the n by n identity matrix. To derive the posterior variance σ_n^2 we first

note that $\overline{J}_n \cdot \overline{J}_n = \overline{J}_n$, $\overline{J}_n \cdot E_n = 0$, and $E_n \cdot E_n = E_n$. Using a variance decomposition method

well known in the analysis of panel data with random effects (Baltagi, p. 14, 1995) we find that

$$(A1) \quad (\Sigma_x^n)^{-1} = \frac{1}{\sigma_y^2 + n\sigma_\varepsilon^2} \cdot \overline{J}_n + \frac{1}{\sigma_y^2} \cdot E_n.$$

Therefore, the likelihood function

$$f(x_1, \dots, x_n, T) = f(x_1, \dots, x_n | T) \cdot f(T) \propto \exp\left(-\frac{1}{2}(x - iT)^T (\Sigma_x^n)^{-1} (x - iT)\right) \cdot \exp\left(-\frac{1}{2} \frac{(T - \mu)^2}{\sigma_T^2}\right)$$

where i is an n by 1 vector whose elements contain the number 1 and x is an n by 1 vector

containing the observations x_i . To determine the posterior variance of T it suffices to collect all

terms that involve T^2 . These are

$$(A2) \quad \left(i^T (\Sigma_x^n)^{-1} i + \frac{1}{\sigma_T^2} \right) T^2.$$

Using (A1), $i^T (\overline{J}_n) i = n$ and $i^T (E_n) i = 0$ equation (A2) can be rewritten as

$$\left(i^T \left(\frac{1}{\sigma_y^2 + n\sigma_\varepsilon^2} \cdot \bar{J}_n + \frac{1}{\sigma_y^2} \cdot E_n \right) i + \frac{1}{\sigma_T^2} \right) T^2 = \left(\frac{n}{\sigma_y^2 + n\sigma_\varepsilon^2} + \frac{1}{\sigma_T^2} \right) \cdot T^2$$

Therefore, the posterior variance of T is equal to

$$\sigma_n^2 = \left(\frac{n}{\sigma_y^2 + n\sigma_\varepsilon^2} + \frac{1}{\sigma_T^2} \right)^{-1}$$

Proof of Proposition 1

The proofs all follow the same idea. Since n is implicitly determined by equation (11) we derive the comparative static results as follows. Suppose we want to know the effect of a variable s on n .

Due to equation (11) it must hold that $\frac{\partial FOC}{\partial n} \cdot \frac{\partial n}{\partial s} + \frac{\partial FOC}{\partial s} = 0$. Since $\frac{\partial FOC}{\partial n} = SOC > 0$, it must

hold that $SOC \cdot \frac{\partial n}{\partial s} = -\frac{\partial FOC}{\partial s}$. Hence, the sign of the effect of s on n is the opposite of the sign of

the effect of s on FOC. We find

- $\frac{\partial FOC}{\partial \sigma_T^2} < 0$ (straightforward)
- $\frac{\partial FOC}{\partial \sigma_\varepsilon^2} > 0$. Since,

$$FOC = -\frac{1}{2} a \sqrt{\frac{2}{\pi}} \left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2} \right)^{-3/2} \left(\frac{\sigma_y^2}{(\sigma_y^2 + n\sigma_\varepsilon^2)^2} \right) + c$$

$$FOC = -K \left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2} \right)^{-3/2} \left(\frac{1}{(\sigma_y^2 + n\sigma_\varepsilon^2)^2} \right) + c,$$

where K collects some terms which are positive

$$\text{sign}\left(\frac{\partial FOC}{\partial \sigma_\varepsilon^2}\right) = \text{sign}\left(K\left(-\frac{3}{2}\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2}\right)^{\frac{1}{2}} \cdot \frac{(\sigma_y^2 + n\sigma_\varepsilon^2)^2}{\left(\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2\right)^2} + 2n\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2}\right)^{\frac{3}{2}}(\sigma_y^2 + n\sigma_\varepsilon^2)\right)\right) =$$

$$\text{sign}\left(K\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2}\right)^{\frac{1}{2}}(\sigma_y^2 + n\sigma_\varepsilon^2) \cdot \left(-\frac{3}{2} \frac{\sigma_y^2 + n\sigma_\varepsilon^2}{\left(\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2\right)^2} + 2n\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2} + \frac{1}{\sigma_T^2}\right)\right)\right) =$$

$$\text{sign}\left(-\frac{3}{2}n\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2}\right) + 2n\left(\frac{1}{\frac{\sigma_y^2}{n} + \sigma_\varepsilon^2}\right) + \frac{2n}{\sigma_T^2}\right)$$

- $\frac{\partial FOC}{\partial c} > 0$ (straightforward)
- $\frac{\partial FOC}{\partial a} < 0$ (straightforward)

References

- Arora, A., 1995, Licensing tacit knowledge: intellectual property rights and the market of know-how, *Economics of Innovation and Technology*, 4, pp. 41-59.
- Arora, A. and A. Gambardella, 1994, Evaluating technological information and utilizing it, *Journal of Economic Behavior and Organization*, 24, pp. 92-114.
- Baltagi, B.H., 1995, *Econometric analysis of panel data*, New York, John Wiley & Sons.
- Barney, J., 1991, Firm resources and sustained competitive advantage, *Journal of Management*, 17, pp. 99-120.
- Barney J., 2001, Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view, *Journal of Management*, 27, pp.643-650.
- Burt, R.S., 1992a, *Structural holes - The social structure of competition*, Cambridge (MA), Harvard University Press.
- Burt, R.S., 1992b, The social structure of competition, in N. Nohria and R.G. Eccles, *Networks and organizations*, Boston (MA), Harvard Business School Press, pp. 57-91.
- Coleman, J.S., 1988, Social capital in the creation of human capital, *American Journal of Sociology*, 94, pp. 95-120.
- Cukierman, A., 1980, The effects of uncertainty on investment under risk neutrality with endogenous information, *Journal of Political Economy*, 88, pp. 462-475.
- DeGroot, M.H., 1970, *Optimal statistical decisions*, New York, McGraw-Hill
- Dosi, G., 1982, Technological paradigms and technological trajectories, *Research Policy*, 11, pp. 147-162.
- Eisenhardt K.M. and C.B.Schoonhoven, 1996, Resource-based view of strategic alliance formation: strategic and social effects in entrepreneurial firms, *Organization Science*, 7, pp. 136-150.

- Erbe, W., 1962, Gregariousness, group membership, and the flow of information, *American Journal of Sociology*, 67, pp. 502-516.
- Freeman, L.C., 1979, Centrality in social networks, *Social Networks*, 1, pp. 215-239.
- Gomes-Casseres, B., 1996, *The alliance revolution - The new shape of business rivalry*, Cambridge (MA), Harvard University Press.
- Goyal S. and J.L. Moraga-González, 2001, R&D Networks, *Rand Journal of Economics*, vol. 32, pp. 686-707.
- Granovetter, M.S., 1973, The strength of weak ties, *American Journal of Sociology*, 78, pp. 1360-1380.
- Gulati, R., 1995a, Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances, *Academy of Management Journal*, 38, pp. 85-112.
- Gulati, R., 1995b, Social structure and alliance formation patterns: a longitudinal analysis, *Administrative Science Quarterly*, 40, pp. 619-652.
- Hagedoorn, J., E. Carayannis and J. Alexander 2001, Strange bedfellows in the personal computer industry: technology alliances between IBM and Apple, *Research Policy*, 30, pp. 837-849.
- Hagedoorn, J. and G. Duysters, 2002, Learning in dynamic inter-firm networks – the efficacy of multiple contacts, *Organization Studies*, 23, pp. 525-548.
- Jackson, M.O. and A. Wolinsky, 1996, A strategic model of social and economic networks, *Journal of Economic Theory*, 71, pp. 44-74.
- Knoke, D. and J.H. Kuklinski, 1982, *Network analysis*, London, Sage Publications.
- Lorrain F. and H.C. White, 1971, Structural equivalence of individuals in social networks, *Journal of Mathematical Sociology*, 1, pp. 49-80.
- Madhok, A. and S.B. Tallman, 1998, Resources, transactions and rents: managing value through interfirm collaborative relationships, *Organization Science*, 9, pp. 326-339.

- Mody, A., 1993, Learning through alliances, *Journal of Economic Behaviour and Organization*, 20, pp. 151-170.
- Nelson, R.R. and S.G. Winter, 1977, In search of useful theory of innovation, *Research Policy*, 6, pp. 36-75.
- Nooteboom, B., H. Berger and N.G. Noorderhaven, 1997, Effects of trust and governance on relational risk, *Academy of Management Journal*, 40, pp.308-338.
- Poppo, L. and T. Zenger, 2002, Do formal contracts and relational governance function as substitutes or complements?, *Strategic Management Journal*, 23, pp. 707-725.
- Powell, W., K. Koput and L. Smith-Doerr, 1996, Inter-organizational collaboration and the locus of innovation: networks of learning in biotechnology, *Administrative Science Quarterly*, 41, pp. 116-145.
- Raiffa, H. and R. Schlaifer, 1961, *Applied statistical decision theory*, Cambridge (MA), MIT Press.
- Ring, P.S., 2002, The role of contract in strategic alliances, in Contractor, F.J., Lorange, P., *Cooperative strategies and alliances*, Amsterdam: Pergamon, pp. 145- 162.
- Rowley, T., D. Behrens and D. Krackhardt, 2000, Redundant governance structures: an analysis of structural and relational embeddedness in the steel and semiconductor industries, *Strategic Management Journal*, 21, pp. 369-386.
- Saxton, T., 1997, The effects of partner and relationship characteristics on alliance outcomes, *Academy of Management Journal*, 40, pp. 483-498.
- Uzzi, B., 1997, Social structure and competition in interfirm networks: the paradox of embeddedness. *Administrative Science Quarterly*, 42, pp. 35-67.
- Walker, G., B. Kogut and W. Shan, 1997, Social capital, structural holes and the formation of an industry network, *Organization Science*, 8, pp. 109-125.

Zellner, A., 1971, *An introduction to Bayesian inference in econometrics*, Chichester, John Wiley and Sons.