
Anand Nair\textsuperscript{a,1}, Sangphet Hanvanich\textsuperscript{b,2}, S. Tamer Cavusgil\textsuperscript{c,*}

\textsuperscript{a}Department of Management Science, Moore School of Business, University of South Carolina, Columbia, SC 29208, USA
\textsuperscript{b}Williams College of Business, Xavier University, 3800 Victory Parkway, Cincinnati, OH 45207-3214, USA
\textsuperscript{c}The Eli Broad College of Business, Michigan State University, N370 North Business Complex, East Lansing, MI 48824-1122, USA

Received 16 January 2007; received in revised form 28 March 2007; accepted 31 August 2007

Abstract

Collaborative ventures—both equity-based partnerships as well as project-based alliances—have dominated the international business scene over the past two decades. By means of this study we investigate the patterns of related and unrelated collaborative venture formation. Using a large database of over 90,000 collaborative ventures formed during the 1985–2001 period, this study clusters collaborative ventures on the basis of the industry group and home country relatedness of the collaborating partners. Self-organizing map technique within neural network methodology is used to accomplish this objective. The clusters obtained from the self-organizing map form the basis for developing taxonomy of collaborative ventures in which the neurons underlying clusters are classified based on the country of origin and industry affiliations of the collaborating partners and the collaborative venture. The distinguishing characteristics of the clusters and the taxonomy help augment our current understanding of the formation of collaborative ventures.

\textsuperscript{*}Corresponding author. Tel.: +1 517 432 4320; fax: +1 517 432 4322.
\textit{E-mail addresses:} nair@moore.sc.edu (A. Nair), Hanvanich@xavier.edu (S. Hanvanich), cavusgil@bus.msu.edu (S. Tamer Cavusgil).
\textsuperscript{1}Tel.: +1 803 777 2648; fax: +1 803 777 3064.
\textsuperscript{2}Tel.: +1 513 745 3379; fax: +1 513 745 3692.

Keywords: Collaborative ventures; Relatedness; Taxonomy; Neural network; Pattern recognition
1. Introduction

Alliance partnerships have been an area of intense scholarly investigations. Researchers have suggested varying rationales for determining the choice of alliance partners. One group of scholars contends that complementary core competencies are central to alliance formation (e.g., Collis & Montgomery, 1995; Hennart & Reddy, 1997; Peteraf, 1993; Silverman, 1999). The underlying rationale here is the synergy that results from resource and skill complementarities between the partners (Das & Teng, 2000). Others argue that the degree of similarity between the partnering firms is positively related to alliance formation and performance (e.g., Akhavan, Berger, & Humphrey, 1997). The similarity argument is rooted in the notion of “relatedness,” both in terms of industry affiliations and home country relatedness.

Industry and home country frequently are cited as factors that are indicative of relatedness among firms engaged in collaborative ventures (e.g., Palich & Gomez-Mejia, 1999; Silverman, 1999). These variables are commonly examined in terms of the closeness between partners’ industries or home country (e.g., Kogut & Singh, 1988) as well as the closeness between parent firms and their collaborative ventures (Merchant & Schendel, 2000). Most researchers argue that relatedness can be regarded as a distance. For example, cultural distance (Kogut & Singh, 1988) captures the degree to which various countries are related. While the relatedness concept has contributed to our knowledge of alliance formation, it also raises a key question: if a joint venture (JV) is in an industry related to one parent but not the other, should this partnership be classified as a related or an unrelated collaborative venture? Some attempt has been made to address this question. Reuer and Koza (2000), for instance, categorize JVs into four groups: parents and JV in the same industry, parents in the same industry but JV in a different industry, JV in the same industry with one parent, and parents and JV in different industries. It was found that only the last two groups are associated with positive shareholder values. This research points to the importance of approaching the relatedness factor from a multidimensional perspective. Furthermore, this issue is compounded by an increasing evidence of alliance formation between partners from somewhat unrelated industries and different countries (Reuer & Koza, 2000; Steensma & Lyles, 2000).

Consensus has yet to be reached as to which of the competing arguments is a valid explanation for alliance formation. Hence, it is important to explore this issue in a generalized setting. The present research represents an attempt towards such an exploratory examination. Specifically, we explore the underlying clusters in collaborative venture formation and their associated characteristics using JV data from 1985–2001. There are three research objectives for this study:

- Explore the underlying patterns in collaborative venture formation during 1985–2001 by allowing the data to self-organize into distinct clusters.
- Develop taxonomy of collaborative ventures formed during the 1985–2001 period based on industry and country of origin of collaborating partners and the collaborative venture.
- Delineate the distinguishing characteristics of the various classes underlying the taxonomy.

A preliminary examination of the dataset indicated that the required assumptions for applying statistical techniques were not met. Hence, the use of econometric and
psychometric approaches was considered inappropriate for clustering the data and to develop the taxonomy. We employ neural networks to detect patterns in collaborative venture formation and to examine the relationship between collaborating partners' and the collaborative JV’s industry and national affiliation. Neural networks have recently gained increasing visibility in business research (e.g. Hu, Zhang, & Chen, 2004; Lin, Chen, & Nunamaker, 1999/2000; Montagno, Sexton, & Smith 2002; Smith & Aggoune 2003; Veiga, Lubatkin, Calori, Very, & Tung, 2000; Wray, Palmer, & Bejou, 1994). Among the various neural network techniques, we use the unsupervised learning technique called self-organizing maps (SOM). This technique does not require any prior model specifications and does not have to satisfy the statistical assumptions, thus making it appropriate for our exploratory study.

The remainder of this article is organized as follows. We begin with a brief review of the literature and then present our conceptual framework in Section 2. In Section 3, we discuss the research design, methods and the data. Sections 4 and 5 presents the results of our analysis, which is followed by a discussion of key implications of these results. Finally, in Section 6, we offer conclusions and suggest directions for future research.

2. Alliance formation in the literature

In this section, we develop the rationale for the present research with a review of the extant literature pertaining to alliance formation. There is a vast body of literature that focus on alliances and the issue of relatedness and diversity among the collaborating firms. Since the 1980s, scholars have observed the proliferation and increased popularity of alliance formation (Chung, Singh, & Lee, 2000; Gulati, 1998; Gulati, Nohria, & Zaheer, 2000; Hagedoorn, 1993, 1995; Kale, Dyer, & Singh, 2002). An inter-firm alliance can be defined as a voluntary arrangement among partnering firms that exchange or share resources and engage in collaboration on some aspect of value chain (Gulati, 1998). Alliances can take different forms, including equity JVs, franchising, reciprocal trade agreements, project-based partnerships, and affiliations in research consortia.

Firms have exhibited an increasing tendency to seek collaborative ventures over the past two decades, coinciding with the newest phase of globalization. Interestingly, the nature and the scope of these collaborative ventures have also changed. While equity-based JVs prevailed prior to the 1980s, the modern times have seen the proliferation of non-equity, project-based, strategic alliances. In addition, Powell, Koput, and Smith-Doer (1996) note, whereas firms have previously engaged alliances for performing relatively simple peripheral activities, in recent years alliances are being used at various stages of R&D, production and marketing, and in almost all types of industries.

Research on alliance formation has focused on several distinguishing criteria for partnership. For instance, Geringer (1988) classified alliance formation based on the relatedness of the firms in terms of: (i) operational skills and resources critical to competitive success (task-related criteria); or (ii) the efficiency and effectiveness of the partner (partner-related criteria). Geringer remarked on the importance of task-related compatibility for successful collaboration. Later, Glaister and Buckley (1996) used the typology suggested by Geringer (1991) for a sample of UK firms and established a relationship between task-related criteria and motivation for alliance formation. Other scholars advocated additional factors such as corporate and national culture, organization identity adaptation, strategic as well as organizational and financial traits of the partner.
(Luo, 1998; Lerpold, 2003). Yet, the extant literature has failed to reach a consensus in terms of a definitive and comprehensive list of criteria for alliance formation. For example, researchers have proposed the importance of unrelatedness in collaborative venture formation by suggesting that complementary resources and skills obtained from an unrelated partner are more important for competitive success (Mowery, Oxley, & Silverman, 1996).

The present investigation focuses on relatedness, or lack thereof, in alliance formation with respect to: (i) industry group of the partner firms and the collaborative venture; and (ii) home countries of the partner firms and the collaborative venture. The following discussion is intended to delineate the relevance of these two issues to collaborative venture formation.

2.1. Industry effects

Industry structure and strategic positioning within the industry traditionally have been emphasized as sources of competitive advantage and drivers of long-term firm performance (Porter, 1985). In cooperative strategy research, industry effect has been approached from the perspective of the collaborative venture, the participating firms, or the relatedness between them. Madhavan and Prescott (1995), for example, suggest that differential synergy, measured through shareholder value creation, is due to different information-processing loads inherited from the industries of participating firms. Likewise, Anand and Khanna (2000) treat the industry factor for alliances as a fixed effect that can partially explain differential abnormal returns and wealth effects of the parent firms.

In an attempt to capture industry synergy more closely, researchers commonly measure the effect of industry relatedness on firm performance. For instance, Merchant and Schendel (2000) suggest that parent and collaborative venture relatedness is positively associated with shareholder value. This argument is commonly based on the synergy provided by the similarity of businesses conducted by participating firms and/or collaborative ventures. As Merchant and Schendel (2000, p. 726) contend, “...greater similarity between businesses of these entities confers scale and/or scope economies upon these firms.” The scale and scope of economies are believed to help firms raise entry barriers against potential entrants (Contractor & Lorange, 1988). Accordingly, firms are more likely to form collaborative ventures with partners in the related industries.

While firms in the related industries may be likely to form partnerships in order to leverage their compatible resources, another research stream suggests that firms from different industries are more probable to form partnerships because the lack of similarity in the core business or capabilities between partners makes the partnerships more likely to succeed (Harrigan, 1988). The latter research stream contends that lack of similarities between firms in the different industry ensures that the collaborative partners bring different but valuable capabilities to the relationships and thus enhances the learning between partners (Mowery et al., 1996). When firms are in an intense competitive environment, learning from the collaborative venture partners becomes even more important because it provides the firms with new knowledge and skills that are vital for their survival (Norman, 2004). This argument suggests that the industry relatedness may not be sufficient for firms seeking stability and reduced uncertainty in a hypercompetitive environment (Rindova & Kotha, 2001).
Advocates of the dynamic capability perspective reason that the industry influences a firm’s resource deployment because it affects resource reconfiguration and effective evolution (Eisenhardt & Martin, 2000). Firms that must cope with hypercompetition, they suggest, use morphing (or rapid process transformation) as a way to pursue competitive advantage in the shifting environment. As a result, firms are encouraged to form collaborative ventures with partners in the unrelated industries as they constantly seek new capabilities to fulfill their process transformations. Consequently, resource leverage and resource transformation will collectively create a pattern in which firms from highly related industries as well as firms from distinctively unrelated industries come together to form partnerships. The two competing lines of thoughts provide the motivation to explore the data to uncover the underlying patterns in collaborative JV formations during an extended period of time spanning 1985–2001. Further, it would be insightful to understand any systematic pattern among the collaborative ventures that are formed among partners and collaborative ventures belonging to related or unrelated industries. This study aims at addressing these important issues pertaining to industry relatedness in collaborative venture formation.

2.2. Home country effects

Rapid changes in competitive environments and proliferation of globalization has substantially increased the number of international alliance formations (Glaister & Buckley, 1994; Harrigan, 1986). With an increase in the cross-border partnerships, the issue of how the differences or dissimilarities between the partnering firms’ home countries affect performances of collaborative ventures and their partnering firms has become a focus in international cooperative strategy research. While some researchers approach cultural dissimilarities from the perspective of organizational culture differences (e.g., Simonin, 1999), others view cultural dissimilarities as cultural differences between the partnering firms’ home countries (e.g., Barkema & Vermeulen, 1997; Merchant & Schendel, 2000). The cultural difference between the partnering firms’ home countries is typically captured in a form of cultural distance, which is an arithmetic average of the squared deviations of each country’s score obtained from Hofstede’s (1983) cultural dimensions (Kogut & Singh, 1988). In recent literature, researchers have used financial regions (akin to trading blocks) to represent relatedness among countries forming collaborative ventures (e.g., Glaister & Buckley, 1996; Hitt, Dacin, Levitas, Edhec, & Borza, 2000). We adopt such a perspective in characterizing related and unrelated home countries of partnering firms in a collaborative venture.

The effect of home country unrelatedness on the collaborative ventures and partnering firms’ performance is rather inconclusive. Oum, Park, Kim, and Yu (2004), for instance, found that the home country unrelatedness did not have significant impact on the partnering firms’ profitability. Likewise, Park and Ungson (1997) found that cultural distance generally did not have an affect on JV dissolution. Interestingly, when the impact of home country unrelatedness was significant, the effect was usually found to be negative. For example, the home country difference has been found to negatively affect IJV survival (Hennart & Zeng, 2002). Home country unrelatedness has also been cited as a cause of conflict and mistrust between alliance partners (Barkema, Bell, & Pennings, 1996), and a factor inhibiting effective operation between partners (Barkema, Shenkar, Vermeulen, & Bell, 1997). Home country unrelatedness may also inhibit effective communication, a
crucial element in the success of collaborative ventures (Phan, Styles, & Patterson, 2005). Consequently, it is believed that firms are more likely to form collaborative ventures with partners from related home country than from unrelated home country.

The argument above, however, is challenged by the notion that firms are often motivated to form collaborative ventures in order to gain knowledge of local business practices or general knowledge of the economy (Beamish, 1987; Inkpen & Beamish, 1997). While participating in collaborative venture with unrelated home country partners was often found to have negative effects on the collaborative venture and collaborating firms’ performance, such collaborations, in the long run, may become beneficial to firms competing in intensely competitive markets. From a dynamic capability viewpoint, collaborative venture with unrelated home country partners could be a valuable source of market knowledge. Such collaborations provide firms with the experience that may be necessary to adapt and evolve in the new business environment. International experience can be a valuable, rare, and difficult to imitate resource that may lead to organizational competitive advantage (Carpenter, Sanders, & Hal, 2001). Firms can gain organizational experience from their international partners, even if some of the collaborations fail. Mistakes are commonly viewed as an end of the learning process, but the dynamic capability perspective regards them as lessons for the future (Arino & de la Torre, 1998). Even though major failures raise the learning barrier, small failures motivate managers to focus on the process, create greater engagement in learning, and do not trigger defenses that inhibit learning (Eisenhardt & Martin, 2000). Therefore, it is likely that firms will seek distinctive international knowledge from collaborative partners that are located in unrelated home countries as well as those in related home countries. To gain further insights into this issue, we undertake an extensive empirical examination of the home country relatedness/unrelatedness among collaborative ventures that were formed during the 1985–2001 period. The evidence would provide directions for further understanding of collaborative venture formation. Fig. 1 depicts the relationships that are the basis of this investigation.

3. Research design

3.1. Data characteristics

We relate the partnering firms involved in collaborative ventures according to their industry or home country. The contention is that any potential synergies derived from these ventures will be reflected in the frequency of their formation. From the Thomson Financial Security Database we collected information on collaborative ventures (both equity JVs and non-equity alliances) in manufacturing and service industries formed between 1985 and 2001. The database dates from 1980 and uses such sources as Securities and Exchange Commission filings and trade publications. The pool of 90,529 collaborative ventures was divided into two time-series samples: those formed between 1 January 1985 and 31 December 1995; and those formed between 1 January 1996 and 31 December 2001.

The dataset has six-dimensions: (i) Partner 1’s industry, (ii) Partner 2’s industry, (iii) Collaborative venture’s industry, (iv) Partner 1’s home country, (v) Partner 2’s home country and (vi) Collaborative venture’s home country. Industries are categorized into 11 groups based on their products/services and the nature of the manufacturing process.
These categories are in line with the standard SIC system. Most recent studies have grouped firms into cultural or financial regions (e.g. Glaister & Buckley, 1996; Hitt et al., 2000). In line with these studies, home countries are categorized into 16 groups based on geographical area and affiliation with trading blocks. The data were mean-centered. Table 1 summarizes the industry groups, and Table 2 lists the home country classifications.

After missing data were deleted, the final samples were: training data (1985–1995): 36,034; testing data (1996–2001): 36,034; and calibration data (1996–2001): 8508. The rationale for the use of these three datasets is provided later in the paper.

3.2. Research method

Econometric and psychometric techniques such as regression and structural equations modeling are well suited for causal modeling and path analysis. A preliminary examination of the dataset reveals a lack of independence among the predictor variables. Moreover, the statistical assumption of multinomial normality is also not satisfied. This makes the available statistical techniques inappropriate for analyzing the dataset either in a regression or a cluster analysis framework.

In contrast, the neural network approach is not constrained by the stated violations of statistical assumptions and is capable of unraveling the clusters in data that are characterized by high dimensionality. The ability of these techniques to depict patterns associated with complex relationships among variables (Haykin, 1999) makes them especially appropriate for the context considered in this study.

Fig. 1. Industry and home country of partners and the collaborative venture.
Table 1  
Industry list

<table>
<thead>
<tr>
<th>Industry category</th>
<th>Category code</th>
<th>Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discrete manufacturing</td>
<td>1</td>
<td>Machinery; Metal and Metal Products; Stone, Clay, Glass, and Concrete Products; Transportation Equipment</td>
</tr>
<tr>
<td>Consumer goods</td>
<td>2</td>
<td>Food and Kindred Products; Measuring; Medical; Photo Equipment; Clocks, Soaps, Cosmetics, and Personal-Care Products; Tobacco Products; Wood Products; Furniture and Fixtures</td>
</tr>
<tr>
<td>High technology</td>
<td>3</td>
<td>Drugs; Computer and Office Equipment; Prepackaged Software; Electronic and Electrical Equipment; Communications Equipment</td>
</tr>
<tr>
<td>Process manufacturing</td>
<td>4</td>
<td>Textile and Apparel Products; Paper and Allied Products; Chemicals and Allied Products; Rubber and Miscellaneous Plastic Products; Leather and Leather Products</td>
</tr>
<tr>
<td>Projects</td>
<td>5</td>
<td>Construction Firms; Aerospace and Aircraft</td>
</tr>
<tr>
<td>Miscellaneous manufacturing</td>
<td>6</td>
<td>Miscellaneous Manufacturing</td>
</tr>
<tr>
<td>Services</td>
<td>7</td>
<td>Transportation and Shipping (except air); Air Transportation and Shipping; Telecommunications; Radio and Television Broadcasting Stations; Printing, Publishing, and Allied Services; Electric, Gas, and Water Distribution; Sanitary Services; Hotels and Casinos; Amusement and Recreation Services; Motion Picture Production and Distribution; Personal Services; Business Services; Advertising Services; Repair Services; Health Services; Legal Services; Educational Services; Social Services; Miscellaneous Services</td>
</tr>
<tr>
<td>Trade</td>
<td>8</td>
<td>Wholesale Trade, Durable Goods; Wholesale Trade, Nondurable Goods; Retail Trade, General Merchandise and Apparel; Retail Trade, Food Stores; Retail Trade, Eating and Drinking Places; Retail Trade, Home Furnishings; Miscellaneous Retail Trade</td>
</tr>
<tr>
<td>Financial</td>
<td>9</td>
<td>Commercial Banks; Bank Holding Companies; Savings and Loans; Mutual Savings Banks; Credit Institutions; Mortgage Bankers and Brokers; Investment and Commodity Firms/Dealers; Insurance; Other Financial</td>
</tr>
<tr>
<td>Natural resources</td>
<td>10</td>
<td>Agriculture; Forestry and Fishing; Mining; Oil and Gas; Petroleum Refining</td>
</tr>
<tr>
<td>Others</td>
<td>11</td>
<td>Holding Companies (except banks); Public Administration; Nonclassifiable Establishments; Unknown</td>
</tr>
</tbody>
</table>

Table 2  
Home country list

<table>
<thead>
<tr>
<th>Home country category</th>
<th>Category code</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australasia</td>
<td>1</td>
<td>American Somoa (US), Australia, Cook Islands, Federated States of Micronesia, Fiji, French Polynesia, Guam (US), Kiribati, Marshall Islands, N. Mariana Islands (US), Nauru, New Caledonia (France), New Zealand, Niue (New Zealand), Norfolk Islands (Australia), Palau, Papua New Guinea, Solomon Islands, Tokelau (New Zealand), Tonga, Tuvalu, Vanuatu (New Hebrides), Wallis/Futuna Islands (France) and Western Somoa</td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th>Home country Category</th>
<th>Category Code</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caribbean</td>
<td>2</td>
<td>Anguilla (UK), Antigua, Aruba, Bahamas, Barbados, Belize, Bermuda, British Virgin Islands (UK), Cayman Islands, Commonwealth of Dominica, Cuba, Dominican Republic, Grenada, Guyana, Haiti, Jamaica, Martinique, Montserrat (UK), Netherlands Antilles, Puerto Rico, St Kitts and Nevis, St Lucia, St Vincent and The Grenadines, Surinam (Suriname), Trinidad and Tobago, Turks/Caicos Islands (UK) and US Virgin Islands (US)</td>
</tr>
<tr>
<td>Central America</td>
<td>3</td>
<td>Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua and Panama</td>
</tr>
<tr>
<td>Central Asia</td>
<td>4</td>
<td>Afghanistan, Armenia, Azerbaijan, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>5</td>
<td>Albania, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Czechoslovakia, East Germany, Estonia, Georgia, Hungary, Latvia, Lithuania, Macedonia, Moldova, Poland, Romania, Russian Federation, Slovak Republic, Slovenia, Soviet Union, Ukraine and Yugoslavia</td>
</tr>
<tr>
<td>Japan</td>
<td>6</td>
<td>Japan</td>
</tr>
<tr>
<td>Middle East</td>
<td>7</td>
<td>Abu Dhabi, Bahrain, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Syria, United Arab Emirates and Yemen</td>
</tr>
<tr>
<td>North Africa</td>
<td>8</td>
<td>Algeria, Egypt, Libya, Morocco and Tunisia</td>
</tr>
<tr>
<td>North America</td>
<td>9</td>
<td>Canada, St. Pierre/Miquelon (France), US of America</td>
</tr>
<tr>
<td>North Asia</td>
<td>10</td>
<td>China, Hong Kong, Macau, Mongolia, North Korea, South Korea and Taiwan</td>
</tr>
<tr>
<td>South America</td>
<td>11</td>
<td>Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Falkland Islands (UK), Paraguay, Peru, Uruguay, Venezuela</td>
</tr>
<tr>
<td>South Asia</td>
<td>12</td>
<td>Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, Sri Lanka</td>
</tr>
<tr>
<td>Southeast Asia</td>
<td>13</td>
<td>Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar (Burma), Philippines, Singapore, Thailand and Vietnam</td>
</tr>
<tr>
<td>Sub-Saharan Africa</td>
<td>14</td>
<td>Angola, Benin, Botswana, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Dem. Rep. of the Congo, Djibouti, Equatorial, Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Ivory Coast, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mayotte (France), Mozambique, Namibia, Niger, Nigeria, Republic of Congo, Reunion (France), Rwanda, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, St. Helena (UK), Sudan, Swaziland, Tanzania, Togo, Uganda, Upper Volta (Burkina Faso), Western Sahara (Morocco), Zaire, Zambia and Zimbabwe</td>
</tr>
<tr>
<td>Supranational</td>
<td>15</td>
<td>Multinational and Supranational</td>
</tr>
<tr>
<td>Western Europe</td>
<td>16</td>
<td>Andorra, Austria, Belgium, Cyprus, Denmark, Faroe Islands, Finland, France, Germany, Gibraltar, Greece, Greenland, Guernsey, Holy See (Vatican City), Iceland, Isle of Man, Italy, Jersey, Liechtenstein, Luxembourg, Malta, Monaco, Netherlands, Norway, Portugal, Republic of Ireland, San Marino, Spain, Svalbard/Jan Mayer Islands (Norway), Sweden, Switzerland, Turkey and UK</td>
</tr>
</tbody>
</table>
3.3. Neural network technique

Artificial neural networks can be traced to the work of Hebb (1949), who proposed a learning law that became the basis of modern neural network training techniques. Rosenblatt (1958), Widrow and Hoff (1960), and many other researchers contributed substantially in varied ways. Werbos (1974) and Parker (1985) proposed network architectures and the generalized Delta rule, which could be used for complex networks. The seminal work by Rumelhart and McClelland (1986) popularized that learning rule and suggested many applications. More recently, White (1992) clarified the relationship between artificial neural networks and fields like statistics and econometrics. Over the years many network architectures, differing in structure and learning methods, have emerged. Some are very general, such as the feed-forward networks with back-propagation and self-organizing Kohonen maps, and others are more problem specific, such as the ART networks (Carpenter & Grossberg, 1987) and counter propagation networks (Hecht-Nielsen, 1987).

In general, neural networks are advanced mathematical and statistical techniques that automate the analysis of large databases. They uncover complex relationships in the data and build predictive models based on those relationships. Their success stems from their ability to perform: (1) statistical analysis and identification of rules underlying economic systems; and (2) simulation and pattern recognition of micro and macro behavior. The present study uses the ability of feed-forward neural networks to uncover patterns underlying a dataset. Feed-forward networks with back-propagation are arguably the most widely adopted by researchers, and they are most suitable for research in business and economics.

We define fields of input and output variables and based on these variables the neural network “learns” the association between the input and outputs sets. The neural network produces \( n \)-dimensional relationships between the two by going through a large number of iterations. In the present inquiry, the input is information about the industry or home country of partner firms and collaborative ventures, and the output is weights representing the relationship patterns of the industries or home countries.

Learning in neural networks entails training the network by presenting patterns from the training set. The aim of training is to set the connection weights and unit thresholds to some proper value, so that the desired network computation is obtained. The learning process is therefore a search in weight space. We start out from some position in that space. A training pattern is then presented to the network, and the network makes a move (i.e., adjusts its weights) according to some transition rule, commonly termed the learning rule. In most learning strategies, the training proceeds by minimizing energy defined as a measure of the error made by the network when presented with the training set.

The strategies of learning are usually divided into two classes: supervised and unsupervised. In the latter, the training set consists of input vectors only, and the network determines the output during the course of training. It is in effect an autonomous, self-organizing process. In supervised training, the output is a part of the training set. After clamping the input units with the input vector, the network computes the output. The network is then told whether its output is right or wrong. In this research, we adopt unsupervised learning. After the learning phase we use the testing data to validate the patterns.
3.4. Neural network model

Prediction of economic time series is by far the most intensely studied application of neural networks in the area of economics and business. In complex prediction tasks, neural networks can also be very useful for performing exploratory data analysis (Kaski & Kohonen, 1995). The central goal in exploratory analysis is to present a dataset in a form that is easily understandable but, at the same time, preserve as much essential information of the original dataset as possible. The exploratory data analysis methods are general-purpose instruments that illustrate the essential features of a dataset, such as its clustering structure and the relations among the data items.

The method is particularly appropriate to an investigation that involves capturing a pattern embedded in time-series data. Most of the other techniques cannot explore the dynamic relationship in collaborative ventures, because they are required to start out with an assumption about the form of the relationship. Neural networks are ideal for capturing patterns without making any modeling assumptions. Such patterns in earlier studies have been avoided as it is a challenge to make assumptions on the expected form of the patterns. Even when researchers are willing to assume nonlinearity, no evidence is provided as to the form or degree of nonlinearity. Unknown nonlinear pattern of association among variables in the data highlights the underlying complexities. Neural networks cater to this need directly and provide a superior analysis technique by letting the data organize themselves, thereby enabling clustering of the data.

In the purview of the neural network techniques, we study the process of collaborative venture formation as SOM. These networks are based on competitive learning; the output neurons of the network compete among themselves to be activated or fired, with the result that only one output neuron, or one neuron per group, is “on” at any point in time. These winning neurons in essence represent the underlying clusters inherent in the original dataset. In a SOM, the neurons are placed at the nodes of a lattice that is usually one- or two-dimensional. Higher-dimensional maps are possible but not as common. The neurons become selectively tuned to various input patterns (stimuli) or classes of input patterns in the course of a competitive learning process. The location of the neuron so tuned (i.e., the winning neurons) becomes ordered with respect to others in such a way that a meaningful coordinate system for different input features is created over the lattice (Kohonen, 1990). A SOM is therefore a topographic map of the input patterns in which the spatial locations (i.e., coordinates) of the neurons in the lattice indicate intrinsic statistical features contained in the input data (Haykin, 1999).

3.5. Neural network specifications

We use SOM_PAK neural network package, which is widely used for analyzing SOM. This freely available package (www.cis.hut.fi/research/som_pak) relates quite closely to the classic SOM algorithm (Kohonen, 1995). The network used in this paper is comprised of 64 neurons, arranged in a two-dimensional array of nodes (x dimensions: 8 and y dimensions: 8). The data can be arranged in rectangular or hexagonal topology. We adopt the more commonly used hexagonal topology (Kohonen, Hynninen, Kangas, & Laaksonen, 1995). The weights associated with these 64 neurons capture the pattern underlying the overall data sample that is being studied. In line with the prescription of splitting the data into equal parts to form the training and testing datasets, the dataset
pertaining to 1985–1995 is used as a training data and that pertaining to 1996–2001 is used as calibration and testing data. A Gaussian learning function is used for training the 64 nodes associated with the neurons (details regarding the learning algorithm can be obtained from Kohonen et al., 1995; Haykin, 1999).

The SOM algorithm goes through various stages as described below:

**Initialization:** The reference vectors (corresponding to the 64 neurons) are first initialized to tentative values. This step also initializes the topology and neighborhood type chosen for analysis. The map is initialized by using a random number.

**Training:** After initializing the map, it is trained by using the training data. Training is done in two phases. In the first phase, the reference vectors of the map are ordered (ordering phase). During the second phase these ordered vectors are fine-tuned. In the beginning the neighborhood radius is taken almost equal to the diameter of the map and decreases to one during training, while the learning rate decreases to zero. Meanwhile, during the second phase the reference vectors in each unit converge to their correct values. The learning rate is decreased to a lower value in this stage. The neighborhood radius is also decreased. The values used for learning rate and the neighborhood radius in the two phases are presented in Table 3. In the first part we use 100,000 iterations, a learning rate of 0.05, and the radius between the neural nodes as 10. Subsequently, as suggested in Kohonen et al. (1995), in the second part we use 100,000 iterations but reduce the training rate and radius to 0.02 and 3, respectively.

**Evaluation of quantization error:** SOM quantize the data space into finite number of code vectors. The squared distance between an observed data and its corresponding code vector is called the quantization error. In essence, the quantization error presents the quality of the resulting map. The quantization error values do not have a prescribed cut-off and are dependent on the specific context in which SOM are used. The SOM_PAK package has an in-built algorithm that searches for good mappings that has lowest possible quantization error, by automatically repeating tests with different random initializing and training procedures. The testing data file is used to determine the quantization error associated with neural network training. The program computes the quantization error over all the samples in the data file and tries to find the map that result in lowest quantization error. This two-phase SOM training and the validation (i.e. evaluation of quantization error) is administered for 100 trial runs.

**Map visualization:** After the 100 trials the neuron weights are obtained. These neuron weights (also called codebook vectors) are calibrated by using the calibration data before we obtain the visualization maps. The cluster maps are generated by using visualization packages provided in SOM_PAK. We present the unified matrix representation (U-matrix) of the SOM that visualizes the distances between the neurons. The distance between the adjacent neurons is calculated and presented with different colorings between the adjacent nodes. A dark coloring between the neurons corresponds to a large distance and thus a gap between the codebook values in the input space. A light coloring between the neurons signifies that the codebook vectors are close to each other in the input space. Light areas can be thought as clusters and dark areas as cluster separators. This can be a helpful presentation when one tries to find clusters in the input data without having any a-priori information about the clusters.

A summary of the SOM algorithm utilized in this paper is presented in Appendix A. Further details can be obtained from Kohonen et al. (1995) and Haykin (1999). Table 3
Table 3
Neural network specifications for self-organizing map

<table>
<thead>
<tr>
<th>Specification</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of trials</td>
<td>100</td>
</tr>
<tr>
<td>Topology</td>
<td>Hexagonal</td>
</tr>
<tr>
<td>Neighborhood</td>
<td>Gaussian</td>
</tr>
<tr>
<td>Number of neurons</td>
<td>64 (X dimensions: 8; Y dimension: 8)</td>
</tr>
<tr>
<td>Training sample size (for industry</td>
<td>36,034</td>
</tr>
<tr>
<td>relatedness)</td>
<td></td>
</tr>
<tr>
<td>Testing sample size (for industry</td>
<td>44,542 (Note: We use the first 8508 cases as calibration data and the</td>
</tr>
<tr>
<td>relatedness)</td>
<td>remaining 36,034 cases as the testing data)</td>
</tr>
<tr>
<td>Data dimensionality</td>
<td>6 (Partner 1 &amp; 2’s industry and country affiliations; Collaborative</td>
</tr>
<tr>
<td></td>
<td>venture’s industry and country affiliation)</td>
</tr>
</tbody>
</table>

First part
- Iterations: 100,000
- Learning rate: 0.05
- Radius: 10

Second part
- Iterations: 100,000
- Learning rate: 0.02
- Radius: 3

Quantization error = 4.409611

presents the neural network specifications used in this paper, the sample sizes used in SOM training algorithm, and the resulting quantization error.

4. Results

The neural networks are modeled as SOM that reduce the dimensionality of the input data by recording the resulting weights as output. In the present study, the input has at least 30,000 cases. The advantage of SOM is their ability to capture the underlying patterns in the weights of the neurons. From the $U$-matrix visualization of the SOM, we can understand the complex patterns in collaborative venture formation and the underlying clusters in the data. Two visualizations of the $U$-matrix are presented.

(i) A characterization of the clusters in terms of relatedness of industries among the two partners and the collaborative venture.
(ii) A characterization of the clusters in terms of relatedness of home country of the two partners and the collaborative venture.

4.1. Industry effects

The $U$-matrix representation in Fig. 2 highlights the collaborative ventures and explicates the specific industries to which the two partners and the collaborative venture belong.

Overall, there are seven distinct clusters in the map. None of the identified clusters have code vectors with the partnering firms and the collaborative ventures all belonging to
distinct industries. From Fig. 2 it can be observed that none of the clusters have all firms from the same industry sector. However, in all clusters at least two firms (among the two partners and the collaborative venture) belong to the same industry. On further examination, we observe that in cluster 6 the two collaborating firms belong to the same industry. In the remaining clusters (1, 2, 3, 4, 5, 7), the collaborative venture is formed in the industry of one of the partners. Cluster 1 comprises of collaborative ventures that are in the services industry. Among the code vectors, two have partner firms belonging to the same industry group (high-technology industry), the remaining code vectors have one of the partnering firms in services the other representing either high-technology projects or discrete manufacturing industry group. In cluster 2, the collaborative ventures belong to projects and services industry group. The collaborative ventures in the services industry are formed either by the partners who are both in services industry or when one of the partnering firm is in services industry and the other is in the projects industry group. The collaborative ventures in the projects industry are similarly formed either when both partners are from projects industry or when one of the partners is from projects industry.
and the other is from the financial industry group. Cluster 3 represents collaborative ventures in the services and natural resources industry group. While the collaborative venture in the natural resources industry group is formed when both partners are from natural resources industry group; the collaborative venture in the services industry group is formed either when the two partners are from financials industry group or when one is from financial industry group and the other is from the services industry group. In cluster 4, the collaborative ventures belong to projects, high technology and process manufacturing industry groups. In the collaborative venture in projects industry group, either the two partners are from services industry group or one is from services and the other is from financial industry group. High-technology collaborative ventures comprise of two partners either both from financials industry group or both from services industry group. High-technology collaborative ventures are also formed when one partner is from services and the other is from either financials industry group or high-technology industry group. The process manufacturing collaborative ventures in this cluster are formed when one of the collaborating partners is from process manufacturing industry group and the other is from either trade or natural resources industry group. In cluster 5, the collaborative ventures are formed in high technology or discrete manufacturing industry groups. The collaborative venture in high-technology industry group comprise of one collaborating partner from high technology and the other from either financial or services industry group; collaborative venture in discrete manufacturing is formed when one of the collaborating partners is in discrete manufacturing industry group and the other is in the trade industry group. In cluster 6, the collaborating partners come from the same industry group and form a collaborative venture in the same industry group. The cluster consists of process manufacturing and high-technology industry groups. Finally, in cluster 7 the collaborative ventures in process manufacturing and discrete manufacturing are formed when both the collaborating partners also belong to these industry groups; meanwhile the collaborative venture in services industry group is formed by the collaborating partners in discrete manufacturing and services industry groups. We observe that industry groups—consumer goods (industry category code 2), miscellaneous manufacturing (industry category code 6) and other (industry category code 11)—are not captured in any of the code vectors suggesting that a cogent underlying structure involving these industries could not be established. Fig. 3 presents the characterization of clusters based on the level of industry relatedness.

4.2. Home country effects

Fig. 4 present the clusters formed by using the home country relatedness data of collaborating partners and the collaborative ventures. The figure highlights specific home countries that fall in each of the clusters.

Similar to the observation in industry effects, none of the clusters have collaborating partners and collaborative venture, all belonging to different home countries. However, unlike in industry effects analysis, clusters 2 and 6 have all code vectors with the collaborating partners and the collaborative ventures belonging to the same industry. In cluster 1, the partners are from North America or Japan and the collaborative venture is formed in North America, Japan or is Supranational. Cluster 2 comprises of partnering firms from Australasia, North Asia and North America forming a collaborative venture in their respective regions. In cluster 3, one code vector has collaborating partners from
North America and South America with the resulting collaborative venture being formed in South America. The remaining code vectors comprise of collaborative ventures that are formed in the country of origin of the collaborating partners (i.e. Australasia, North America and North Asia). In cluster 4, similarly, except for two code vectors in which the collaborating partners from Japan form a collaborative venture in North Asia and in which collaborating partners from Australasia and North America form a collaborative venture in Australasia, other code vectors are characterized by the collaborating partners and the collaborative venture belonging to the same country (Australasia, North America and Japan). In cluster 5, one-code vector represent a collaborative venture formed in Australasia that is formed by collaborating partners in North America and Australasia. The other two collaborative ventures are formed in North America and Japan by collaborating partners in the same countries. In cluster 6, the code vectors are characterized by collaborating partners and the collaborative venture all belonging to the same country (North America, Japan and North Asia). Finally, in cluster 7, firms in Western Europe collaborate with firms in North America, North Asia and South America to form collaborative ventures in North America, Western Europe, North Asia and South America. In Fig. 5, the characterization of clusters based on home country of the partnering firms and the collaborative venture is presented.

A summary of the overall integrated information obtained regarding industry effects and the home country effects is presented in Table 4.
Overall, the relatedness could be associated with similarities in partner profiles. Being from the same industry implies that the partners operate within the same competitive environment, with the same customers, distributors and suppliers. Moreover, it is reasonable to expect that firms from the same industry would have similar key success factors. It is probable that industry relatedness translates into similar value chains, comparable technological portfolios, similarity in manufacturing facilities and distribution channels. Meanwhile, relatedness in terms of home country is rooted in similarities of cultural, socio-economic and other backgrounds of the partners. The related alliances are characterized by resources brought by allies that are similar and potentially substitutable. Collaborative venture aims to accumulate identical resources to achieve scale efficiencies and to enhance market power. Moreover, by forming an alliance, the risk associated with the business is also shared by the partners. This type of synergy underlies collaborative venture formations in Cluster 2. Heritage Train Company provides an example of collaborations in this cluster. Formed in 1999, the company is the alliance between two Australian firms, Queensland Rail and Venice Simplon Orient Express. Understanding
their customers and business environments in Australian market, the two companies have successfully revived the grand style of luxury train travel in the local market through Heritage Train Company.

It is important to note that collaborative venture formation between firms sharing related profiles also has some drawbacks. The absence of differentiation limits the learning process. When two firms have related technology portfolios, similar organizational know-how, comparable managerial processes, associated with industry relatedness; or when they are governed by similar cultural and socio-economic and geo-political environments, the alliance offers limited opportunities for learning. Moreover, a collaboration in a related industry and country also carries with it a risk of inertia that inhibits future risk-taking behavior and engaging in diversification that takes advantage of available opportunities. There is also the risk of opportunism, lock-in, and potential competition between collaborative venture partners.

Companies sharing related profiles in terms of country may therefore opt to avert such inertia through industry diversification. Neural network recognizes three patterns of industry diversifications in which alliances operate in the industry that is different from the industry of both or at least one of the partners. Clusters 1 and 7 demonstrate industry diversification pattern in which manufacturing firms diversify to the services industry through alliance units. For example, in July 1998 a US software manufacturer Cincom Systems, Inc. formed a joint marketing alliance with a US-based telecommunication manufacturer Bohemia to form Periphonics, a joint collaboration designed to enter service industry. Cluster 3 shows industry diversification pattern in which firms in financial industry moving into services industry through alliances. For instance, an Asian financial firm Delta Asia formed a collaborative venture with a Hong Kong-based
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Code vectors</th>
<th>Partner 1 nation</th>
<th>Partner 2 nation</th>
<th>CV nation</th>
<th>Partner 1 industry</th>
<th>Partner 2 industry</th>
<th>CV industry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>High technology</td>
<td>High technology</td>
<td>Services</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Services</td>
<td>Discrete manufacturing</td>
<td>Services</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>Supranational</td>
<td>High technology</td>
<td>High technology</td>
<td>Services</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>High technology</td>
<td>Services</td>
<td>Services</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>Supranational</td>
<td>Projects</td>
<td>Services</td>
<td>Services</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
<td>High technology</td>
<td>Services</td>
<td>Services</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Services</td>
<td>High technology</td>
<td>Services</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>Australasia</td>
<td>Australasia</td>
<td>Australasia</td>
<td>Services</td>
<td>Services</td>
<td>Services</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>North Asia</td>
<td>North Asia</td>
<td>North Asia</td>
<td>Services</td>
<td>Projects</td>
<td>Services</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Projects</td>
<td>Projects</td>
<td>Projects</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Projects</td>
<td>Services</td>
<td>Services</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>North Asia</td>
<td>North Asia</td>
<td>North Asia</td>
<td>Projects</td>
<td>Financial</td>
<td>Projects</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Australasia</td>
<td>Australasia</td>
<td>Australasia</td>
<td>Projects</td>
<td>Financial</td>
<td>Projects</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>Australasia</td>
<td>Australasia</td>
<td>Australasia</td>
<td>Services</td>
<td>Financial</td>
<td>Services</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Financial</td>
<td>Financial</td>
<td>Services</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>North Asia</td>
<td>North Asia</td>
<td>North Asia</td>
<td>Services</td>
<td>Financial</td>
<td>Services</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>Australasia</td>
<td>Australasia</td>
<td>Australasia</td>
<td>Natural resources</td>
<td>Natural resources</td>
<td>Natural resources</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Natural resources</td>
<td>Natural resources</td>
<td>Natural resources</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>North America</td>
<td>South America</td>
<td>South America</td>
<td>Natural resources</td>
<td>Natural resources</td>
<td>Natural resources</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>Australasia</td>
<td>Australasia</td>
<td>Australasia</td>
<td>Services</td>
<td>Services</td>
<td>Projects</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>Australasia</td>
<td>Australasia</td>
<td>Australasia</td>
<td>Financial</td>
<td>Financial</td>
<td>High technology</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Services</td>
<td>Services</td>
<td>High technology</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Services</td>
<td>Projects</td>
<td>Projects</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
<td>Services</td>
<td>Financial</td>
<td>High technology</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>Japan</td>
<td>Japan</td>
<td>North Asia</td>
<td>Process manufacturing</td>
<td>Trade</td>
<td>Process manufacturing</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>Australasia</td>
<td>North America</td>
<td>Australasia</td>
<td>Services</td>
<td>Services</td>
<td>High technology</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Services</td>
<td>High technology</td>
<td>High technology</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>Natural resources</td>
<td>Process manufacturing</td>
<td>Process manufacturing</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>North America</td>
<td>Australasia</td>
<td>Australasia</td>
<td>High technology</td>
<td>Financial</td>
<td>High technology</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>North America</td>
<td>North America</td>
<td>North America</td>
<td>High technology</td>
<td>Services</td>
<td>High technology</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
<td>Discrete manufacturing</td>
<td>Trade</td>
<td>Discrete manufacturing</td>
</tr>
</tbody>
</table>
Table 4 (continued)

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Code vectors</th>
<th>Partner 1 nation</th>
<th>Partner 2 nation</th>
<th>CV nation&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Partner 1 industry</th>
<th>Partner 2 industry</th>
<th>CV industry&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>33</td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
<td>High technology</td>
<td>High technology</td>
<td>High technology</td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>North Asia</td>
<td>North Asia</td>
<td>North Asia</td>
<td>High technology</td>
<td>High technology</td>
<td>High technology</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>Western Europe</td>
<td>North America</td>
<td>Western Europe</td>
<td>Discrete manufacturing</td>
<td>Services</td>
<td>Services</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>Western Europe</td>
<td>North Asia</td>
<td>North Asia</td>
<td>Discrete manufacturing</td>
<td>Discrete manufacturing</td>
<td>Discrete manufacturing</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>Western Europe</td>
<td>South America</td>
<td>South America</td>
<td>Discrete manufacturing</td>
<td>Discrete manufacturing</td>
<td>Discrete manufacturing</td>
</tr>
</tbody>
</table>

<sup>a</sup>CV nation = collaborative venture nation.

<sup>b</sup>CV industry = collaborative venture industry.
telecommunications holding SmarTone in November 2000 as a way to expand its position in the telecommunications service sector. In addition, an industry diversification is probable as traditional services or financial firms moved into high technology or projects sectors, through alliances. This diversification pattern is captured in Cluster 6. As an example, in October 1999, Deloitte & Touche Consulting harnessed its business-application services to Sprint’s global communications networks to offer network services, data-center operations that ranged from electronic commerce to enterprise applications. The alliance thus enabled Deloitte Consulting and Sprint’s to diversify from the traditional service firms to high-technology sector.

The drawbacks of absence of differentiation may also necessitate collaborative venture formation among firms with unrelated resource profiles. Firms from unrelated industries typically do not have same suppliers, distributors and customers. Collaborative venture enables the partners to take advantage of the complementary effects of completely different value chains, different technologies, different manufacturing facilities and different expertise at the commercialization stage. It also allows firms to take advantage of the partner’s knowledge of prevailing economic, market, legal, cultural and political situation in a country to enable a smooth entry into the market. Collaborative ventures between firms from unrelated industries and countries provide opportunities of learning for the partners. Clusters 3, 4, 5 and 7 provide examples of situations in which the partnering firms are from different industries and different countries.

The learning can occur in diverse areas like, industrial know-how, technological assets, organizational abilities, commercial and marketing competencies, cultural issues, and market profiles (Hamel, 1991; Kogut, 1988; Mowery et al., 1996; Richter & Vettel, 1995). Meanwhile, it is important to manage the procedural and cultural gaps that exist due to unrelated industry and home countries of the partnering firms. Cluster 5 captures the underlying synergy related to firms in non-manufacturing sectors diversifying into manufacturing sectors through collaborations with unrelated country partner.

It is intuitive that the specific motive for collaborative venture formation is likely to have an impact on the partner selection process. The valuation of the capabilities of a potential partner is based on this initial motive. As an example, if the objective of alliance formation is to reduce costs by achieving economies of scale, selection criteria associated with access to customers, materials, or core knowledge may be most important. It can be argued that in this situation the partnering firms will be related in terms of industry. A collaborative venture from Cluster 6 provides an example for this type of collaboration. In November 2000, a JV “Gigaphoton” was formed between Ushio and Komatsu, both of Tokyo, Japan, to produce a new argon fluoride (ArF) excimer laser. With the combined knowledge from the two parent firms, the collaborative venture was expected to be able to launch the new product and gain as much as 50% of the world market within less than 1 year. The partnering firms were therefore able to gain from the synergy arising from combining their similar critical knowledge.

If, on the other hand, the reason for collaborative venture formation is to access global resources or market entry, partner selection criteria will include the partner’s natural resources or knowledge of the local market. In this situation complementary resources of the partnering firm from an unrelated country become an important determinant of collaborative venture formation. Cluster 3 reflects this type of underlying synergy between partners and is exemplified by the extra-heavy crude oil project—a collaboration between Exxon Corporation of the United States and Corpoven of Venezuela. Formed in 1997, the
joint project enables the US partnering firm to access critical natural resources that would be difficult to do so otherwise.

5.1. Further exploration of code vectors underlying clusters

The central focus of this research is to explore the pattern in collaborative ventures formed during 1985–2001 period by examining the relatedness and unrelatedness of partnering firms and the collaborative venture in terms of industry and home countries. We observe a complex pattern in alliance formations. In this section we elaborate on the findings by presenting taxonomy of collaborative ventures. In the taxonomy, we classify the codebook vectors into categories formed by considering the industry and home country relatedness continuum.

The patterns recognized by neural network provide an insight into different types of synergies underlying collaborative venture formations. Due to the changes in the imperfections generated in markets and in firms, firms concentrate and reassess their diversification strategies (Pearce, 1993). Consequently, it is often very difficult to classify firms’ strategies over long periods of time as being only of related or unrelated diversification. Nevertheless, despite the unique ways through which firms respond to those imperfections (Nelson, 1991), it is possible to find common patterns in their diversification strategies. Often, these are characterized in terms of double diversification such that firms diversify into unrelated industries and countries (Pearce, 1993). In Fig. 6, we present taxonomy of varying levels of relatedness that would aid in further enhancement of knowledge and practice in the area of collaborative venture formation. Specifically, the taxonomy presents 16 categories in which collaborative ventures formed by two collaborating partners can be placed. These categories are formed by considering the varying levels of relatedness in terms of industry as well as home country of the collaborating partners and the collaborative venture. The categories span the continuum of

<table>
<thead>
<tr>
<th>Relatedness of Home Country</th>
<th>Collaborating partners and the collaborative venture are from the same home country</th>
<th>Collaborating partners are from the same industry group but the collaborative venture is in a different home country</th>
<th>One collaborating partner and the collaborative venture are from the same home country but the other collaborating partner is from a different home country</th>
<th>Collaborating partners and the collaborative venture are all from different home countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborating partners and the collaborative venture are from the same industry group</td>
<td>Code Vectors: 8, 10, 12, 18, 22, 33, 34 (Category I)</td>
<td>Code Vector: 26 (Category VIII)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaborating partners are from the same industry group but the collaborative venture is in a different industry group</td>
<td>Code Vectors: 1, 15, 20, 21, 22 (Category II)</td>
<td>Code Vector: 3 (Category V)</td>
<td>Code Vector: 25, 5 (Category VI)</td>
<td></td>
</tr>
<tr>
<td>One collaborating partner and the collaborative venture are from the same industry group but the other collaborating partner is from a different industry group</td>
<td>Code Vectors: 2, 4, 6, 7, 9, 11, 12, 13, 14, 16, 23, 27, 28, 30, 31 (Category III)</td>
<td>Code Vectors: 29, 36, (Category IX)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collaborating partners and the collaborative venture are all from different industry groups</td>
<td>Code Vector: 24 (Category IV)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 6. Taxonomy of collaborative ventures based on industry and home country relatedness.
collaborative ventures that are completely related and those that are completely unrelated. We place the codebook vectors underlying the clusters identified in our study into suitable categories within the taxonomy.

In total, we have 38 code vectors that have been categorized based on the relatedness of the collaborating partners and collaborative venture. We observe that the majority of code vectors (about 39%) belong to the III category within the taxonomy. The category represents a very high degree of home country relatedness among collaborating partners and the collaborative venture. However, within the continuum of relatedness in terms of industry, it was observed that the pattern is characterized by collaborating partners who are from distinct industries but who form a collaborative venture within the industry group of one of the partners. Category I has the second highest number of code vectors (about 18%) suggesting that a fairly large number of collaborative ventures are formed with very high degree of relatedness on both industry and home country dimensions. The next highest number of code vectors belongs to category II (about 13%). This category is characterized in terms of a very high degree of relatedness in terms of home country; in these collaborative ventures the partners belong to the same industry but the collaborative venture is formed in an unrelated industry group. In Category VII, there is a very high degree of relatedness in terms of industry; however, the collaborating partners belong to different home countries that form a collaborative venture in the home country of one of the collaborating partners. This category represents the fourth most common pattern (about 11%) observed in collaborative ventures. Categories VI and IX have about 5% code vectors. In category VI, collaborating partners come from different industries and the collaborative venture is formed in the industry group of one of the collaborating partners. In this category collaborating partners come from the same home country but form a collaborative venture in an unrelated home country. Category IX is similar to Category VI when viewed from the perspective of industry relatedness. However, in Category IX the collaborating partners come from unrelated home countries and form a collaborative venture in the home country of one of the partner. Categories IV, V and VIII, each contain one code vector. Categories V and VIII have collaborating partners from the same industry group but the collaborative venture is formed in an unrelated industry. From the perspective of home country relatedness, Category V has collaborating partners from the same home country but the collaborative venture is formed in an unrelated home country. Meanwhile, Category VIII comprise of collaborating partners from unrelated home countries who form a collaborative venture in the home country of one of the collaborating partners. Category IV represents a pattern in collaborative ventures that are characterized in terms of highly related geographical region but unrelated industry groups. We did not find dominant patterns in the gray-shaded cells within the taxonomy. The gray-shaded cells provide some interesting insights. First, the gray-shaded column for the highest degree of unrelatedness in terms of home country suggests that collaborative venture formations that are characterized in terms of complete unrelatedness among the home countries of the collaborating partners and the collaborative venture is very uncommon. Second, the row representing highest degree of unrelatedness in terms of industry suggest that collaborative venture formations with completely unrelated industry groups of the collaborating partners and the collaborative venture was observed only in situations when the collaborating partners come from the same home country and form the collaborative venture within the region. Intriguingly, we find an empty cell within the taxonomy that is characterized in terms of
completely related industry group of the collaborating partners and the collaborative venture; when viewed from the perspective of home country this cell comprise of collaborating partners from the same home country who form a collaborative venture is an unrelated home country.

The taxonomy expands our understanding of collaborative venture formation when viewed from the relatedness continuum of industry groups and home countries. It presents several avenues for further theoretical and empirical investigations. Theoretical studies could provide deeper understanding of the classes of related and unrelated collaborative ventures. Further studies are needed to examine why some categories present a dominant pattern of collaborative venture formation while others fail to find adequate representation. Empirical investigations could subsequently validate the relationships among the various facets underlying the collaborative venture formations.

6. Conclusion

The present study advances our knowledge of collaborative venture formation in two important ways. First, we provide an empirical investigation of a large dataset of collaborative venture formation for an extended period of time to unravel the underlying patterns. We present self-organizing maps as an approach to examine these patterns and delineate the clusters comprising of code vectors that are strongly related. We explain the underlying characteristics of these clusters and present some insights regarding collaborative venture formation during the 1985–2001 time period. Second, we present taxonomy, which explicitly considers the continuum of relatedness when viewed from the perspectives of industry groups and home countries. The taxonomy provides an understanding of the overall pattern in which collaborative ventures were formed by grouping the code vectors into appropriate categories.

There are a few limitations of this study. First, although we examined data from more than 65 industries and 200 countries, we had to keep the input data at a manageable level by grouping industries and countries into categories. This introduced some degree of subjectivity into the analysis. Second, the research offers a novel approach for recognizing industry and home country patterns, but it lacks predictive ability. Third, the data consists mostly of publicly traded companies hence the results cannot be generalized for privately held corporations. Nevertheless, the extensive dataset and the exploratory nature of the study fulfills its task by pointing to the patterns that underlie collaborative ventures for a large number of organizations.

In the future, neural network methods can be expanded to cover pattern recognition for collaborations involving more than two partners. Moreover, several other possible antecedents for collaboration should be explored, such as specific management styles, prior alliance experience, and political and macroeconomic conditions in partners’ respective countries (Nath, 1988). In addition, researchers can incorporate predictive features by using such neural network techniques as feed-forward networks to predict the success of alliances under various contingencies. The results reported in this paper present an initial step towards a conceptual framework to understand collaborative ventures based on the relatedness dimension. The framework can be refined in future studies to incorporate richer measures of relatedness that extend beyond industry and home country relatedness.
Appendix A

A.1. Self-organizing map (SOM) algorithm

The SOM algorithm is unique in that it combines the goals of the projection and clustering algorithms. It can be used at the same time to visualize the clusters in a dataset and to represent the set on a two-dimensional map in a manner that preserves the complex nonlinear relations of the data items; nearby items are located close to each other on the map. Moreover, even if no explicit clusters exist, “ridges” and “ravines” are revealed. The former are open zones with irregular shapes and a high clustering tendency, whereas the latter separate datasets that have a different statistical nature.

The essence of Kohonen’s (1990) SOM algorithm is that it substitutes a simple geometric computation for the more detailed properties of the Hebb-like rule and lateral interactions. The essential ingredients/parameters of the algorithm are:

- A continuous input space of activation patterns that are generated in accordance with a certain probability distribution. We use the specific characteristics in a JV formation as our input space.
- A network topology in the form of a lattice of neurons, which defines a discrete output space.
- A time-varying neighborhood function, \( h_{j,i}(x)(n) \), which is defined around a winning neuron \( i(x) \).
- A learning-rate parameter \( \eta(n) \) that starts at an initial value (\( \eta_0 \)) and then decreases gradually with time, \( n \), but never goes to zero.

For the neighborhood function we use the expression:

\[
    h_{j,i}(x)(n) = \exp \left( -\frac{d_{j,i}^2}{2\sigma^2(n)} \right), \quad n = 0, 1, 2, \ldots
\]  

(A.1)

For cooperation among neighborhood neurons to hold, it is necessary that topological neighborhood \( h_{j,i} \) be dependent on lateral distance \( d_{j,i} \) between winning neuron \( i \) and excited neuron \( j \) in the output space, rather than on some distance measure in the original input space.

There are three basic steps involved in the application of the algorithm after initialization: sampling, similarity matching and updating. These steps are repeated until the feature map is completed. The algorithm is summarized as follows.

1. **Initialization**: Choose random values for the initial weight vectors \( w_j(0) \). The only restriction here is that the \( w_j(0) \) be different for \( j = 1, 2, \ldots, l \), where \( l \) is the number of neurons in the lattice. It may be desirable to keep the magnitude of the weights small. Another way to initialize the algorithm is to select the weight vectors \( \{w_j(0)\}_{j=1}^{l} \) from the available set of input vectors \( \{x_i\}_{i=1}^{N} \) in a random manner.
2. **Sampling**: Draw a sample \( x \) from the input space with a certain probability; the vector \( x \) represents the activation pattern that is applied to the lattice. The dimension of vector \( x \) is equal to \( m \).
(3) **Similarity matching**: Find the best matching (winning) neuron \( i(x) \) at time step \( n \) by using the minimum-distance Euclidian criterion:

\[
i(x) = \arg \min_j ||x(n) - w_j||, \quad j = 1, 2, \ldots, l.
\]

(4) **Updating**: Adjust the synaptic weight vectors of all neurons by using the update formula:

\[
w_j(n + 1) = w_j(n) + \eta(n)h_{j,i(x)}(n)(x(n) - w_j(n)),
\]

where \( \eta(n) \) is the learning-rate parameter and \( h_{j,i(x)}(n) \) is the neighborhood function centered around the winning neuron \( i(x) \);

(5) **Continuation**: Continue with step 2 until no noticeable changes in the feature map are observed.

**References**


