Annals of Operations Research 87(1999)213–232

Estimating the performance of Sino–Hong Kong joint ventures using neural network ensembles

Michael Y. Hu, Ming S. Hung, B. Eddy Patuwo and Murali S. Shanker
College of Business Administration, Kent State University, Kent, OH 44242-0001, USA
E-mail: epatuwo@bsa3.kent.edu

Prediction of the performance of international joint ventures remains a relatively under-researched area, yet its importance is well recognized due to the tremendous surge in joint venture activities in the past decade. Data on 1,463 Sino–Hong Kong joint ventures were gathered and those that appeared on the Honor Roll of the China Association of Enterprises with Foreign Investment are identified. Neural network models were used to relate the posterior probability – the probability that a venture gets on the Honor Roll – and the seven economic variables. A simple and yet powerful method called the ensemble method was used to estimate the posterior probability. The results indicate that every variable, except one, has influence on the probability of success. Perhaps more important, the results demonstrate that the modeling approach is able to mine useful information from the data set.

1. Introduction

It has been argued that for multinationals to venture into other countries, the returns must be high enough to compensate for the additional costs incurred in dealing with their new environment. At the same time, these venturing firms should also possess specific advantages for them to invest abroad rather than at home. Particularly in the case of Hong Kong, investors will be likely to take advantage of the abundance of low-cost labor, land and materials in China. The tremendous surge of joint venture activities in China, particularly in the past ten years, allows the opportunity for such investigation.

How well a business performs can be measured in many ways. This paper employs artificial neural networks to model the factors affecting one such measure of Hong Kong joint ventures in China. It constitutes the first large-scale economic analysis of performance in this area. Researchers are constantly seeking better ways to model the economic determinants of performance. In this case study, we use neural networks to approximate the probability that a venture will be successful (according to the measure employed) by estimating the relationship between this probability and the economic factors associated with the venture. Managerial implications of such a relationship can then be studied more closely.

The next section provides the rationale for studying the performance of Sino–Hong Kong joint ventures and a description of the data used in this study. Section 3 introduces statistical classification, posterior probability, and how neural networks can be used to estimate the posterior probability directly. A simple and yet powerful approach called the ensemble method is explained. Model building using neural networks is described in section 4, and section 5 discusses the results along with economic interpretations. Finally, section 6 offers some concluding remarks.

2. International joint ventures

International joint ventures are formed to create corporate synergism. Local partners, particularly from developing countries, would benefit from the technological know-how, management skills, and capital brought in by their foreign partners. Multinationals can benefit from their local partners' knowledge of the marketplace and position in the existing distribution network. Economies of scale can be achieved through trust and cooperation among the partners (Contractor and Lorange [5]). While improving efficiency through combining resources, joint ventures have also created management problems through different goals, values and cultures. The existing literature on joint venture performance indicates the success rate of international joint ventures has been dismal, ranging from 30% to 70% (Geringer and Hebert [11]). Prior studies focused their attention on the performance of joint ventures in developed countries. As more joint ventures are formed in developing countries, it is important to investigate the performance of these ventures.

In the past decade, foreign investments in China have increased significantly. Previous studies rely on interviews and surveys, securing information from foreign investors on whether or not they are satisfied with their operations in China [6,30, 31,33]. After successfully gathering a large-sample data set from reliable sources, we fill the void in the literature by conducting the first large-scale study on the performance of Sino–foreign joint ventures (Hu et al. [15]). Multinationals extend substantial resources in the formation and implementation of these ventures. The profit potentials are huge. They have an obvious interest in accurately predicting the success of these ventures. Even a moderate improvement in model prediction will be of importance to these investors.

2.1. *Performance factors*

A unified theory explaining the performance of international joint ventures has yet to be developed. Prior literature has found that partner commitment, control exercised by the foreign partners, number of partners, product characteristics and host country environment, have a major influence on the success of an international joint venture.

2.1.1. Partner commitment

A joint venture cannot exist if the partners are not committed to the shared objectives as prescribed in the agreement. A lack of commitment leads to an ill-defined set of corporate objectives. As a result, the whole organization will suffer. Killing [17] provided empirical evidence for the impact of commitment on performance using a sample of international joint ventures from developed countries.

Beamish and Banks [4] interviewed 46 senior executives from 66 international joint ventures in the manufacturing sector, located in the Caribbean economies. Most of the multinationals surveyed are from the US, Canada, and UK. Performance is defined as whether there is a mutual agreement between the partners on overall satisfaction. They concluded that commitment was positively related to the amount of contribution made by each partner; thus, greater contribution leads to higher level of performance.

2.1.2. Control

Control in an international joint venture can be defined as the process through which parent companies ensure that the way a joint venture is managed conforms to their interest [2,3,9,10,12,17,18,29,32]. Control is a complex concept that includes several dimensions: (1) mechanisms of control (equity ownership, representation in management bodies, technical superiority and management skills, etc.), (2) extent of control (whether one or more partners play an active role in decision making), and (3) focus of control (the scope of activities over which parents exercise control). These dimensions are complementary and interdependent. Research to date on control-performance has been inconclusive on the direction of the relationship.

Killing [17] proposes that dominant partner joint ventures tend to be more successful than shared management ventures. Shared ownership often inflicts managerial conflicts (Gomes-Casseres [12]), thus exerting a negative effect on performance. This line of reasoning is consistent with transaction cost theory. Dominant foreign ownership provides the mechanism for keeping transaction costs to the minimum (Hennart 13]). As a result, stability of an international joint venture can be better achieved. Killing found that dominant partner joint ventures were more successful than shared management ventures. However, the suggested relationship has not always been consistent with empirical evidence. Tomlinson [32] reported that when parents showed a more relaxed attitude to control over joint ventures, their level of profits was higher. Franko [9] also found that joint ventures were more stable when parents demanded less control. Beamish [3] reported that among twelve international joint ventures formed in less developed countries, performance was negatively related with the level of control from the foreign partner. Reynolds [28] in his investigation of US joint ventures in India concluded that control had not become an issue. In a study on joint ventures in developed countries, Kogut [18] found no relationship between control and performance.

In sum, the research to date has produced conflicting results due to the complex nature of the issue and lack of consensus on the measure of joint venture performance. Thus, the direction of the control–performance relationship has yet to be established, tested, and clarified in the existing literature (Geringer and Hebert [10]).

2.1.3. Number of partners

The problem of shared management is heightened as the number of partners to a joint venture increases. Schaan and Beamish [29] note that the task of managers in international joint ventures is complicated because they are often expected to accommodate the interests of two or more parties. Thus, the role theory would prescribe that as the number of parent firms increases, the rate of failure among international joint ventures will increase (Franko [9]). Shenkar and Zeira [30] studied the relationship between role conflict of chief executive officers in international joint ventures in Israel. Based on 44 in-depth interviews, they reported a surprising finding that role conflict is lower when the number of parent firms is high, and justified their findings through non-transferability of the role theory developed in the uninational setting to the multinational environment like that of the international joint ventures.

This expanded problem in shared management would have to be tempered with the benefits to be gained through coordination and cooperation (Gomes-Casseres [12]). Joint venture is not a viable channel for transferring a firm’s capabilities if there exist only costs without benefits. Corporate synergism as discussed earlier will provide benefits to forming a joint venture. The benefits from cooperation and coordination may outweigh the costs when the number of partners is small, thus yielding a positive relationship between the number of partners and performance. However, as the number of partners increases beyond a certain level, a negative relationship should persist. Thus, on the whole a curvilinear relationship is expected between the number of partners and level of performance.

2.1.4. Product characteristics

Kogut [19] studied the stability of 92 US–foreign joint ventures in the United States. Stability is measured in terms of the likelihood of termination. It is found that the amount of ties (whether the joint venture partners have other licensing agreements), R&D and scale intensity (percent of tangible assets primarily engaged in production) contribute positively to stability. High technology products due to product differentiation typically command a stronger position in local or foreign markets. These factors allow firms to charge a higher price and engage in non-price competition. Therefore, a positive relationship should persist between technological characteristics of a product and performance of foreign investment enterprises.

Lecraw [20] indicates that many of the multinational companies in his sample sold products that they have already imported in the finished form in addition to manufacturing products locally. Although their capital investments in developing countries may be small, these firms can earn high profits on their imports as well as their locally produced items. Products that are not capital intensive tend to be labor intensive. By producing labor intensive products, foreign investors can take advantage of the abundance of low cost labor in China. Therefore, it is proposed that the relationship between performance and capital intensity is negative, while it is positive between performance and labor intensity.

2.1.5. Host country environment

In a developing country like China, major business centers are more developed and better equipped to conduct business operations. In addition, the Chinese government has also established special economic, technological development and trade zones to foster the development of these activities. In these areas, tax rates are low and facilities are plentiful. It is envisioned that joint ventures should benefit from the better infrastructure and favorable tax structure. Thus, it is proposed that performance of foreign subsidiaries is positively related to the economic development and tax incentives provided by the Chinese government.

2.2. Sino–Hong Kong joint ventures

The country of origin of foreign investment has been found to be useful in explaining performance. This study focuses on the performance of joint ventures formed by Hong Kong investors in China. While most previous studies use multiple host countries because of the small number of foreign ventures in any one host country, by limiting the scope of our study to one host and home country, we eliminate the variations between countries from these two sources. In addition, since Hong Kong investors account for roughly 60% of foreign investment projects in China, limiting the scope of our study to Hong Kong will cause little loss of generality.

2.3. The data

The data for this study come from two sources: (1) Statement of Sino–Foreign Joint Ventures, 1979–1990 in Almanac of China's Foreign Economic Relations and Trade, and (2) Sino–foreign joint ventures honored for their outstanding performance by the China Association of Enterprises with Foreign Investment, a nonprofit organization for Sino–foreign joint ventures. The Statement contains a description of 3,071 Sino–foreign joint ventures, providing names of the ventures, their parent companies' names, geographic location in China, number of years the venture is to exist, amount of total investment, percent of foreign investment, and the primary scope of the business.

In 1989 and 1990, the China Association of Enterprises with Foreign Investment evaluated operating joint ventures and selected 1,187 outstanding manufacturing enterprises. To be included on the Honor Roll, a joint venture must (1) produce high-quality products, (2) have a before-tax profit exceeding one million yuans, and (3) export value greater than two million US dollars. This study uses those commended on the Honor Roll as successful joint ventures because they must simultaneously produce high-quality goods, profits, and export revenue. Profits, understandably, are important for a foreign joint venture. This measure was adopted by various researchers (Lecraw [20]). Producing high-quality products for export is also beneficial to all parties involved. For foreign partners, it is an indication that products produced in China are of good quality and competitive in international markets.

Of the 3,071 joint ventures, 2,416 fall in the manufacturing category. Matching the 2,416 joint ventures in the Statement and the 1,187 in the Honor Roll results in 220 joint ventures which received commendation in 1989 or 1990. Of the 2,416 ventures, 1,463 (60.6%) are with Hong Kong investors and of these, 140 are on the Honor Roll. The average investment of these ventures amounts to US$2.442 million, which is relatively small compared to the average of US$3.087 million for all Sino–foreign ventures.

Several observations can be made here. First, given the average size of the joint ventures, the criteria set forth by the China Association of Enterprises with Foreign Investment are unduly stringent. Thus, less than ten percent of the sample joint ventures are on the Honor Roll. Secondly, one can expect only a handful of ventures formed in 1989 or 1990 to be on the Honor Roll since they would have been in operation in China for only one or two years. Thirdly, businesses and joint ventures alike are likely to under-report their earnings. All these factors will lead one to conclude that the commended ventures are indeed "successful" in their business operations. Yet, among those that have failed to be on the Roll, many may have actually surpassed the success criteria at the time of commendation.

2.4. Variables

The dependent variable, performance of joint venture, is coded as 1 for "successful", those listed on the Honor Roll, and 0 otherwise. It will be denoted by SUCCESS.

Independent variables used in this study include level of commitment, foreign control, number of partners, industrial characteristics of joint venture products, and joint venture location in China.

- The level of commitment involved is measured using duration (number of years as specified in the joint venture contract), which is transformed by a natural logarithm and denoted by LDUR. A positive relationship is expected between performance and this variable.
- Percent foreign ownership (PFINVEST) is used as a proxy of the level of a foreign investor's control. A larger share of equity ownership would give a partner greater control of a joint venture.

- The number of partners is represented by its natural logarithm LPARTNER.
- In order to examine the product-related effects, the 2,416 manufacturing joint ventures are classified into high or low technology (TECH), and labor or capital intensive (CAP). This classification scheme, developed by Dunning [8] and later used by Lee [21], is based on the two-digit standard industrial classification (SIC) codes. By this scheme, (1) labor-intensive and low-technology industries include: flowers and plants, food and kindred products, tobacco, textile products, apparel products, lumber and wood products, furniture, paper, printing, rubber and plastics, leather, fabricated metal and miscellaneous manufacturing industries; (2) labor-intensive and high-technology industries include: nonelectric machinery, electrical and electronic products, and measuring equipment; (3) capital-intensive and low-technology industries include: stone, clay, glass and concrete products, and primary metal industries; (4) capital-intensive and high-technology industries include: chemicals and allied products, petroleum refining and transportation equipment.
- Economic risk due to differences in joint venture location is examined at the regional level (LOC). There are three regions for LOC. The first are areas with both economic and tax incentives. Corporate income tax rate for joint venture in China is 30%. A local surtax of 10% of this income tax is levied by the local government, thus bringing the effective tax rate to a total of 33%. In the Special Economic Zones and the Economic and Technology Development Zones in the 14 cities, this tax rate is reduced to 15% for foreign investors. LOC = 1 when a venture is in either special zone. The second are areas with economic but no tax incentives. The coastal regions in China are more economically developed, providing facilities for conducting business operations. A venture in a coastal area outside of the special zones has LOC = 2. The third region (LOC = 3) consists of inland provinces, typically lagging in economic development and offering no tax and economic incentives.
- It can be argued that it is more likely for joint ventures with large sums of investment to be included on the Honor Roll. Thus, the natural logarithm of the total amount of investment, LINVEST, will be included as an additional independent variable to remove variability in the dependent measure associated with the size of the venture.

The categorical variables are coded using indicator variables. A 1, –1 coding is used for each of the two industry related factors (TECH, CAP). TECH is coded as “1” (“–1”) for high (low) technology, and CAP as “1” (“–1”) for capital (labor) intensive. LOC is coded using two variables, (LO1, LO2), with (1, 0) denoting LOC = 1, (0, 1) for LOC = 2, and (–1, –1) for LOC = 3.

Given that the dependent variable in this study is categorical in nature, either successful or not successful, the problem evolves into one of classification. That is to say, the economic factors will be first used to predict the probability that a joint venture belongs to one or the other category, and then the resulting functional relationship will be used to estimate the impact of each economic factor on performance.

2.5. Previous studies

Hu and Chen [14] analyzed the data on 2,442 Sino–foreign joint ventures for the relationship between the economic variables and performance. Logistic regression was employed. Variables that have statistically significant (at the 0.01 level) positive influence on performance are LDUR, LPARTNER and LINVEST. Thus, as duration of the venture, the number of partners and the size of the joint venture increase, it is more likely for the venture to become more successful. The other variables, TECH, CAP and LOC are found to be insignificant.

Hu et al. [15] used neural networks to analyze this sample of 2,442 Sino–foreign joint ventures. The primary objective, however, was to assess the prediction accuracy in classification of successful/not-so-successful ventures. It was found that neural networks outperformed logistic regression. Logistic regression simply assigned all cases to the not-so-successful group, whereas neural networks were able to identify some successful ventures.

Patuwo et al. [25] did a similar study on 1,463 Sino–Hong Kong joint ventures, using the same data set as in the present study. Again, the primary objective was to evaluate the prediction accuracy of neural network models. Neural networks provided superior classification results to logistic regression. As in the previous study, logistic regression assigned all cases to the not-so-successful group. A procedure of backward elimination was developed to select the significant variables. The variables selected were LPARTNER, LINVEST, PFINVEST, CAP and TECH. Backward logistic regression retained only three of the five neural network variables, LPARTNER, LINVEST and TECH, as significant explanatory variables for performance.

2.6. Preliminary data analysis

In order to gain some ideas about the possible relationships between the economic factors of a joint venture and whether the venture was on the Honor Roll, scatter plots of the continuous variables are prepared. Figures 1–4 show the plots for variables LDUR, PFINVEST, LINVEST, and LPARTNER, respectively. Each plot contains all the 1,463 observations and, unfortunately, the pixels are not proportional to the number of duplicate points. From these figures, it is difficult to discern any clear or strong relationship between any variable and performance. Take LDUR, for example. The location and distribution of this variable for the successful ventures are quite similar to those of the unsuccessful ones. The linear regression function is overlaid, only to suggest the direction of relationship between the variables since the regression coefficients are not meaningful. For this example, the relationship between LDUR and

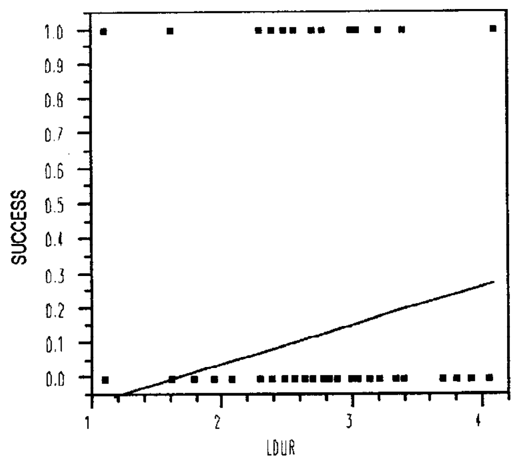

Figure 1. SUCCESS versus LDUR.

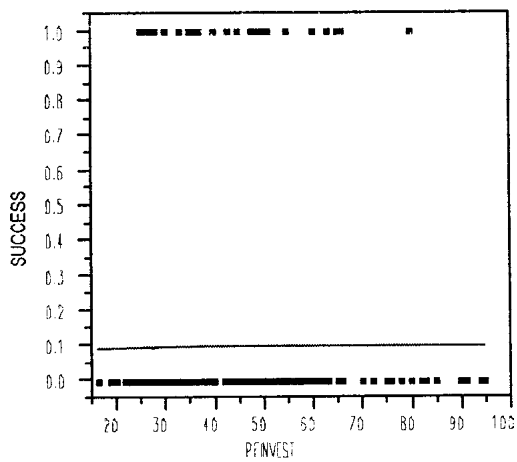

Figure 2. SUCCESS versus PFINVEST.

SUCCESS seems to be positive. Indeed, the relationship seems positive also for LINVEST, PFINVEST, and LPARTNER, although the line for PFINVEST is almost horizontal.

Variables TECH, CAP, and LOC are categorical, and scatter plots would have few points and thus be uninteresting. Frequency tables were therefore prepared. In each cell, the observed frequency and the expected frequency (in parentheses) are shown. For example, in table 1 there are 107 low tech (TECH = –1) ventures on the Honor Roll, whereas the expected frequency (assuming independence between TECH and SUCCESS) is 91. This frequency distribution indicates that low-tech firms have a disproportionately higher tendency to be on the Honor Roll. The next panel in table 1

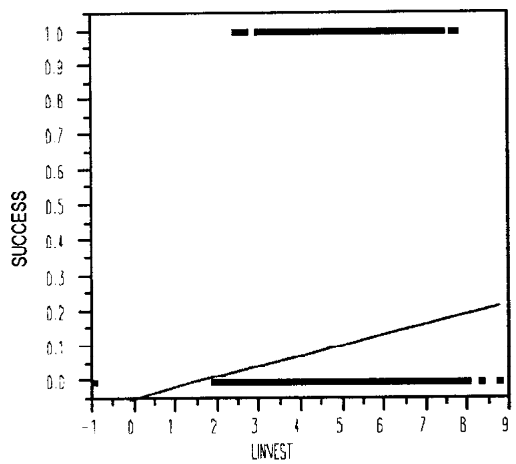

Figure 3. SUCCESS versus LINVEST.

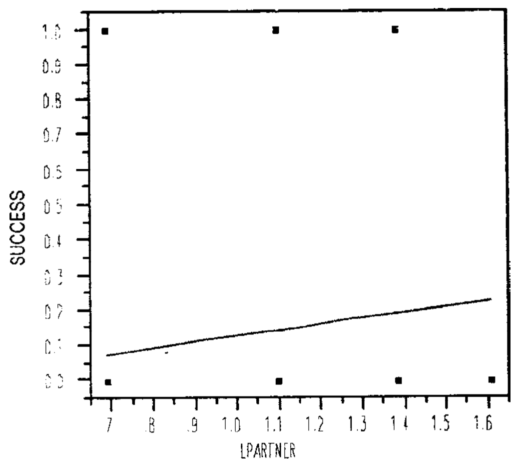

Figure 4. SUCCESS versus LPARTNER.

seems to indicate that labor intensive (CAP = –1) firms also have a disproportionately higher – although not as high as TECH – tendency to be on the Honor Roll. Location does not seem to have any effect on performance, as seen from the last panel.

In summary, the data set contains large overlaps between the successful and not-so-successful joint ventures in each variable. Traditional statistical analyses of means or distributions, or discriminant analyses, would have difficulty identifying the relationships between performance and the economic variables, and predicting the performance of a venture. This is the main motivation behind our use of neural networks.

Table 1

Tabulation of Sino–Hong Kong joint ventures [frequency (expected frequency)].

	Variable	Joint venture: Successful	Joint venture: Not successful
TECH	High	33 (49)	479 (463)
	Low	107 (91)	844 (860)
CAP	Capital intensive	19 (24)	236 (231)
	Labor intensive	121 (116)	1087 (1092)
LOC	Special	67 (63)	594 (598)
	Coastal	49 (50)	475 (474)
	Inland	24 (27)	254 (251)

3. Classification theory and neural networks

The following is a very brief synopsis of the statistical classification problem, which deals with assigning an object, based on a set of variables, to one of a finite number of groups. Since the performance variable of a venture is either successful or not, this is a two-group classification problem. Therefore, the description below is limited to two-group classification of joint ventures.

Let x_i be the vector of economic factors of joint venture i and ω be the state variable that a venture is successful. According to the Bayesian classification theory (Duda and Hart [7]), $P(\omega \mid x_i)$ is the probability that venture i belongs to the successful group and is called the *posterior probability*. This probability is extremely useful. For example, to classify a venture based on observed data x_i, let c_i^1 be the cost of wrongly assigning venture i to the successful group and c_i^2 be the cost of wrongly assigning it to the unsuccessful group. Then the optimal *Bayesian decision* is to assign venture i to the successful group if

$$c_i^1(1 - P(\omega \mid x_i)) \le c_i^2 P(\omega \mid x_i).$$

If the costs are equal, then the decision rule is to assign it to the successful group if $P(\omega \mid x_i) \ge 1/2$.

Typically, the posterior probability is a nonlinear function of x_i and cannot be derived directly. Fortunately, it has been shown (Hung et al. [16], among others) that least-squares estimators can provide an unbiased estimate of this probability. The formulation is as follows.

Let t_i denote group membership, with $t_i = 1$ if the venture is successful and $t_i = 0$ if it is not. For a given model (a neural network, for example), let a_i be the estimate of $P(\omega \mid x_i)$. Then a least-squares problem can be posed:

$$\min \sum_i (a_i - t_i)^2. \tag{1}$$

The solution to this problem satisfies the unbiasedness condition of statistical estimation, namely,

$$E(a_i) = P(\omega \mid x_i),$$

where E stands for expectation.

Neural networks offer a very convenient way to carry out this computation. Feedforward networks are used here. For two-group classification as this application calls for, only one output node is necessary. For each observation x_i, the economic factors are inputs to the network and, to train the network, the 0–1 target t_i is supplied to the output node. The networks have one hidden layer. The activation function at the output and hidden nodes is logistic. The output node has an additional scalar called the *bias*. A special feature is that the input nodes are connected not only to the hidden nodes, but also to the output node. The reason for this is to make it as close to logistic regression as possible.

To estimate the posterior probability, an approach called the *ensemble method* is used and is briefly described below.

3.1. Ensemble neural networks

Instead of relying on one network to estimate a function, Perrone and Cooper [27] suggest using a number of networks. This simple and yet powerful method is based on straightforward reasoning and is not limited to neural networks only. In simplified notation, let $f(x)$ be the unknown (vector) function to be estimated for observation (input vector) x. Suppose N neural networks (or models, in general) have been built, resulting in $f_j(x)$, $j = 1,\ldots,N$ predictions. The *basic ensemble method* simply uses the mean

$$\bar{f}(x) = \frac{1}{N} \sum_{j=1}^{N} f_j(x)$$

to estimate $f(x)$.

The justification is as follows. Define the error vector $e_j(x)$ as $e_j(x) = f(x) - f_j(x)$ and the variance as $\sigma_j^2 = E[e_j^2]$, where the expectation is taken over all observations x. Similarly, denote the variance of the basic ensemble method by $\sigma_{BEM}^2 = E[(f(x) - \bar{f}(x))^2]$. Further, let

$$\bar{\sigma}^2 = \left(\frac{\sum_{j=1}^{N} \sigma_j^2}{N} \right)$$

and

$$\sigma_{jk} = E(e_j e_k).$$

That is, $\bar{\sigma}^2$ is the average variance and σ_{jk} is the covariance of the errors between networks, $j \neq k$. It is straightforward to show that

$$\sigma_{BEM}^2 = \frac{1}{N^2} \left(\sum_{j=1}^{N} \sigma_j^2 + 2 \sum_{j=1}^{N} \sum_{k>j} \sigma_{jk} \right).$$

Assuming the errors are independent (i.e., $e_j(x)$ is independent of $e_k(x)$), then

$$\sigma^2_{BEM} = \frac{\bar{\sigma}^2}{N},$$

which shows that the variance of the ensemble method is only $1/N$ of the average variance of the N models. If the errors are not independent, one can show that

$$\sigma^2_{BEM} \le \bar{\sigma}^2.$$

The implication is that the ensemble estimator is likely to be more accurate (as measured by variance) than individual networks. Indeed, if one can create negative covariances, then σ^2_{BEM} can even be smaller than $\bar{\sigma}^2/N$. This is the idea behind several schemes to improve network prediction accuracy. See, for example, Munro and Parmanto [23].

In this study, we use independent samples to build the neural networks so as to increase the chances that the errors are independent.

4. Model building

The objective of this study is to discover the relationship between the economic factors of a Sino–Hong Kong joint venture and the probability that it is on the Honor Roll. Neural networks were built and estimation was based on the basic ensemble method.

4.1. Samples

The entire data set of 1,463 observations was partitioned into a training set and a test set. The former has 240 observations, leaving the latter with 1,223 cases. In order to be better able to distinguish the successful ventures (Group 1) from the not-so-successful ones (Group 2), the training sample has an equal number of both groups. First, 120 observations from each group were randomly chosen. Next, each sample was randomly partitioned into five subsamples. At the end of this step, there are five subsamples, each with 48 observations, half of them from the successful group. The distribution of the data set is shown below.

	Training set	Test set
Group 1	120	20
Group 2	120	1,203

The test set is never seen by the neural networks before they are trained. It is not totally desirable from a modeling point of view to have such a large test set (as compared to the training set); it is a result of the lopsided distribution between the two groups and the decision to have equal representation in the training set.

4.2. Neural networks

As mentioned before, the neural networks employed here are feed-forward types with one output node, eight input nodes, and a layer of hidden nodes. There is an arc connecting each input node to every hidden node and to the output node. Each hidden node is connected to the output node. The hidden nodes and the output node have logistic activation function. The output node has an extra variable called bias. The number of parameters to be estimated through training is equal to the number of arcs plus the biases.

To train a network, the sum of squared errors (SSE) in equation (1) is the objective function and the algorithm used was developed by Ahn [1] and employs a minimization method called Limited Memory Quasi-Newton of Nocedal [24] and starting values for network parameters are based on the data. The algorithm starts with a network with zero hidden nodes and uses starting values the same as those found in linear regression. Then a hidden node is added. The weights of the arcs in the previous network are used as starting weights. The starting weights for the new arcs are determined by a scheme designed to maximize the reduction in SSE. This approach of using deterministic starting values improves the chances that for similar data sets on the same network, the solutions for the network parameters will be close to each other. It is a desirable property and as in the ensemble method, multiple subsamples from the same data set are used.

How well a network can predict an unseen case is termed *generalizability*. Since our training samples are small, the entire data set is used to measure the generalizability of a network. The criterion is the mean square error (MSE), defined as

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{df},$$

where the degree of freedom df = number of observations – number of network parameters. The number of observations is 1,463. Since our networks have eight input nodes, one hidden layer, arcs between every pair of input nodes and hidden nodes and every pair of input nodes and the output node, and a bias at the output node, the number of network parameters is equal to $9(H + 1)$, where H denotes the number of hidden nodes. So the degree of freedom is $1454 - 9H$.

For each subsample, neural networks with hidden nodes ranging from 0 to 8 were trained. For each number of hidden nodes, all five networks were used as an ensemble to estimate the posterior probability of every joint venture. In other words, each a_i in equation (1) is the mean output from the five networks for observation i. The MSE was then computed, and the architecture with the smallest MSE was chosen. It turns out that MSE is fairly stable across the number of hidden nodes, ranging from 0.272 to 0.286. The architecture selected has six hidden nodes with an MSE of 0.272.

5. Results

The posterior probabilities estimated from the ensemble method with all 1,463 observations are plotted against each predictor variable, as shown in figures 5–8. A linear regression line is superimposed onto the scatter plot to signify a general relationship. (The stripes along multiples of $1/5$ of posterior probability are the results of averaging from five network outputs which tend to be 0 or 1.)

The first impression is the richness of information these new plots provide, as compared to those in figures 1–4. It is obvious that the relationship between each predictor variable and the performance variable is more complex than that which can be captured by a linear regression function. But the direction of the relationship is much clearer than that in the original data set. It is interesting to note that the direction is generally the same for the same variable. For example, for LDUR, both figures 1 and 5 show a positive relationship with the dependent variable. The only exception to this rule is PFINVEST, which has a slight positive slope in figure 2, but a slight negative slope in figure 6.

The regression coefficients in figures 5, 7 and 8 are significant at the 0.01 level. However, because of the large number of observations (1,463), statistical significance per se is not too important. What is more meaningful are the sign and the magnitude of the coefficient.

- LDUR. Figure 5 shows that as the duration of a joint venture increases, the likelihood of success also goes up. The estimated regression coefficient has a value of 0.4484. Previous research [4,17] identified partner commitment to a venture as one of the key performance factors. Results reported here clearly support this finding.
- PFINVEST. A small downward relationship is shown in figure 6. As the level of Hong Kong investors' equity control goes up, the posterior probability of success decreases. The pattern of decline as estimated by the regression coefficient takes on a value of –0.00123, with a p-value of 0.0145. This finding is in line with what was reported by Beamish [3].
- LINVEST. It is included in the model to remove the variation in SUCCESS due to the size of the joint venture. One would expect that a joint venture of a considerable amount of equity investment would more likely be commended by the China Association of Enterprises with Foreign Investment. A generally positive relationship is depicted in figure 7, with a coefficient of 0.08691.
- LPARTNER. A strong positive relationship is indicated in figure 8. The linear regression coefficient has a value of 0.71061. Hu and Chen [14] also reported a positive logistic regression coefficient for LPARTNER. Primary benefits of joint ventures are through corporate synergism. Yet, joint venture organizations also suffer the problems shared management. Thus, the performance of joint ventures is a net tradeoff between these two factors.

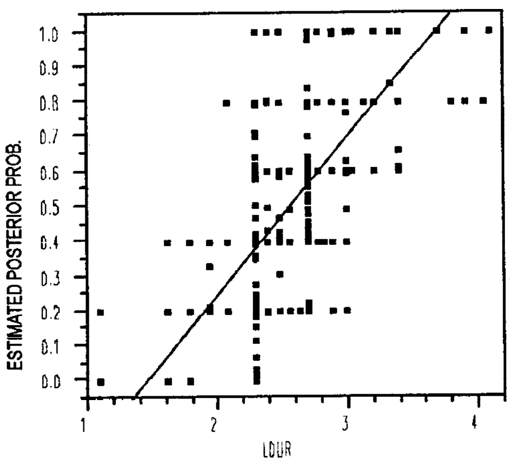

Figure 5. Estimated posterior probability versus LDUR
(posterior = –0.655 + 0.4484 LDUR).

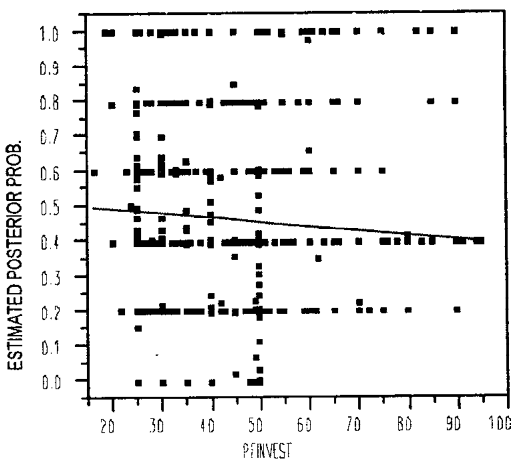

Figure 6. Estimated posterior probability versus PFINVEST
(posterior = 0.51618 – 0.00123 PFINVEST).

For the categorical variables, mean posterior probability for each value is reported in table 2. Statistical significance for the difference among means is also computed. The differences are all significant at the 0.01 level.

- TECH. High-technology industries have an average output value of 0.4138 and low-technology industries 0.4977. This finding that low-tech industries are more likely to be on the Honor Roll is contrary to the a priori assumption that technology improves the chance of good performance. In the case of China, however, it has been reported that technologies being transferred are very much at the

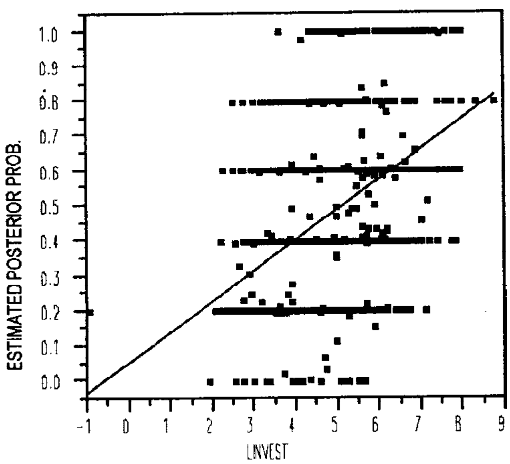

Figure 7. Estimated posterior probability versus LINVEST (posterior = 0.0505 + 0.08691 LINVEST).

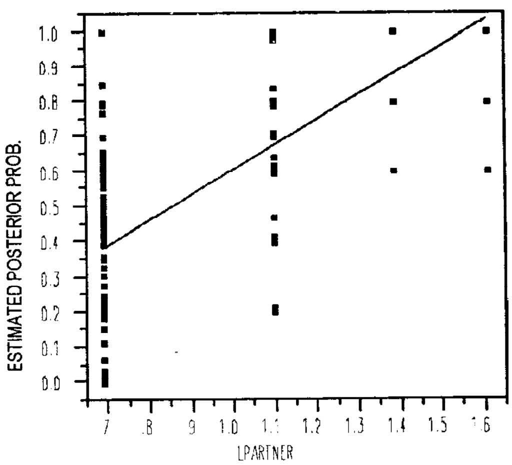

Figure 8. Estimated posterior probability versus LPARTNER (posterior = –0.1087 + 0.71061 LPARTNER).

last stage of the product life cycle. Furthermore, transferred technology into a developing economy typically entails two major problems. First, technology to be transferred may be too advanced to be useful for the host economy. Second, successful transfer of advanced technology will take a longer time than low-tech knowledge. Thus, one should not be surprised by the findings reported in this study.

- CAP. The average posterior probability for capital intensive ventures is 0.5339 and for labor-intensive 0.4545. This finding, again, is contrary to the a priori

Table 2

Estimated average posterior probabilities for categorical variables.

Variable		Average posterior probability
TECH	High	0.4138
	Low	0.4977
CAP	Capital intensive	0.5339
	Labor intensive	0.4545
LOC	Special	0.4480
	Coastal	0.4950
	Inland	0.4664

assumption for the performance of joint ventures in China. This reflects the differences between the two groups of industries. But it could also be related to the amount of investment as capital-intensive ventures typically entail more equity investment and there is a positive relationship between LINVEST and SUCCESS.

- LOC. The average posterior probability for LOC = 2, coastal regions, is higher than the averages for the other two regions. An even bigger surprise is that ventures formed in the Special Economic Zones have lower probability than those in the inland provinces. It indicates that all the special incentives do not seem to contribute to the chances of success.

6. Conclusions

This paper reports the analysis of the performance of Sino–Hong Kong joint ventures. Seven economic factors, with three being categorical, were used to predict whether a venture would be on the Honor Roll. An ensemble of neural networks was built to do the prediction. The mean of the outputs from five networks was used to estimate the posterior probability.

A natural question is whether the ensemble method was necessary. To evaluate this question, all five subsamples were put together into one training set and a neural network of six hidden nodes was trained. This network was used to predict the posterior probability of all 1,463 ventures. The MSE of this single network is 0.344, which is about 26% greater than that of the ensemble method. So as predicted by theory, the ensemble method indeed produces more accurate predictions.

The results produced by the ensemble networks provide much more information than the original data. In a sense, this report represents a case study of successful data mining and management decision support.

References

[1] B. Ahn, Forward additive neural network models, Ph.D. Dissertation, Kent State University, 1996.

[2] E. Anderson and H. Gatignon, Modes of foreign entry: A transaction cost analysis and propositions, Journal of International Business Studies 11(1986)1–26.

[3] P. Beamish, Joint venture performance in developing countries, unpublished Doctoral Dissertation. The University of Western Ontario, 1984.

[4] P. Beamish and J. Banks, Equity joint ventures and the theory of the multinational enterprise, Journal of International Business Studies (Summer, 1987)1–16.

[5] F. Contractor and P. Lorange, *Cooperative Strategies in International Business*, Lexington Books, Boston, 1988.

[6] W. Davidson, Creating and managing joint ventures in China, California Management Review 29(1987)77–94.

[7] R.O. Duda and P. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

[8] J. Dunning, Explaining changing patterns of international production: In defence of the eclectic theory, Oxford Bulletin of Economics and Statistics (November, 1979)269–296.

[9] L. Franko, *Joint Venture Survival in Multinational Corporations*, Free Press, New York, 1971.

[10] M. Geringer and L. Hebert, Control and performance of international joint ventures, Journal of International Business Studies 20(1989)233–253.

[11] M. Geringer and L. Hebert, Measuring performance of international joint ventures, Journal of International Business Studies 22(1991)249–263.

[12] B. Gomes-Casseres, Ownership structures of foreign subsidiaries, theory and evidence, Journal of Economic Behavior and Organization 11(1989)1–25.

[13] J. Hennart, A transaction costs theory of joint ventures, Strategic Management Journal 9(1988) 361–374.

[14] M.Y. Hu and H. Chen, An empirical analysis of factors explaining foreign joint ventures performance in China, Journal of Business Research 35(1996)165–173.

[15] M.Y. Hu, M.S. Hung, M.S. Shanker and H. Chen, Using neural networks to predict the performance of Sino–foreign joint ventures, International Journal of Computational Intelligence and Organizations 1(3)(1996)134–143.

[16] M.S. Hung, M.Y. Hu, M.S. Shanker and B.E. Patuwo, Estimating posterior probabilities in classification problems with neural networks, International Journal of Computational Intelligence and Organizations 1(1996)49–60.

[17] P. Killing, *Strategies for Joint Venture Success*, Praeger, New York, 1983.

[18] B. Kogut, Joint ventures: Theoretical and empirical perspectives, Strategic Management 9(1988) 319–332.

[19] B. Kogut, The stability of joint ventures: Reciprocity and competitive rivalry, The Journal of Industrial Economics 38(1989)183–198.

[20] D. Lecraw, Performance of transnational corporations in less developed countries, Journal of International Business Studies 14(1983)15–33.

[21] C. Lee, International production of the United States and Japan in Korean manufacturing industries: A comparative study, Weltwirtschaftliches Archiv 119(1983)745–753.

[22] Ministry of Foreign Economic Relation and Trade, various editions, *Almanac of China's Foreign Economic Relations and Trade*, Beijing, China.

[23] P. Munro and B. Parmanto, Competition among networks improves committee performance, in: *Advances in Neural Information Processing Systems 9*, eds. D. Touretzky, M. Mozer and M. Hasselmo, MIT Press, Cambridge, MA, 1997.

[24] J. Nocedal, Updating quasi-Newton matrices with limited storage, Mathematics of Computation 35(1980)752–773.

[25] E. Patuwo, M. Hu, M. Shanker and M. Hung, Neural network analysis of performance of Sino–Hong Kong joint ventures, in: *International Conference on Management Science and The Economic Development of China*, Hong Kong, 1996.

[26] People's Daily, Commendation for outstanding foreign investment enterprises in China from China Association of enterprises with foreign investment, July 16, 1991.

[27] M. Perrone and L. Cooper, When networks disagree: Ensemble methods for hybrid neural networks, in: *Artificial Neural Networks for Speech and Vision*, ed. R. Mammone, Chapman and Hall, London, 1993, pp. 127–142.

[28] J. Reynolds, The "pinched shoe" effect of international joint ventures, Columbia Journal of World Business (1984)23–29.

[29] J. Schaan and P. Beamish, Joint venture general managers in LDCs, in: *Cooperative Strategies in International Business*, eds. F.J. Contractor and P. Lorange, Lexington Books, Boston, 1988, pp. 279–299.

[30] O. Shenkar and Y. Zeira, Role conflict and role ambiguity of chief executive officers in international joint ventures, Journal of International Business Studies 23(1992)55–75.

[31] M. Teagarden and M. Von Glinow, Sino–foreign strategic alliance type and related operating characteristics, International Studies of Management and Organization 20(1990)99–108.

[32] J. Tomlinson, *The Joint Venture Process in International Business: India and Pakistan*, MIT Press, Cambridge, MA, 1970.

[33] US-China Business Council, *Special Report on U.S. Investment in China*, The China Business Forum, Washington, DC, 1990.