International Journal of Innovative
Computing, Information and Control
Volume 9, Number 4, April 2013

ICIC International ©2013 ISSN 1349-4198
pp. 1493–1510

ASSESSMENT OF IC CLUSTERING EVOLUTION BY USING A NOVEL DIFFUSION MODEL AND A GENETIC ALGORITHM

Bi-Huei Tsai

Department of Management Science
National Chiao Tung University
No. 1001, Ta-Hsueh Road, Hsinchu 30010, Taiwan
bhtsai@faculty.nctu.edu.tw

Received February 2012; revised June 2012

Abstract. *The evolution of industrial clusters is a critical factor in the strategic development of locations for high-tech industries. Most previous studies have neglected the quantitative aspects of industrial clustering and their access to crucial data may have been limited. This work developed a novel diffusion model to illustrate the extension of IC clusters from Taiwan to China and included foreign direct investment (FDI) as a quantitative indicator of industrial clustering. We also modified the conventional Bass model by incorporating profitability as an incremental factor in conjunction with regulatory FDI limits as constraints for parameter estimation. A genetic algorithm was used with the fourth-order Runge-Kutta algorithm for parameter optimization. Finally, t-statistics were used to compare clustering features among three IC sub-industries: design, packaging/testing and memory modules. Simulation results demonstrate negligible standard deviation in the optimized parameters, which confirms the reliability of our findings. Furthermore, comparison results demonstrate that the proposed model is more stable and accurate than the conventional Bass model. The proposed approach is applicable to other high-tech industries and other locations around the world.*
Keywords: Genetic algorithms, Diffusion model, t-statistics, Parameter stability, Vertical disintegration

1. **Introduction.** This work applied a genetic algorithm (GA) to explore the parameters involved in the movement of industrial clusters from Taiwan to China. We then compared incentives for the clustering of various industrial sectors along the IC value chain. For over ten years, Taiwanese firms have dominated the global IC industry and currently account for $> 70\%$ of the market share worldwide. The Taiwanese IC industry is largely concentrated in the Hsinchu Science and Industrial Park (HSIP), where numerous producers are efficiently interconnected in production, marketing, logistics and technical development [1,2]. The location of Taiwanese IC clusters must be taken into account when any semiconductor-related industry develops strategies for the establishment of factories or sales offices. Since 2000, China has provided inexpensive land, cheap labor, preferential tax treatment, and administrative support for Taiwanese IC enterprises establishing subsidiaries in China. As a result, understanding the clustering behavior of Taiwanese IC firms in China has become an issue of critical importance in the semiconductor industry. The purpose of this study was to examine the process of IC clusters moving from Taiwan to China, while correcting for the skewed results obtained in previous studies. This research was motivated by the need for a diffusion model capable of minimizing bias through the incorporation of factors related to profitability. We also obtained a unique data set suitable for a quantitative investigation of this phenomenon and optimized the parameters with a high degree of accuracy using a genetic algorithm. This work then

applied the model to the analysis of upstream and downstream IC sub-industries involved in industrial clustering in China.

This study began by developing a novel diffusion model to transform value-added data into parameter values to quantify various aspects of clustering in the IC industry. Most previous studies have employed qualitative methods to map industrial clusters [3] and determine the spatial concentration of related firms [4]. Not only did these studies neglect the quantitative aspects of this issue, but their access to crucial data may also have been limited. This work adopted foreign direct investment (FDI) as a quantitative indicator to evaluate the extension of industrial clusters. In this study, FDI is defined as the capital investment from firms in one country used to build facilities in another. As Taiwanese IC companies have expanded into China through investment in production facilities, the amount of FDI transferred to China has increased, thereby leading to the expansion of industrial clusters located in China. The amount of FDI invested by Taiwanese companies in China is a reliable gauge of the extent to which industrial clusters from Taiwan have expanded into China. In most countries, data related to FDI in specific regions is difficult to obtain; however, the Taiwan Economic Journal (TEJ) compiles data related to the actual FDI invested by Taiwanese IC firms in China. This work adopted this unique database for our diffusion model.

Our proposed diffusion model incorporates profitability as an incremental factor and regulatory FDI limits as a constraint in the estimation of parameters. The IC industry is capital-intensive and initial investments can only be recovered through successful long-term operations. Newly established Chinese subsidiaries are not always profitable in the early stages and companies that suffer serious initial losses are inclined to consider the FDI of other companies when deciding whether to pursue further expansion into China. This trend has a powerful influence on the evolution of industrial clustering from Taiwan to China. Despite the importance of the profitability of Chinese subsidiaries in predicting the formation of IC industrial clusters, previous researchers adopted the conventional Bass model, which overlooks factors related to profitability. In contrast, this study specifies time-varying profitability as a function of industrial clustering in our proposed model. In addition, no previous studies have considered the influence of investment limits on the process of extension through investment in industrial clusters. Because the Taiwanese central government has placed an upper limit on FDI to China, this study adopts FDI regulatory limits as a constraint in the estimation of parameters used in simulations.

With regard to parameter optimization, previous studies have employed ordinary least squares regression to estimate parameters [5]; however, multi-collinearity between variables often results in biased estimations. Other studies have used the nonlinear least squares (NLS) approach to optimize parameters. Tsai, Li and Lee [6] described how this approach uses a sequential search technique to estimate parameter values. The use of iterative procedures in the NLS method requires that users provide initial values for unknown parameters prior to optimization. Unfortunately, the initial values must be reasonably close to unknown parameter estimates; otherwise, optimization will not necessarily converge. Erroneous initial values can result in convergence to a local solution instead of a global solution. By contrast, Goldberg [7] indicated that GAs perform well in identifying solutions to all types of problems because no assumptions are made regarding the topic of investigation. Gen and Chang [8] demonstrated that GAs are capable of providing a real number evaluation function within the optimal solution and many subsequent researchers have combined additional methods with a GA to improve forecasting accuracy. To avoid the drawbacks of the NLS method, a GA can be used to identify the initial values; GAs function according to internal rules, and make no assumptions regarding the shape of the objective function. This enables GAs to scan and enhance a vast set of solutions, which eventually evolve into an optimal solution within a reasonable number of computational iterations. This work also adopted the GA with the fourth-order Runge-Kutta algorithm [9] to provide a numerical solution to the proposed nonlinear ordinary differential equation. This work ran the GA simulation through 800 iterations to obtain 800 sets of parameters. We then determined whether the estimated parameters were located within acceptable intervals. Our combination of a genetic algorithm (GA) with the fourth-order Runge-Kutta algorithm is a novel approach to the numerical optimization of parameters related to the evolution of industrial clusters.

This is the first study to use statistical analysis to compare the characteristics of clustering among three IC sub-industries: design, packaging/testing and memory modules. No previous studies have attempted to quantify the difference between various sectors of the IC industry under a vertical disintegration framework. Vertical disintegration refers to the subdivision of the IC industry into several levels along the value chain, to facilitate the specialization of firms at each level [10]. In the Taiwanese IC industry, independent upstream firms specialize in IC design, midstream firms focus on IC manufacturing, and downstream firms deal with IC packaging and testing, or IC memory modules. Because the Taiwanese government prevents manufacturing firms from establishing factories in China, this study focused on design, packaging/testing and memory modules, comparing the expansion of firms in these IC sectors to identify the dominant factors in the process of industrial clustering. This work seeks to identify which sector depends most heavily on Chinese promotions and which sector seeks to imitate the actions of other Taiwanese IC firms. This is the first study to use statistical analysis to compare the motivation behind clustering between upstream and downstream operations.

In summary, our novel modified diffusion model incorporates profitability as an incremental factor and FDI regulatory limits as a constraint, and employs a GA with the fourth-order Runge-Kutta algorithm to perform parameter optimization. Finally, we compared the forecasting accuracy and model stability of the proposed model with that of the conventional Bass model. The simulation results clearly demonstrate that early adopters (the first enterprises to establish branches in China) have a significant impact on IC clustering. The success of these enterprises persuades other firms to imitate their expansion into China and the evidence of this influence is most pronounced in the IC packaging/testing sector. Our results prove that profitability influences the expansion of industrial clusters into China and provides evidence as to the superiority of the modified model (which considers profitability) over the conventional Bass model with regard to prediction accuracy and stability. The modified model is capable of providing accurate predictions related to the clustering of Taiwanese IC firms in China.

The remainder of this paper is organized as follows. Section 2 provides a background of IC clusters moving from Taiwan to China. Section 3 outlines the methodology used to interpret the expansion of IC clusters from Taiwan to China. Section 4 presents data of actual FDI from Taiwan to China, comparing upstream with downstream processes. Section 5 reports the estimation of parameters, the stability of parameters and forecasting accuracy. Finally, conclusions are provided in Section 6.

2. Background.

2.1. Formation of IC industrial clusters in China.

China provides foreign enterprises with access to inexpensive land, low labor costs, preferential tax treatment, political stability and administrative support [11]. Pressure from labor costs in Taiwan and opportunities in China have prompted many Taiwanese IC firms to relocate production facilities to China. Product life cycle theory asserts that the manufacturing of products is initially performed in developed countries and is later moved to emerging countries [12]. Once production techniques have matured, even low-wage laborers in China can be trained to perform production techniques quickly and easily.

The number of suppliers of raw materials and manufacturers of instruments and equipment for IC firms has increased considerably in China. For instance, semiconductor material suppliers, such as Applied Materials and T.E.L Ltd.; semiconductor instrument suppliers, such as Varian, Axcelis and KLA-Tencor; and mask lithography suppliers, such as ASML, Nikon and Canon, are establishing operations in China to service nearby IC firms. As emphasized by Porter [13,14], the formation of industrial clusters creates a platform on which the pool of workers with specialized skills may exchange experiences and share information. Such an atmosphere is highly conducive to productivity, enabling suppliers to attract other Taiwanese IC firms to establish factories in China. Mainland China is gradually transforming from the world's factory into the world's market. In accordance with location theory, corporations establishing foreign subsidiaries must consider the distance their products must be shipped to reach markets [15]. We infer that Taiwanese firms have established Chinese branches to service Chinese customers. Chinese subsidiaries send finished goods directly to nearby downstream assemblers, who fabricate products to be sold to local customers. In addition, China imposes tariffs on imported products but not on products manufactured in China. As a result, global integrated device manufacturers (IDMs), such as Intel, Hynix, Qimonda and NEC, as well as the competitive IC manufacturers (such as Samsung and LG), have established strong IC industrial clusters in mainland China. All production processes in IC-related industries in Taiwan, (ranging from upstream IC design to the downstream IC memory module industry), have imitated large semiconductors corporations in the establishment of divisions in China. Shanghai, Shenzhen, and neighboring regions (i.e., Suzhou, Kunshan and Wujiang) have become the primary locations for cross-strait investment from Taiwan [16].

2.2. **Clustering firms in IC design, packaging and testing, and memory modules.** The development of industrial clusters differs according to manufacturing processes, functional characteristics, and the sector along the IC value chain. Because China is the largest market in the world, IC design firms must respond to the demands of the Chinese market as rapidly as possible. In the diffusion theory of Bass [5], "innovation" refers to Chinese promotions offering preferential terms for investment in China, and "imitation" refers to successive firms investing in China, after hearing of the benefits of the Chinese manufacturing environment from the first companies to establish facilities there.

Taiwanese IC design companies have established divisions to obtain information concerning consumer preferences in China and continually seek to manufacture products capable of meeting those demands. Conversely, downstream processes (i.e., packaging/testing) involve standardized procedures and machinery. Similarity in production processes allows new employees to become familiar with operations and an increasing number of branches in China perform in-house training of employees. As a result, downstream firms are inclined to imitate preceding firms and establish divisions in China to take advantage of low wages and increase profits.

Thus, the imitation effects on industrial clustering into China are greater in downstream firms than in upstream firms. The differences in technological skills and manufacturing features among IC design, packaging/testing, and memory modules firms lead to the differences in the innovation and imitation effects on industrial clusters in various sectors along the IC value chain.

2.3. **Taiwanese regulatory limits on investment in China.** To prevent Taiwanese IC firms from taking root in China and to curb the flow of skills out of the country, the Taiwanese central government has prohibited IC enterprises from moving production facilities to China. Under the "Regulations Governing the Approval of Investment or Technical Cooperation in Mainland China", Taiwanese firms are required to submit applications to the Investment Commission, Ministry of Economic Affairs (MOEA), Taiwan, prior to the commencement of construction in China. The Investment Commission determines whether to approve an application and assigns a maximum investment limit, which is enforced by the Taiwanese central government. Nonetheless, IC firms have sought to establish Chinese divisions because IC firms in Taiwan are losing their competitive advantage, due to changes in the operating environment, increased labor costs, a shortage of skilled employees and a lack of industrial land.

3. **Methodology.**

3.1. **Model formulation.**

3.1.1. *Conventional Bass model.* This work utilizes the total FDI of the Taiwanese IC industry into China as the primary indicator of IC industrial clustering. The amount of FDI represents the flow of capital out of Taiwan into China, and indicates the degree to which industrial clusters extend into China. This investigation applies the conventional Bass model to the process of clustering as expressed in Equation (1):

$$n(t) = \frac{dN(t)}{dt} = g(t)(M - N(t)), \tag{1}$$

where $N(t)$ is the cumulative FDI of the Taiwanese IC industry into China since time t; $n(t)$ is the amount of FDI at time t; and parameter M is the amount of capital available to flow from Taiwan into China, such that the difference between M and $N(t)$ is the remaining flow of capital at time t. The incremental flow capital at any given point $\frac{dN(t)}{dt}$ is directly proportional to the flow of remaining capital $(M - N(t))$. The growth rate, $g(t)$, can be written as Equation (2):

$$g(t) = \left(p + q\frac{N(t)}{M}\right). \tag{2}$$

By substituting Equation (2) into Equation (1), and using the assumptions mentioned above, the rate of dynamic change of clustering $g(t)$ is proportional to the cumulative flow of capital out of Taiwan to China $N(t)$. Thus, the dynamics of capital flow from Taiwan to China is expressed as

$$n(t) = \frac{dN(t)}{dt} = \left(p + q\frac{N(t)}{M}\right)(M - N(t)). \tag{3}$$

According to diffusion theory, innovation parameter p and imitation parameter q are used to capture incentives for industrial clustering. Innovation parameter p comprises incentives unrelated to the clustering experience of firms with Chinese divisions. Innovation parameter p is determined by Chinese promotions designed to encourage foreign corporations to establish factories or subsidiaries in China. As a response to the demands of the Chinese market, other corporations may also be tempted to develop divisions in China. As innovation parameter p increases, the rate of change in clustering $g(t)$ increases, such that the speed at which capital flows from Taiwan into China, $N(t)$ is accelerated.

Conversely, imitation parameter q describes how experienced firms that have already established Chinese divisions inspire other IC corporations to expand into China. Taiwanese firms that already have factories in China communicate the advantages of low labor costs in China to other firms, who imitate their predecessors by setting up Chinese divisions. Interpersonal communication and intra-industry interaction convey the information that motivates expansion into China. This work utilizes two coefficients (innovation p and imitation q) to interpret the incentives for the evolution of industrial clusters.

3.1.2. *Modified model incorporating profitability as incremental variable and governmental FDI limits as a constraint for parameter estimation.* Previous studies have shown that imitation strongly influences the diffusion of products [17-19]. We infer that many enterprises joining clusters in China are imitating the actions of firms already established there. As the profitability of Chinese subsidiaries increase, the amount of capital required from Taiwanese firms decreases. Thus, profitability bankrolls the continuous expansion of factories in China, such that Taiwanese firms no longer have to invest as heavily in the new operations. Conversely, the IC industry is capital-intensive and the initial investment can only be recovered through successful long-term operations. In addition, newly established Chinese subsidiaries are not necessarily profitable in the early stages. Once Taiwanese firms have decided to establish a presence in China, they cannot afford to retreat from their Chinese operations. The IC firms that have suffered losses in Chinese branches are more likely to follow other companies' decisions to determine the FDI amount to further devote to their Chinese subsidiaries and to maintain their ability to compete with other IC firms in China. Hence, the tendency to imitate other firms is affected by the profitability of operations in China – the prospect of future gains might encourage firms to investment more despite initial losses. Because the time-varying incentive to imitate other firms is correlated to profits, estimations based on the conventional Bass model which does not consider profitability may be biased. Hence, this work relaxes the restrictive assumption of constant incentives to imitate other firms, specifying them as a function of profitability. The incentive to imitate others in joining profitability factors is revised as Equation (4) in the proposed model.

$$q_t = q_1^{\exp(EPS_t \times \delta)}. \tag{4}$$

where EPS_t is the additional contribution of Chinese subsidiaries to the earnings per share (EPS) of Taiwanese firms and is the indicator of profitability. Because the Taiwanese government prevents IC manufacturing firms from moving manufacturing facilities to China, the IC manufacturing sector is not an issue. Thus, the total sample of IC firms in this investigation was divided into three sub-industries: design, packaging/testing and memory modules. We calculated EPS_t for each sub-industry by dividing the total additional contribution of all sample firms by the number of firms in the sub-industries.

Clearly, the term $\exp(EPS_t \times \delta)$ in Equation (4) indicates how profitability influences the incentive to imitate other firms through investment in China. The coefficient of imitation at period t, q_t, varies with profitability over time, making it a function of profit variable EPS_t. Parameter δ denotes the marginal influence of profitability on the incentive to imitate other firms. If q_1 is positive and < 1 ($0 < q_1 < 1$), a positive coefficient δ indicates that profit deters successive remittance in China by diluting the flow of capital into China. Conversely, it is less likely that coefficient δ is negative. IC firms do not initiate new projects until they have established a long-term plan and secured the funds required for an extended commitment. IC firms in the beginning stages of expansion require a great deal of capital for machines, materials and facilities; the cost of depreciation is also large. Therefore, profits and revenues are not immediately apparent and the coefficient δ is less likely to be negative.

This work uses coefficient δ to capture how earning factors influence IC clustering dynamics. Thus, Equation (4) is input into Equation (3), yielding Equation (5):

$$n(t) = \left(p + q_1^{\exp(EPS_t \times \delta)} \times \frac{N(t)}{M}\right)[M - N(t)]. \tag{5}$$

In Equation (5), the incentive to imitate other firms depends on the ability of foreign branches to earn a profit and is not always constant. The Taiwanese central government has established maximum investment limits for Taiwanese firms establishing branches in China. Hence, the amount of capital flowing from Taiwan to China, M, in the modified model is assumed to be the maximum permitted limit – roughly NT$9.7 billion, NT$28 billion and NT$17 billion for the IC design, IC packaging/testing and IC memory module industries, respectively.

3.2. **Simulation procedure.**

3.2.1. *Basic concept of the genetic algorithm.* This work utilizes a GA to determine the initial value of parameters. GAs search for a global solution with a value approaching the maximum determined by an evaluation function [7]. GAs use a parallel, evolutionary search algorithm to locate parameters to optimize an objective function [21]. Based on Darwin's theory of natural selection and evolution, a chromosome (or an individual) in a GA can be defined as an enciphered entity of a candidate solution, which is expressed as a set of variables. Initially, a population of chromosomes is created randomly and the chromosomes in the following generation are produced through selection, crossover, and mutation operations based on the values of fitness functions. The GA then generates an initial set of solutions that eventually evolve into a final set of solutions, while constantly improving in accordance with fixed criteria [22]. Holland [23] provided mathematical justification of the success of a GA. The advantages of a GA are two-fold. First, a GA does not require rules for problems and can provide solutions for highly complex search spaces. GAs can quickly scan a vast solution set using its own internal rules. Second, GAs are capable of escaping a local optimum. Many studies have demonstrated the superiority of GAs over traditional gradient search algorithms.

3.2.2. *Computational procedure.* Figure 1 presents a flowchart of the proposed computational procedure and the steps we took in our investigation. The GA was used to identify the initial values. The fourth-order Runge-Kutta algorithm was then implemented to numerically solve the nonlinear ODE. Previous studies have utilized ordinary least squares regression estimates [4,5], but the multi-collinearity between variables causes bias. Conversely, numerical simulation is capable of producing good estimates of unknown model parameters with a relatively small dataset. Numerical simulation is fairly well developed for the computation of confidence, prediction, and calibration intervals to answer questions. Thus, we employed numerical simulation for the small sample set of quarterly FDI data to solve the nonlinear ordinary differential equation (Equations (3) and (5)). Once the computed solution was obtained, it was calibrated with real data to generate an accurate simulation. If the tolerance of the computational results exceeded that of an error, optimal parameters and results were obtained. Otherwise, a least squares optimization technique was used to update parameters for following iteration. The GA estimation procedure was performed through 800 iterations, generating different initial values with each iteration, resulting in a matrix of 800 estimated parameters p, q, q_1 and δ. We then applied t-statistics to examine whether parameters p, q, q_1 and δ differed significantly from zero to determine which factor most strongly influenced the clustering process. After obtaining the final parameter estimations separated into IC design, packaging/testing, and memory module industries, we compared the parameters representing clustering incentives and evaluated the satiability of parameters, and the accuracy of forecasting using the conventional Bass model and the proposed model. The following sections discuss in detail how the investigations of parameter comparison, parameter satiability, and forecasting ability were conducted.

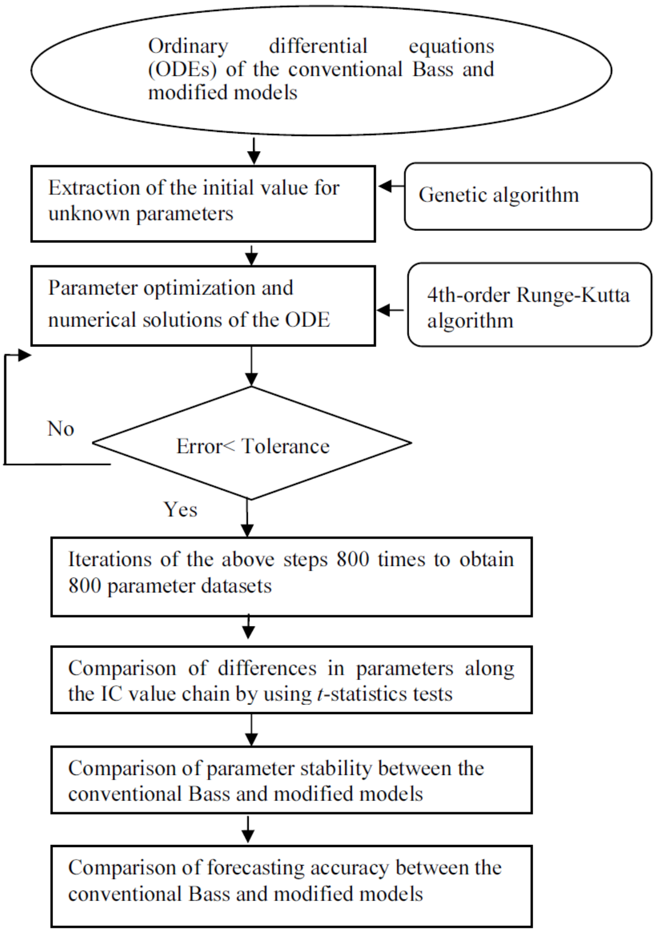

FIGURE 1. Flowchart of the computational procedure and investigation steps in this work

3.2.3. *T-statistics test.* This section provides an in-depth examination of the difference between imitation and innovation coefficients $(q - p)$ in the three IC sub-industries. This study employed the conventional Bass model and a modified model incorporating factors related to profitability to estimate the marginal effect of incentives to imitate other firms, q, in Equation (3) or q_1 in Equation (5) and the marginal effect of innovation incentives p, on industrial clustering. The t-statistics tests were then conducted to determine which incentive(s) dominate clustering dynamics. The incentives were compared among the three sub-industries by examining p, q and q_1. We initially grouped the parameters of the three industries into pairs and used a t-statistics test to examine the differences $(q - p)$ among them. We determined which sector was the quickest to cluster in China.

3.3. **Parameter stability.** Following parameter calculations, we evaluated parameter stability and forecasting accuracy to determine whether the modified model (incorporating factors of profitability and regulatory FDI limits Equation (5)), is capable of predicting the process of expansion through IC industrial clusters. We employed the approach developed by Golder and Tellis [24] to evaluate parameter stability between the conventional Bass model and our proposed modified model. We re-estimated the parameter values over different sample periods, from the initial period to four terminal periods. The four terminal periods were the second quarter of 2008, the third quarter of 2008, the fourth quarter of 2008, and the first quarter of 2009. In other words, parameter optimization began with a data series from the initial stage to the second quarter of 2008 and added subsequent periods to the other three sets to compare the four sets of the parameter values. If all four sets of optimized parameters using different sample periods for parameter optimization were similar, the stability of the model would be confirmed. The two measures of parameter stability are expressed as Equations (6) and (7):

$$STAB1 = \frac{mean\left(\overset{\wedge}{\beta}\right)}{Std\left(\hat{\beta}\right)}, \tag{6}$$

and

$$STAB2 = \sum \left| \frac{\overset{\wedge}{\beta}_t - \overset{\wedge}{\beta}_{t-1}}{mean\left(\overset{\wedge}{\beta}_t\right)} \right| \frac{1}{K}, \tag{7}$$

where $\overset{\wedge}{\beta}$ represents parameter estimates, K is the number of intervals, and $Std\left(\overset{\wedge}{\beta}\right)$ and $mean\left(\overset{\wedge}{\beta}\right)$ are the standard deviation and mean of parameters, respectively. Hence, *STAB1* captures fluctuations in the overall mean. This measure is the mean of the multiple parameter estimates divided by their standard deviation, where multiple estimates are obtained by adding an additional quarter. Notably, *STAB2* captures period-to-period fluctuations, as the average period-to-period change standardized by the mean of parameters. In theory, a high *STAB1* value combined with a low *STAB2* value would indicate that a model has good parameter stability. This work calculated *STAB1* and *STAB2* to compare their parameter stability of our modified model with that of the conventional Bass model.

3.4. **Prediction ability.** Mean absolute percentage error (MAPE) was used to compare the prediction accuracy of the conventional Bass model with that of our modified model, while considering the impact of profitability and the regulatory FDI limit on expansion through IC clusters. MAPE is calculated as follows:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \frac{\left|Y_t - \hat{Y}_t\right|}{Y_t}, \tag{8}$$

where Y_t and $\hat{Y}_t$ represent the real and predicted cumulative FDI at time t, respectively. The parameters of these models were estimated using the quarterly FDI sent to China up to the first quarter of 2008, beginning with the first year that each of the following Taiwanese IC industries entered China: IC design (2000), IC packaging/testing (2000) and IC memory modules (1998). We compared the forecasted quarterly FDI from the second quarter of 2008 to the first quarter of 2009 with actual quarterly FDI figures.

4. **Sample and Data.** We collected real data covering 86 public IC firms in Taiwan, involved in IC design (52), IC packaging/testing (13) and IC memory modules (21). As of 2002, no IC manufacturing firms except for the Winbond Corporation were permitted by the Taiwanese government to relocate to China. As a result, we excluded the IC manufacturing industry from this sample. We obtained quarterly data for the period when investment in China began. Because firms in each sector entered China at different times, the initial period differed among these IC sub-industries. The sample period began from the initial period to the first quarter of 2009. Sample data were divided into two subsamples: training samples and test samples. All prediction models were developed

using the training sample, ranging from the initial period of each IC industry to the first quarter of 2008. The accuracy of the conventional Bass model and our modified model was evaluated by comparing predicted values with real data using the test sample from the second quarter of 2008 to the first quarter of 2009. Data of permitted and actual FDI into China were collected from the Taiwan Economic Journal (TEJ) database. Actual FDI is defined as the net FDI to China; i.e., the FDI sent to China minus the amount returning to Taiwan.

5. **Empirical Results.**

5.1. **Parameter estimation results.**

This work employed total FDI in China from Taiwan as the primary indicator of industrial clustering. Table 1 lists descriptive statistics for the net FDI into China in the three IC sub-industries. The mean and median of each firm's EPS are negative (Table 1). Generally, the IC industry is capital-intensive, due to the need to re-locate entire factories, which generated considerable depreciable expenses. Because all IC design and packaging/testing firms in this sample began clustering in China after 2001, a considerable amount of depreciable expenses were allocated to Chinese facilities during the sample period. Furthermore, in the initial stage of clustering, few orders were received and revenue was low, such that the mean of earnings variable EPS_t in the quarterly data did not exceed zero.

Table 2 shows the analytical results using the conventional Bass model and our modified model. The 800 parameter estimates were evaluated using t-statistics to determine whether they differed significantly from zero. The reported estimates are the mean of parameter values over 800 iterations, as shown in Table 2. Using the conventional Bass model, the coefficient values of the incentive to imitate other firms (q) were 0.1499, 0.2233 and 0.0569 for IC design, IC packaging/testing and IC memory module sub-industries, respectively. The test results indicate that the coefficient values of the incentive to imitate

Table 1. Descriptive statistics of two variables in three IC industries: the design, packaging and testing, and memory module industries. The first variable is quarterly cumulative FDI amount, $N(t)$. The second variable is the additional contribution of Chinese subsidiaries to the earnings per share (EPS) of Taiwanese firms, EPS_t.

	$N(t)$ in Equation (5) (unit: NT thousand dollars)		
	Design	Packaging and Testing	Memory Module
Mean	2,996,788	6,643,664	2,418,012
Median	2,933,664	2,954,681	1,822,775
Maximum	7,595,920	22,648,874	5,631,584
Minimum	9,837	29,127	820,904
Standard Deviation	2,445,657	7,371,731	1,499,742
	EPS_t in Equation (5) (unit: NT dollars)		
	Design	Packaging and Testing	Memory Module
Mean	–0.0506	–0.0469	–0.0495
Median	–0.0500	–0.0238	–0.0358
Maximum	–0.0463	0.0100	0.1300
Minimum	–0.1400	–0.1936	–0.3025
Standard Deviation	0.0380	0.0539	0.1130

Table 2. Parameter estimates by the GA for the conventional Bass model and modified model. Reported estimates are parameter means over 800 iterations. This work computes the t-statistics value to denote the significance level at which the coefficient differs significantly from zero. The t-value is calculated as $\frac{\bar{X}}{\sigma/\sqrt{N}}$, where $\bar{X}$ is the mean of estimated parameters, σ is the standard deviation of estimated parameters, and $N = 800$ as the simulation procedure is iterated 800 times.

	Conventional Bass model		
	Design	Packaging and testing	Memory module
p	0.0218***	0.0136**	0.0853***
q	0.1499***	0.2233***	0.0569***
M	18,952,555***	20,455,760***	30,106,304***
R^2	99%	99%	89%
	Modified model		
	Design	Packaging and testing	Memory module
p	0.0291***	0.0875***	0.0001***
q_1	0.0853***	0.1039***	0.0394***
δ	0.7946***	2.4739***	1.1052***
R^2	99%	99%	92%

Notes:
1. *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level.
2. The potential amount of capital flowing from Taiwan to China, M, in the proposed modified model is assumed to be the maximum permitted amount (NT\$9.7 billion, NT\$28 billion and NT\$17 billion for the IC design, IC packaging and testing, and IC memory module industries.

other firms for all sub-industries were significantly different from zero. Consistent with clustering theory, all sub-industries demonstrated a considerable positive influence from internal communication. As long as Taiwanese IC firms clustered together and constructed factories in China, their suppliers established the channels needed in China. Thus, new IC firms got a free ride with regard to supply channels, resulting in a considerable number of firms imitating early adopters by setting up divisions in China.

Because the modified model incorporates profitability and FDI limits, the profitability coefficient δ is positive for IC design, packaging/testing and memory module sub-industries. Equation (4) was utilized to calculate the time-variant imitation coefficient q_t over time, $q_t = q_1^{\exp(EPS_t \times \delta)}$, such that a positive coefficient δ would indicate that capital flowing into China is negatively related to profitability. An increase in the earnings of Chinese branches reduces the infusion of capital from Taiwanese firms. Once Taiwanese IC firms decide to establish branches in China, they are no longer able to immediately liquidate assets related to property and facilities. Taiwanese firms send additional capital to China, despite the fact that new Chinese subsidiaries suffer losses. By contrast, profits from Chinese branches can provide adequate capital for their continued operation.

5.2. **Results of T-statistics tests.** T-tests were used to compare differences in clustering incentives between firms in upstream and downstream sectors along the IC value chains under a vertical disintegration framework. Table 3 shows the t-statistics for in-pairs differences in clustering incentives between the three IC sub-industries. Using the conventional Bass model, the incentive to imitate other firms was the strongest for firms in the IC packaging/testing industry. In addition, the difference in the imitation coefficients between the IC packaging/testing industry and other IC industries was significant at 1% levels, according to t-statistics results (Table 3). Because production processes are mature, new Chinese employees can be easily trained in the skills of production. Thus, firms in IC packaging/testing can employ laborers at lower wages than other IC industries, providing greater incentives for expansion into China. These empirical results are consistent with the tenets of the product life cycle theory.

A comparison of the coefficient of innovation incentive with that of the incentive to imitate other firms indicates that the latter is higher than the former using the conventional Bass model. The relationship $p < q$ holds and the difference $(q - p)$ is statistically significant for the IC design and packaging/testing industries in the conventional Bass model (Table 4). This suggests that the incentive to imitate other firms dominates the clustering dynamics of Taiwanese firms in China, further indicating that communications and intra-industry interactions are factors key to IC clustering in China. In the past, high economic risk and uncertainty in China prompted many Taiwanese firms to avoid investing in China; however, after learning of the experiences of other enterprises in China, they re-evaluated their tolerance for risk in the context of pooling resources through clustering in China. IC industrial clusters are gradually established after firms communicate with early adopters.

5.3. **Results of calculated parameters stability.** Parameter stability (Subsection 5.3) and forecasting accuracy (Subsection 5.4) were assessed for the conventional Bass model and our modified model. In theory, as the value of STAB1 increases, parameter stability increases, and as the value of STAB2 decreases, parameter stability increases (i.e., STAB2 captures instability). Table 5 compares the parameter stability of the conventional Bass model and the proposed modified model. Innovation incentive parameter, p, captures a firm's initiative to cluster in China, while parameter q_1 in the modified model represents the degree to which enterprises imitate companies by investing capital in China. The

TABLE 3. Results of t-statistics for comparing in pairs differences in clustering incentives between the three IC industries: the IC design industry, IC packaging and testing industry, and IC memory module industry. The value of t-statistics is calculated as $\frac{(\bar{X}_1-\bar{X}_2)}{\sqrt{\frac{S_1^2}{N}+\frac{S_2^2}{N}}}$, where $\bar{X}_1$ and $\bar{X}_2$ are the mean of estimated parameters for paired industries, S_1 and S_2 are standard deviations of estimated parameters, and $N = 800$ as the simulation is iterated 800 times.

Conventional Bass Model	p		q	
	Difference	t-statistics	Difference	t-statistics
Test vs. design	–0.0083	–0.3671	0.0734	12.4727***
Memory vs. design	0.0634	2.2017***	–0.0930	–10.0484***
Test vs. memory	–0.0717	–2.5361***	0.1664	17.7330***
Modified Model	p		q_1	
	Difference	t-statistics	Difference	t-statistics
Test vs. design	0.0584	1.9514**	0.0185	2.9331***
Memory vs. design	–0.0290	–1.6502*	–0.0460	–14.5071***
Test vs. memory	0.0874	3.6069***	0.0645	11.3249***

Notes: *Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level.

Table 4. T-statistics results of the difference between the imitation and innovation coefficients, $(q-p)$, within the three industries by the conventional Bass model and proposed modified model. The t-statistics value is calculated as $\frac{(\bar{q}-\bar{p})}{\sqrt{\frac{S_q^2}{N}+\frac{S_p^2}{N}}}$, where $\bar{q}$ and $\bar{p}$ are the mean of the estimated imitation and innovation incentives, q and p, respectively; S_p and S_q are the standard deviations of estimated parameters q and p, respectively and $N = 800$ as the simulation procedure is iterated 800 times.

	Conventional Bass model		Modified model	
	Difference	t-statistics	Difference	t-statistics
Design	0.1281	7.5942***	0.0562	3.1554***
Packaging and testing	0.2097	13.1141***	0.0164	0.6574
Memory module	–0.0284	–1.1303	0.0393	34.9394***

*Significant at the 10% level. **Significant at the 5% level. ***Significant at the 1% level.

Table 5. The analytical results of parameters stability indicators, *STAB1* and *STAB2*, for the conventional Bass model and the modified model for the three IC industries: the IC packaging and testing industry, IC memory module industry and IC design industry. Parameter stability increases as *STAB1* increases and *STAB2* decreases. Parameters of innovation incentive, p, profitability coefficient, δ, and potential market, M, are estimated systematically. With the conventional Bass model, imitation coefficient, q, is assumed constant, while time-variant q_t is specified in the modified model, which incorporates profitability factors, namely, $q_t = q_1 \times (P_0 - P_t)^{\delta}$.

	Modified model					
	Stab1_p	*Stab1_*q_1	*Stab1_*δ	*Stab2_p*	*Stab2_*q_1	*Stab2_*δ
Design	50.3983	22.6588	1.6816	0.0158	0.0632	0.9329
Memory	0.5160	46.7137	8.6736	2.2775	0.0254	0.1041
Test	7.9622	10.8815	4.5248	0.1264	0.1180	0.3076
	Conventional Bass model					
	Stab1_ p	*Stab1_q*	*Stab1_m*	*Stab2_ p*	*Stab2_q*	*Stab2_m*
Design	18.3444	8.1207	9.6698	0.0327	0.0750	0.0630
Memory	0.4691	59.0270	5.1503	2.3843	0.0164	0.2236
Test	1.8738	9.3591	8.3338	0.3167	0.0635	0.0715

conventional Bass model does not discern the degree to which profits made by China branches prompt Taiwanese enterprises to decrease their investment in China. Briefly, q_1 represents the imitation of other firms without considering the influence of profit. The function of parameter q in the conventional Bass model differs from that of parameter q_1 in the modified model; therefore, a direct comparison of parameter stability (*STAB1* or *STAB2*) between parameters q and q_1 is theoretically meaningless. In the IC memory module and IC packaging/testing sub-industries, *STAB1* of parameter p is greater for the modified model than for the Bass model; *STAB2* of parameter p is smaller for the modified models than for the conventional Bass model. In the IC design industry, *STAB1* of parameters p and q_1 are greater for the modified model than for the conventional Bass model; *STAB2* of parameters p and q_1 are smaller for the modified model than for the conventional Bass model. In summary, for the IC memory module and packaging/testing sub-industries, parameter p of the proposed model is more stable than that of the conventional Bass model; for the IC design industry, parameters for imitation and innovation incentives in the proposed model are more stable than those of the conventional Bass model. These results further demonstrate the superior stability of the proposed model, which takes into account factors of profitability, for interpreting the motivation to imitate other firms in industrial clustering.

5.4. **Analysis of prediction accuracy.** This study further assessed the accuracy of the conventional Bass model and modified model in predicting IC clustering from Taiwan to China. Parameters of both models were estimated using quarterly FDI in China from Taiwan during the training period. After parameter estimation, the predicted amounts of FDI were calculated using estimated parameters. Both the predicted and actual flow of capital into China increased in both of the models – the modified model (Equation (5)) and the conventional Bass model (Equation (3)). The increased flow of capital into China also demonstrates that the tendency toward industrial clustering is transferred from Taiwan to China. Moreover, this evidence refutes the claim that the Taiwanese government has closed its doors to international transactions or prohibited its enterprises from investing in China.

Next, this study compared forecasted quarterly FDI in the test period with actual quarterly FDI. Figures 2-4 compare actual shipments versus the median of the 800 parameter estimates using the conventional Bass and modified models. Forecast results using our modified model displayed a trend similar to that of real time-series data related to capital flowing from Taiwan to China. Conversely, forecast results using the conventional Bass model differ from those of real time-series data. These results were consistent for all three of the sub-industries in this study.

This study also measured forecasting error for each model using MAPE (Table 6). As shown in Table 6, the MAPE of the Bass model is 0.1261, 0.1461 and 0.0629 for the IC design, packaging/testing, and memory module sub-industries, respectively. The MAPE of the modified model is only 0.0832, 0.0979 and 0.0186 for the IC design, packaging/testing, and memory module sub-industries, respectively. The MAPE of the Bass model is nearly

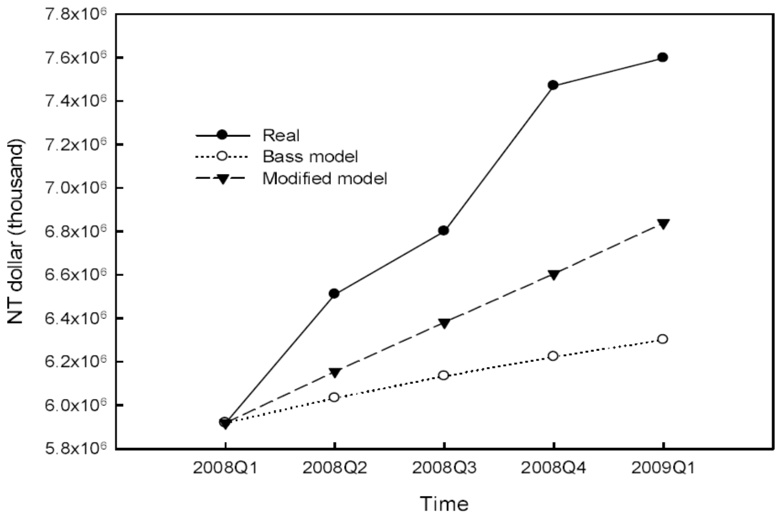

Figure 2. Comparison of real capital amounts invested in China by the Taiwanese IC design industry with predicted amount by the conventional Bass model and modified model, which incorporates profitability and the regulatory FDI limits

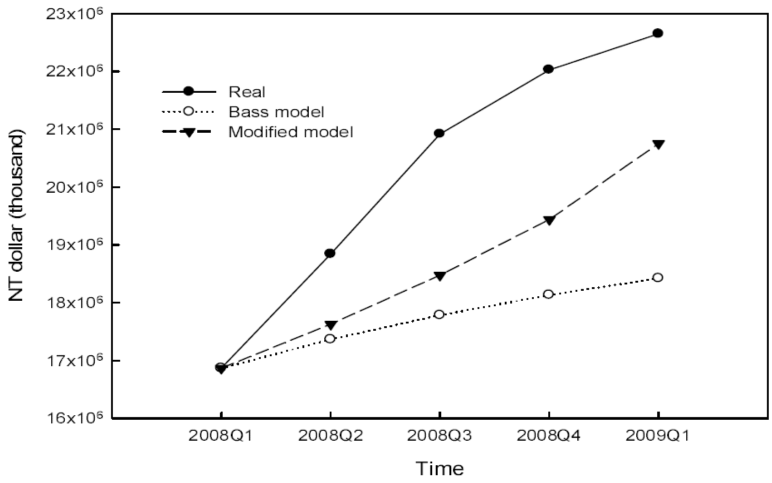

FIGURE 3. Comparison of real capital amounts invested in China by the Taiwanese IC packaging and testing industry with predicted amount by the conventional Bass model and modified model, which incorporates profitability and the regulatory FDI limits

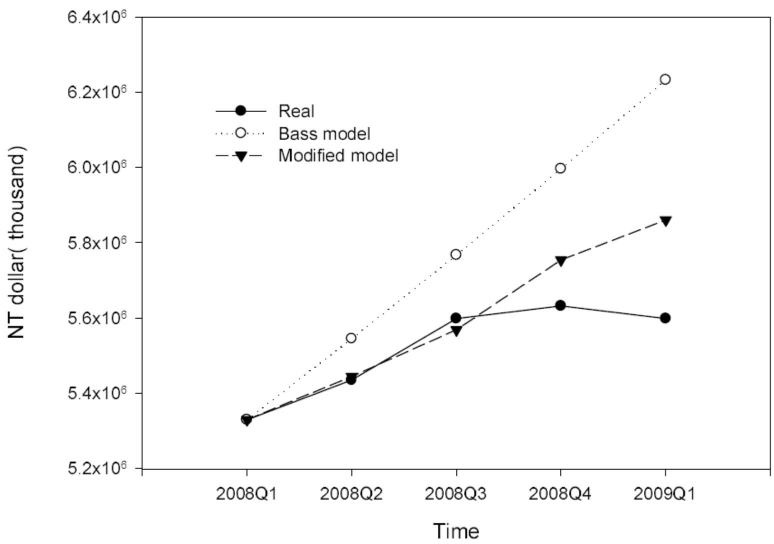

FIGURE 4. Comparison of real capital amounts invested in China by the Taiwanese IC memory module industry with predicted amount by the conventional Bass model and modified model, which incorporates profitability and the regulatory FDI limits

1.5 times that of the modified model for firms in IC design and packaging/testing and more than 2 times higher for firms dealing with IC memory modules. Comparison results suggest that prediction accuracy improves dramatically when regulatory limits and profitability are taken into account. We strongly recommended using the proposed modified diffusion model, which incorporates profitability and regulatory constraints, when analyzing the evolution of clustering from Taiwan to China.

6. **Conclusions.** This work focuses on the extension of clustering by IC firms from Taiwan into China between the fourth quarters of 1998 to the first quarter of 2009. We employed a GA to perform numerical simulations to clarify the impact of profitability on the clustering behavior of IC firms. The proposed modified diffusion model takes into

Table 6. Mean absolute percentage error (MAPE) for the conventional Bass model and our modified model from the second quarter of 2008 to the first quarter of 2009 in IC design, IC packaging and testing and IC memory module industries

MAPE	IC design	IC packaging and testing	IC memory module
Modified model	0.0832	0.0979	0.0186
Conventional Bass model	0.1261	0.1461	0.0629

account the influence of innovation, imitation, and profit on IC clustering dynamics. This study then compared the parameter stability and forecasting accuracy of our modified model with those of the conventional Bass model.

Empirical results indicate that the flow of capital from Taiwanese IC firms into China has increased over time, demonstrating that industrial clusters in China are expanding. Earnings obtained from Chinese subsidiaries may have reduced the flow of capital into China; however, investment continues to flow. Analysis results prove that "Regulations Governing the Approval of Investment or Technical Cooperation in Mainland China" has not deterred Taiwanese IC firms from developing production centers in China. The tendency to cluster in China depends more on the strategic goals of firms than on governmental regulations. Our results also demonstrate that the impact of intra-industry communication is a dominant factor in the evolution of clustering in China. Taiwanese IC firms tend to imitate others when clustering in China. Finally, t-statistics indicate that the incentive to imitate other firms is more pronounced in the IC packaging/testing sub-industry than in other sectors of the IC value chain. Consistent with product life cycle theory, IC packaging/testing firms are strongly inclined to imitate firms in transferring production facilities to China, because Chinese employees can rapidly acquire the low operational skills required for IC packagers and testers.

Parameter stability results indicate that the modified model, which takes into account the influence of profitability and regulatory limits on the evolution of clustering, is more stable with regard to coefficient estimation than the conventional Bass model for all three IC sub-industries. This again demonstrates the reliability of out modified model for interpreting the evolution of IC clustering. In terms of forecasting accuracy, prediction errors were much smaller for the modified model than for the conventional Bass model. Parameter stability and prediction error analysis strongly support the importance of considering profitability and regulatory FDI limits.

The contribution of this research includes the development of a novel diffusion model through which academics and practitioners can better understand the characteristics of industrial clustering beyond what was possible through previous studies [4,25]. The proposed diffusion model resembles a knowledge-based system capable of transforming value-added data related to FDI into parameter values that describes the features of IC industrial clusters. Our novel diffusion model represents a major step forward in the fields of knowledge discovery and management science, highlighted by the following four aspects: unique data, model specification, parameter optimization, and applicability in industrial analysis. In terms of unique data, the novel diffusion model utilizes FDI as a quantitative indicator for measuring the extension of industrial clusters. Most countries lack data related to FDI in specific regions, but the Taiwan Economic Journal (TEJ) database has compiled data on real FDI implemented by Taiwanese IC firms in China.

This unique data was utilized to quantify the IC clustering extent as the basis of strategic management.

This study is the first to construct a model incorporating the impact of profit on the development of industrial clusters. The Chinese branches in capital-intensive IC industries suffering serious losses in the early stages are more likely to follow other companies' FDI decisions determine the amount of funds to further devote to their Chinese subsidiaries. Thus, the profitability of IC industries is closely related to clustering. In contrast to the conventional Bass model used in previous researches, we eliminated the restrictive assumption that IC clustering is unrelated to profit. Our results demonstrate the superior accuracy of the proposed model (incorporating profitability) over that of the conventional Bass model.

As for parameter optimization, we combined a genetic algorithm (GA) and the fourth-order Runge-Kutta algorithm to optimize parameters to avoid the shortcomings of previous studies using NLS, in which poor initial values converge to a local instead of global solution. Our empirical results demonstrate that the standard deviation of the estimated parameters is very small, confirming that the optimized parameters remain similar, despite the fact that 800 different sets of initial values were used in our GA simulations. This finding suggests that the proposed approach generates stable, reliable parameters with very little deviation. Furthermore, the parameters using different sample periods optimized using the proposed model are more similar than those optimized using the conventional Bass model, indicating that the proposed model is more stable in terms of parameter optimization.

As for applicability in industrial analysis, this work compares differences in clustering incentives among the sub-industries of IC design, packaging/testing, and memory modules under the vertical disintegration structure. According to our results, the IC packaging/testing sub-industry is the most likely to imitate other firms in moving facilities from Taiwan to China. This has not been statistically demonstrated in previous studies. The significant difference between various sub-industries with regard to incentives to cluster in China can be attributed to their different skill level and technical requirements. Our analytical results provide a reference for upstream IC equipment suppliers and downstream electronics manufacturers in establishing global strategic goals.

In addition, our proposed modified model could be applied to forecast the evolution of clusters in other high-tech industries. China is currently the world's factory as well as the world's largest market; therefore, other high-tech industries, such as TFT-LCD, computers, communications, information technology, and consumer electronics are investing heavily in China. These industries display the same characteristics as the IC industry with regard to capital overhead and long-term cost recovery. Thus, the proposed model is applicable to forecasting industrial clustering in other industries. In addition, the phenomenon of clustering is not limited to China, and this approach could be applied in other regions of the world. This work makes a considerable contribution to the applications related to knowledge discovery, knowledge-based systems, management science, and industrial analysis.

Acknowledgment. The author would like to thank National Science Council, Taiwan, for financially supporting this research under Contract No. NSC-99-2410-H-009-016-MY3. The author also gratefully acknowledges the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] B.-H. Tsai, Does litigation over the infringement of intellectual property rights hinder enterprise innovation? An empirical analysis of the Taiwan IC industry, *Issues and Studies*, vol.46, no.2, pp.173-203, 2010.

[2] B.-H. Tsai, Exploring the influence of ERP implementation on corporate performance using a modified DEA approach, *African Journal of Business Management*, vol.5, no.14, pp.5435-5488, 2011.

[3] M. E. Porter, Clusters, and the "new" economics of competition, *Harvard Business Review*, vol.33, pp.77-90, 1998.

[4] B. T. McCann and T. B. Folta, Who enters, where and why? The influence of capabilities and initial resource endowments on the location choices of de novo enterprises, *Journal of Management*, vol.34, pp.532-561, 2008.

[5] F. M. Bass, A modified product growth model for consumer durable, *Management Science*, vol.15, pp.215-227, 1969.

[6] B.-H. Tsai, Y. Li and G.-H Lee, Forecasting global adoption of crystal display televisions with modified product diffusion model, *Computers and Industrial Engineering*, vol.58, no.4, pp.553-562, 2010.

[7] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.

[8] M. Gen and R. Chang, *Genetic Algorithms and Engineering Optimization*, Wiley, New York, 2000.

[9] R. L. Burden and J. D. Faires, *Numerical Analysis*, 8th Edition, Thomson Corporation, Boston, MA, USA, 2005.

[10] P. L. Chang and C. T. Tsai, Finding the niche position-competition strategy of Taiwan's IC design industry, *Technovation*, vol.22, no.2, pp.101-111, 2002.

[11] TEEMA, *A Survey Report on the Evaluation of Investment Environment and Risks in China*, Taipei, 2001.

[12] R. Vernon, International investment and international trade in the product cycle, *Quarterly Journal of Economics*, vol.80, pp.190-207, 1966.

[13] M. E. Porter, The Adam Smith address: Location, clusters, and the "new" microeconomics of competition, *Business Economics*, vol.33, pp.7-13, 1998.

[14] M. E. Porter, Location, competition, and economic development: Local clusters in global economy, *Economic Development Quarterly*, vol.14, pp.15-34, 2000.

[15] D. M. Smith, *Industrial Location: An Economic Geographical Analysis*, 2nd Edition, John Wiley, New York, 1981.

[16] C.-K. Lee, How does a cluster relocate across the border? The case of information technology cluster in the Taiwan-Suzhou region, *Technological Forecasting and Social Change*, vol.76, pp.371-381, 2009.

[17] E. M. Rogers, *Diffusion of Innovations*, 5th Edition, Free Press, New York, 2003.

[18] N. Meade and T. Islam, Modelling and forecasting the diffusion of innovation: A 25-year review, *International Journal of Forecasting*, vol.22, pp.519-545, 2006.

[19] E. Ruiz and P. S. H. Leeflang, Diffusion of franchising as an innovation of managerial organization, *Marketing JRM*, pp.65-75, 2006.

[20] S. Kalish and S. K. Sen, Diffusion models and the marketing mix for single products, in *Innovation Diffusion Models of New Product Acceptance*, Mahajan and Y. Wind (eds.), Ballinger, MA, Cambridge, 1986.

[21] R. Venkatesan, T. V. Krishnan and V. Kumar, Evolutionary estimation of macro-level diffusion models using genetic algorithms: An alternative ton nonlinear least squares, *Management Science*, vol.23, pp.451-464, 2004.

[22] B.-H. Tsai, Predicting the diffusion of LCD TVs by incorporating price in the extended Gompertz model, *Technological Forecasting and Social Change*, vol.80, no.1, pp.106-131, 2013.

[23] J. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge, MA, USA, 1975.

[24] P. N. Golder and G. J. Tellis, Beyond diffusion: An explanatory approach to modeling the growth of durables, *Journal of Forecasting*, vol.17, pp.259-280, 1998.

[25] B.-H. Tsai and Y. Li, Cluster evolution of IC industry from Taiwan to China, *Technological Forecasting and Social Change*, vol.76, pp.1092-1104, 2009.