An e-commerce performance assessment model: Its development and an initial test on e-commerce applications in the retail sector of China

Jinghua Huang a,*, Ximin Jiang b, Qian Tang c

a Research Center for Contemporary Management, School of Economics & Management, Tsinghua University, Beijing 100084, China
b HongKong and Shanghai Banking Corporation (HSBC), China
c McCombs School of Business, The University of Texas at Austin, Austin, TX 78712, USA

1. Introduction

Every year, various groups rank enterprises according to the effectiveness of their e-commerce (EC) efforts; for example, the National Informatization Evaluation Center of China offers an evaluation indicator system with six criteria (IT strategy, IT infrastructure, IT implementation, human resources, information security, and IT benefits) that it uses to evaluate companies according to their EC applications [11]. Other organizations use different metrics and models. Though some studies have addressed the impact of EC on business performance, none has developed an EC performance measurement model. To understand the impact of EC on firms, we need effective metrics to measure it and a strong model to evaluate its effect.

This need is particularly acute for retail industry because it has substantial potential benefits. Retailers deploy B2B EC in their value chains to manage inventory and procurement, strengthen their relationship with distributors, increase sales, obtain access to new markets, improve customer service, reduce cost, and compete more effectively. Various ITs, such as point-of-sale and back-end systems and management software, constitute the primary ways of digitizing business processes. At the same time, retailing is developing rapidly and especially in China. By the end of 2006, total retailing in social commodities in China increased by 13.7% to 7641 billion RMB Yuan. As a result, fixed asset investment in China’s retail industry increased; by 41.3% in 2004, 37.1% in 2005, and 23% in 2006. But as the industry expanded, it experienced radical change; most Chinese retailers now employ a combination of online and offline offerings. However, many managers still remain unaware of EC potential benefits and therefore are hesitant to make significant investment in it. Therefore, a model evaluating EC effectiveness can provide useful feedback that determines weaknesses in system implementation, such as poor customer service.

We carefully considered the question of how to measure EC performance by addressing the following questions:

- What indicators, and their weights are appropriate?
- What evaluation methods are available to calculate EC performance?
- What is a reasonable EC performance assessment model?
- What is the relative EC performance in the retail sector of China?

We drew on literature pertaining to EC value, majority aggregation processes for indicator weights, and multiattribute decision-making.
methods (MADM) in answering these questions and developed an EC performance assessment model that included a set of metrics, a majority aggregation process method based on a majority additive-ordered weighted averaging operator (MA-OWA), and an evaluation method containing seven sub-methods and comparison criteria.

2. Literature review

Previous IS and EC performance assessment research included performance measures and methods that relied on divergent points of view such as: Zhuang and Lederer [21] proposed a 27-item instruments to measure business benefits of EC retailing, whereas Barua et al. [2] examined operational and financial performance data, DeLone and McLean used six dimensions to measure IS success (systems quality, information quality, use, user satisfaction, individual impact, and organizational impact) and later updated to deal with e-commerce success [3,4], and other researches [7,10]. Because different EC indicators contribute differently to a firm’s overall EC performance, research had to take indicator weights into account. Furthermore, various approaches employed in group decision making aggregated multiple persons’ preferences, such as OWA, OWG (ordered weighted geometric averaging), and MA-OWA. Among them, these methods have been used in studies that collect subjective data about experts’ preferences. MA-OWA operators can aggregate personal subjective information collected using Likert scales that make the questionnaire simpler, clearer, and easier to understand.

Researchers also employed several evaluation methods within MADM, such as simple additive weighting (SAW), multiplicative exponent weighting (MEW), the technique for order preference by similarity to ideal solution (TOPSIS) [19], etc., grey relational analysis (GRA), concordance analysis, and fuzzy comprehensive evaluation. Some multivariate analysis methods also apply to comprehensive evaluations [6,18]. Data envelopment analysis (DEA) can evaluate the efficiency of alternatives [16].

Studies investigating applications of MADM included those of Stewart and Mohamed [14], who applied a tiered IT balanced scorecard (BSC) construct to IS evaluation systems by deriving indicator weights using AHP and obtaining utility values on the basis of multiattribute utility theory and Schubert [13], who developed an extended Web assessment method to evaluate EC Web sites using simple added, etc. In general, because of the complexity of any evaluation situation, researchers have been forced to select several methods, compare their results, and use the most effective method, according to some evaluation criteria.

3. Performance assessment model for EC

We considered EC performance as being measured as the overall organizational impact after the enterprise has implemented EC. Our performance assessment model is shown in Fig. 1, which consisted of four steps: (1) establishing performance indicators for EC, (2) calculating indicator weights, (3) calculating assessment scores using seven methods, and (4) selecting the optimal evaluation methods.

3.1. Establishing performance indicators for EC

Based on [5,9,15], we proposed several EC performance indicators and hypotheses that reflected the relationships among the indicators. The financial benefits of EC, such as increased profitability, increased revenue, reduced expenses, and increased turnover, represent the purpose of a company’s implementation of EC, but these are not the only benefits. Nonfinancial benefits also should be considered in the model.

EC has an impact on marketing and sales, customer service, and supply chain efficiency. Because the goal of retailing is to sell products, marketing and sales remains the primary motive. EC can open a new market, lower the costs of entering it, and make competition transparent. Furthermore, EC enables a company to understand customer demands and thus provide effective service quickly and accurately; the Internet also allows companies to develop and improve their relationships with suppliers. Finally, EC helps digitalize consumer transactions with efficient supply chains to reduce inventory levels and deliver products to customers on time. We therefore proposed 16 indicators of benefits, categorized into four groups: marketing and sales, customer service, supply chain efficiency, and financial performance.

3.2. Calculating indicator weights

Because the objective weighting method is influenced by indicator scores, the weights tend to be stochastic and inconsistent. Therefore, we adopted an experts’ scoring-based subjective weighting method, which involved three steps to determine the weights:

(1) Gathering experts’ score of each indicator’s significance.
(2) Synthesizing experts’ judgment scores and obtaining each indicator’s aggregate value.
(3) Normalizing all aggregate values to obtain the weight of each indicator.

Experts should have both EC theoretical knowledge, and practical experience. CIOs were therefore qualified as experts. We adopted MA-OWA operators for the aggregate processes to calculate the aggregate values [12]. This method assigned greater weights to experts who shared the same opinions, thus mitigating the influence of those who held a different opinion. The majority aggregation process therefore involved four steps:

(1) Ordering the experts’ judgment scores and classifying them into groups; this placed those with the same scores in one group.
(2) Selecting an element from each group and aggregating and averaging the elements to create a new group with only one element that represents their average. Then delete the element from its original group.
(3) Eliminating groups with no element left.
(4) Repeating steps 2 and 3 until only one group with one element remained.

This final element provided the majority aggregation value. The MA-OWA operator can realize this efficiently by the procedure:

\[ F_{MA}(a_1, a_2, \ldots, a_n) = \sum_{j=1}^{n} \omega_j b_j, \]  

where \( n \) is the number of experts, \( a_i \) is the \( i \)th expert’s judgment score, \( F_{MA}(a_1, a_2, \ldots, a_n) \) is the majority aggregation value, and \( b_j \) is the \( j \)th largest of the \( a_n \), and

\[ \omega_j = f_j(b_1, b_2, \ldots, b_n) = \frac{1}{\prod_{k \neq j} g_k(b_1, b_2, \ldots, b_n)}, \]

where \( g_k \) is a function that indicates when the aggregation process uses \( b_j \) and represents the score’s sequence number in the ascending classified group.

In Eq. (2), \( h_k \) is a function that indicates the number of elements in each step in the aggregation process.
We then calculate \( h_k \) as

\[
h_k(b_1, b_2, \ldots, b_n) = \begin{cases} 
\sum_{j=1}^{n} p_{kj} & \text{if } k = 1 \\
\sum_{j=1}^{n-k+1} p_{kj} + 1 & \text{otherwise } k \neq 1
\end{cases}
\]  

(3)

\[
p_{kj} = \begin{cases} 
1 & \text{if } b_j = b_{j+k-1} \text{ and } b_j \neq b_{j-1} \text{ and } j \neq 1 \\
1 & \text{if } j = 1 \text{ and } k \geq 1 \text{ and } b_j = b_k \\
0 & \text{otherwise}
\end{cases}
\]  

(4)

Finally, after obtaining every indicator’s \( F_{MA} \) \((a_1, a_2, \ldots, a_n)\), we calculate its weight as

\[
w_i = \frac{F_{MA,i}}{\sum_{j=1}^{q} F_{MA,j}} \quad i = 1, 2, \ldots, q.
\]  

(5)

where \( q \) is the number of indicators and \( w_i \) is the \( i \)th indicator’s weight.

3.3. Calculate assessment scores

After weighting the indicators, we calculated the general assessment scores for the EC applications of the target enterprises. Specifically, we considered and compared seven evaluation methods: SAW, MEW, TOPSIS, GRA, concordance analysis, FA-PCA, and DEA, of which the first four are the most common: they are convenient to apply. Concordance analysis is also popular when selecting optimal alternatives. Among the multivariable statistical analysis methods, factor and principal component analyses have received most attention: the FA-PCA method has the advantages of both. However, traditional evaluation methods cannot avoid subjectivity; therefore, DEA tries to determine an optimal weight assignment for each unit [1]. Details of the seven methods are given in Appendix A.

3.4. Choose optimal evaluation methods

After obtaining assessment scores for EC performance, we conducted an in-depth comparative analysis and thereby selected...
the optimal methods. Our comparative analysis relied on four criteria. The first pertained to the correlations among the seven methods. The second required correlations of every evaluation method with the overall evaluation or the score that a respondent provided to override it; we used the Spearman rank correlation coefficient to calculate both of these correlations [20]. The third referred to the compatibility of each evaluation method, that is, the average of the rank correlation coefficients between the assessment scores of the focal method and all other methods. Finally, the fourth entailed computing the difference between each evaluation method, or the average number of enterprises that fell outside the rank range that resulted from all other evaluation methods.

In our study, we chose compatibility as the primary criterion for selecting evaluation methods, because we focused on overall ranking results.

4. Performance assessment model application

4.1. Instrument design and data collection

To test the proposed EC performance assessment model, we designed a questionnaire that featured our performance indicators and collected data from the retail sector of China. We reviewed each item in the questionnaire for content validity. The first part of the questionnaire pertained to the respondent’s current status, whereas the second referred to the company’s current EC status, including the number of IT staff, the number of years since the inception of its EC effort, and the main function of the company’s EC systems. The third part consisted of two kinds of questions designed to gather the 16 performance indicators by investigating both the significance of each indicator and the actual effect of EC, as shown in Appendix B. Using a five-point Likert scale from 5 (very important) to 1 (very unimportant), we assessed each indicators’ significance. To measure the actual application effect of EC, we used a Likert scale anchored by 5 (Strongly agree) and 1 (Strongly disagree). Finally, the fourth part contained only one question about the respondents’ overall evaluation of EC in an enterprise, again scored on a five-point Likert scale.

We conducted the survey in April and May 2006, mailing 240 questionnaires to randomly chosen traditional retailers engaged in EC; we received 70 valid responses from companies using EC, for a response rate of approximately 30%. The firm size, measured by the number of employees, reflected a balance of large and small businesses, and the distribution of firm locations covered the east, west, north, and south regions of China (see Table 1). Thus, respondents reflected the distribution of retailers across China. To minimize potential bias, we assured respondents that their responses and identity would remain confidential and that we would provide only aggregate information as feedback to respondents, without any identifying details. The eligible respondents were those persons best qualified to speak about the firm’s overall EC situation, normally CIOs, IS/IT managers, or business

<table>
<thead>
<tr>
<th>Measures</th>
<th>Exploratory factor analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales area</td>
<td>0.65</td>
</tr>
<tr>
<td>Marketing cost</td>
<td>0.78</td>
</tr>
<tr>
<td>Sales expenses</td>
<td>0.78</td>
</tr>
<tr>
<td>Market share</td>
<td>0.71</td>
</tr>
<tr>
<td>Customer satisfaction</td>
<td>0.83</td>
</tr>
<tr>
<td>More services to customers</td>
<td>0.90</td>
</tr>
<tr>
<td>Customer service expenses</td>
<td>0.78</td>
</tr>
<tr>
<td>Corporate image</td>
<td>0.76</td>
</tr>
<tr>
<td>Inventory turnover</td>
<td>0.77</td>
</tr>
<tr>
<td>Collaboration with partners</td>
<td>0.70</td>
</tr>
<tr>
<td>Procurement cost</td>
<td>0.75</td>
</tr>
<tr>
<td>Procurement expenses</td>
<td>0.90</td>
</tr>
<tr>
<td>Sales revenue</td>
<td>0.57</td>
</tr>
<tr>
<td>Gross profit</td>
<td>0.62</td>
</tr>
<tr>
<td>Net profit to operational expenses ratio</td>
<td>0.86</td>
</tr>
<tr>
<td>Working capital turnover</td>
<td>0.84</td>
</tr>
<tr>
<td>KMO</td>
<td>0.86</td>
</tr>
<tr>
<td>Bartlett’s value</td>
<td>1130 (p &lt; 0.0001)</td>
</tr>
<tr>
<td>Total variance explained</td>
<td>82.0%</td>
</tr>
</tbody>
</table>

Spearman rank correlation coefficients among latent variables

| Financial performance | 1.00 |
| Marketing and sales | -0.03 |
| Customer service | 0.12 |
| Supply chain efficiency | -0.03 |
| Marketing and sales | 1.00 |
| Customer service | 0.14 |
| Supply chain efficiency | 0.08 |
| Financial performance | 1.00 |
| Marketing and sales | 1.00 |
| Customer service | 1.00 |
| Supply chain efficiency | 1.00 |
managers. The positions of the respondents reflected the strong quality of this data source.

4.1.1. Instrument validation

We used exploratory factor analysis (EFA) to test for construct validity. According to the results, we could extract four factors (see Table 2), which mirrored the initial four aspects. This showed that performance measures for EC consisted of marketing and sales, customer service, supply chain efficiency, and financial performance. We further calculated the correlations among these four factors; the results indicated no significant correlations.

4.1.2. Indicator weights in the retail sector

We next used the conceptual model as the basis on which we conducted an EC performance assessment of the enterprises. To aggregate the judgment of the respondents, we use MA-OWA operators to obtain the weight of every indicator (see Table 3). The top five, in terms of weight, were “corporate image”, “sales area”, “more services to customers”, “customer satisfaction”, and “collaboration with partners”, similar to the result derived from a performance assessment of Chinese enterprises in the software industry, where the top four indicators were “corporate image”, “marketing cost”, “customer satisfaction”, “more services to customers” [8]. This suggests that the MA-OWA was valid.

The weighting result indicated that Chinese retailers mainly wanted to upgrade their corporate image, expand their enterprise sales area, and facilitate communications with their business partners. However, they do not attach much importance to EC financial benefits.

4.2. Performance assessment of EC

After obtaining the weight of every indicator, we calculated the EC performance assessment scores and thus derived the scores for all the enterprises. The three highest scoring ones are shown in Table 4. According to concordance analysis, the enterprise with the maximum priority index also achieved the minimum discordance priority index. Generally speaking, we found no significant differences among the evaluation results based on the seven methods. Enterprise 28 scored highly on every indicator; that is, its EC performance assessment score placed it at the top. Other enterprises may use it as a benchmark for their own EC practices.

4.3. Comparison of evaluation methods

In Table 5, we show the Spearman rank correlation coefficients among the assessment scores for all seven methods; all are significant ($p < 0.001$). This showed that no significant differences existed among the methods, and the ranking order by SAW was the same as that established by the discordance priority index.

Furthermore, with the assumption that the overall evaluation scores offered by the respondents were reliable, we explored the correlations among the assessment scores provided by each method.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Indicator weights.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construct</td>
<td>Indicator</td>
</tr>
<tr>
<td>Marketing and sales</td>
<td>Sales area</td>
</tr>
<tr>
<td></td>
<td>Marketing cost</td>
</tr>
<tr>
<td></td>
<td>Sales expense</td>
</tr>
<tr>
<td></td>
<td>Market share</td>
</tr>
<tr>
<td>Customer service</td>
<td>Customer satisfaction</td>
</tr>
<tr>
<td></td>
<td>More services to customers</td>
</tr>
<tr>
<td></td>
<td>Customer service expense</td>
</tr>
<tr>
<td></td>
<td>Corporate image</td>
</tr>
<tr>
<td>Supply chain efficiency</td>
<td>Inventory turnover</td>
</tr>
<tr>
<td></td>
<td>Collaboration with partners</td>
</tr>
<tr>
<td></td>
<td>Procurement cost</td>
</tr>
<tr>
<td></td>
<td>Procurement expense</td>
</tr>
<tr>
<td>Financial performance</td>
<td>Sales revenue</td>
</tr>
<tr>
<td></td>
<td>Gross profit margin</td>
</tr>
<tr>
<td></td>
<td>Net profit to operational expenses ratio</td>
</tr>
<tr>
<td></td>
<td>Working capital turnover</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Top three enterprises’ assessment score under different methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of enterprise</td>
<td>SAW</td>
</tr>
<tr>
<td>28</td>
<td>4.85</td>
</tr>
<tr>
<td>60</td>
<td>4.72</td>
</tr>
<tr>
<td>67</td>
<td>4.68</td>
</tr>
</tbody>
</table>

Note: $p = 0.8$.

$^a$ The discordance priority index is the value in the table divided by $M$ (the maximum number of indicators in the discordance sets). This did not influence the rank order of the enterprises.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Spearman rank correlation coefficients between assessment scores of different methods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>0.98***</td>
</tr>
<tr>
<td>C</td>
<td>0.99***</td>
</tr>
<tr>
<td>D</td>
<td>0.98***</td>
</tr>
<tr>
<td>E</td>
<td>0.99***</td>
</tr>
<tr>
<td>F</td>
<td>1.00***</td>
</tr>
<tr>
<td>G</td>
<td>0.94***</td>
</tr>
</tbody>
</table>

Note: A = SAW; B = MEW; C = TOPSIS; D = GRA; E1 = concordance priority index; E2 = discordance priority index; F = DEA; G = FA-PCA. $^p < 0.001$. The discordance priority index is the value in the table divided by $M$ (the maximum number of indicators in the discordance sets). This did not influence the rank order of the enterprises.
method and the overall evaluation scores. All correlation coefficients between the method-specific assessment scores and the overall evaluation scores were approximately 0.65 and significant (p < 0.001) (see Table 6). The concordance priority index method indicated the maximum coefficient, and the GRA method revealed the minimum coefficient.

We show the degree of compatibility and difference in each evaluation method in Table 7. Three methods – SAW, discordance priority index, and TOPSIS – had the greatest degree of compatibility, which indicated that they are more representative and reliable. Therefore, we considered these three approaches to be the optimal evaluation methods in terms of the degree of compatibility. According to the degree of difference, three methods – SAW, discordance priority index, and FA-PCA – had the lowest degree, making them the optimal methods from the degree of difference perspective.

If an evaluation emphasized an overall ranking order, the degree of compatibility provided the best criterion; however, if it focused on a partial ranking order (e.g., to find the top 10 enterprises in EC performance), the degree of difference represented the best criterion. Therefore, when considering the degree of compatibility and difference, as well as correlation with the overall evaluation scores, we identified discordance analysis and SAW as the optimal evaluation methods for the enterprises. This result matched other’s belief that SAW is the best method for MADM [17].

### 5. Findings and discussion

By surveying EC activities of retailers in China, we applied our model to obtain the indicators and their weights for the EC performance scores for a set of Chinese retailers. Our study revealed that the weight of the financial performance indicator was the lowest of the constructs.

Instead, customer service attained the greatest weight, demonstrating that customer service and corporate image were the main drivers prompting China retailers to implement EC. Many retailers initiated EC because others had done so. In China, most retailers remain small or medium-sized, without brand naming; thus strengthening the company image can increase its business value. Through EC, firms can sell their products to customers who would have been unreachable by a traditional business.

We calculated the scores of EC for each company using seven methods and the results were relatively consistent among them. We note that absolute score did not necessarily mean much to a company; however, its comparison with the competition was key. Despite tremendous efforts to develop MADM models to solve different types of decision problems, the different methods still tend to produce divergent outcomes when it comes to selecting or ranking a set of decision alternatives having multiple attributes. Therefore, by including both the concordance and discordance priority indices, FA-PCA, and DEA, we produced four criteria that enabled us to select suitable methods. This analysis revealed that discordance analysis and SAW provided the optimal evaluation methods for the enterprises cooperating in our study.

### 6. Implications and limitations

Customer service was most important to Chinese retailers in evaluating their EC performance. Among all indicators of EC, they hoped to improve their corporate image, expand their sales area, and collaborate with partners. This called for a more comprehensive way to evaluate EC performance. We therefore provided a guideline by which institutes or governments could compare and rank companies on their EC performance.

However, our research was subject to some limitations. First, the firm performance data was self-reported and this could induce certain biases, though we employed several checks and balances during the data collection process. The questionnaire also offered a “do not know” choice for respondents unsure of the performance. Second, we tested our results only in the Chinese retail sector, whereas indicators and their weights vary among different industries and cultures. Our assessment model cannot therefore be generalized into other industries. Third, no standard exists for selecting evaluation methods, and the selection criterion must depend on the specific context.

### Appendix A

#### A.1. Details of the processes for the seven evaluation methods

\[ \Omega = \{X_1, X_2, \ldots, X_m\} \] is defined as the set of \( m \) enterprises’ EC to be evaluated, where \( X_i \) is the \( i \)th enterprises’ EC. \( U = \{u_1, u_2, \ldots, u_n\} \) is defined as the set of performance indicators, where \( u_j \) is the \( j \)th indicator. Thus, the original assessment matrix \( Z \) can be expressed as

\[
Z = \begin{bmatrix}
    x_{11} & x_{12} & \cdots & x_{1n} \\
    x_{21} & x_{22} & \cdots & x_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    x_{m1} & x_{m2} & \cdots & x_{mn}
\end{bmatrix},
\]

(A-1)

where \( x_{ij} \) represents the performance ratings of the \( i \)th enterprise’s EC with respect to the \( j \)th indicator (\( i = 1, 2, \ldots, m, j = 1, 2, \ldots, n \)). Then, \( r_{ij} \) is column normalized by \( x_{ij} \) as

\[
r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{l=1}^{m} x_{lj}^2}}.
\]

In turn, we get

\[
R = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1n} \\
    r_{21} & r_{22} & \cdots & r_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{m1} & r_{m2} & \cdots & r_{mn}
\end{bmatrix}.
\]

(A-2)
Assume that the weight of each indicator, \( W = \{ \omega_1, \omega_2, \ldots, \omega_n \} \), has been obtained through a weighting method. The evaluation methods we use to obtain the EC assessment scores \( Y_i \) for each enterprise are defined as follows:

1. **SAW**: The EC assessment score for each enterprise is obtained by
   \[ Y_i = \sum_{j=1}^{n} \omega_j r_{ij}; \quad i = 1, 2, \ldots, m. \]  
   \[ \text{(A-3)} \]

2. **MEW**: The EC assessment score for each enterprise is obtained by
   \[ Y_i = \prod_{j=1}^{n} r_{ij}^{\omega_j}; \quad i = 1, 2, \ldots, m. \]  
   \[ \text{(A-4)} \]

3. **TOPSIS**: The positive ideal solution is calculated as \( A^+ = \{ r_1, r_2, \ldots, r_n \} \), where \( r_j^+ = \max_{i \in \Omega} r_{ij} \), whereas the negative ideal solution is calculated as \( A^- = \{ r_1^-, r_2^-, \ldots, r_n^- \} \), where \( r_j^- = \min_{i \in \Omega} r_{ij} \).

   The weighted Euclidean distance between the original ratings for the \( i \)th enterprise and the positive and negative ideal solutions can be calculated, respectively, as
   \[ S_i^+ = \sum_{j=1}^{n} \omega_j (r_{ij} - r_j^+) ; \quad i = 1, 2, \ldots, m. \]  
   \[ \text{(A-5)} \]
   \[ S_i^- = \sum_{j=1}^{n} \omega_j (r_{ij} - r_j^-) ; \quad i = 1, 2, \ldots, m. \]  
   \[ \text{(A-6)} \]

   The EC assessment score for each enterprise is obtained by
   \[ Y_i = \frac{S_i^-}{S_i^- + S_i^+} ; \quad i = 1, 2, \ldots, m. \]  
   \[ \text{(A-7)} \]

4. **Concordance analysis**: For each pair of the \( i \)th enterprise’s EC and \( i’ \) (\( i = 1, 2, \ldots, m, i \neq i’ \)), the concordance set is defined as \( C_{ii'} = \{ j | x_{ij} > x_{i'j} \} \), or the set of indicators for which the \( i \)th enterprise’s EC is better than the \( i’ \)th enterprise’s EC. The concordance set is defined as \( D_{ii'} = \{ j | x_{ij} < x_{i'j} \} \), which is the set of indicators for which the \( i \)th enterprise’s EC is no better than the \( i’ \)th enterprise’s EC.

   In addition, the concordance and discordance matrices consist of \( c_{ii'} \) and \( d_{ii'} \), respectively, where
   \[ c_{ii'} = \sum_{j \in C_{ii'}} \omega_j ; \quad i = 1, 2, \ldots, m, \text{ and} \]  
   \[ d_{ii'} = \sum_{j \in D_{ii'}} \omega_j (x_{ij} - x_{i'j}) / (\max_{j \in C_{ii'}} x_{ij} - x_{i'j}) \];  
   \[ i = 1, 2, \ldots, m. \]  
   \[ \text{(A-8)} \]
   \[ \text{(A-9)} \]

   Therefore, \( c_{ii'} \) implies the sum of the relative degree of importance for the indicators in the concordance set and \( d_{ii'} \) implies the extent that the \( i \)th enterprise’s EC is worse than the \( i’ \)th enterprise’s EC; \( M \) is the maximum number of indicators in the discordance set.

   The concordance index for the \( i \)th enterprise’s EC is defined as
   \[ c_i = \sum_{i = 1}^{m} c_{ii'} - \sum_{i 
eq i'} c_{ii'}; \quad i = 1, 2, \ldots, m, \]  
   \[ \text{(A-10)} \]
   such that \( c_i \) is the row sum, minus the respective column sum in the concordance matrix.

   The discordance index for the \( i \)th enterprise’s EC is defined as
   \[ d_i = \sum_{i = 1}^{m} d_{ii'} - \sum_{i 
eq i'} d_{ii'}; \quad i = 1, 2, \ldots, m, \]  
   \[ \text{(A-11)} \]

where \( d_i \) is the row sum, minus the respective column sum in the discordance matrix.

Therefore, a greater value for \( c_i \) and a smaller value for \( d_i \) implies better EC.

5. **Grey relational analysis**: The ideal referenced EC is defined as \( X_0 = (x_{01}, x_{02}, \ldots, x_{0n}) \), where
   \[ x_{0j} = \max_{1 \leq i \leq m} x_{ij} (j = 1, 2, \ldots, n). \]

When we calculate the absolute value of the difference between the ideal referenced EC \( X_0 \) and the original assessment matrix \( Z \), we get the following matrix:
   \[ \Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{m1} & \Delta_{m2} & \cdots & \Delta_{mn} \end{bmatrix}, \]  
   \[ \text{(A-12)} \]

where \( \Delta_{ij} = |x_{ij} - x_{0j}| \). In this matrix, \( \Delta_{\text{max}} \) is the maximum element and \( \Delta_{\text{min}} \) is the minimum element, such that
   \[ \max_{1 \leq i < j \leq n} \Delta_{ij} = \Delta_{\text{max}} \text{ and} \]  
   \[ \min_{1 \leq i < j \leq n} \Delta_{ij} = \Delta_{\text{min}}. \]  
   \[ \text{(A-13)} \]
   \[ \text{(A-14)} \]

The relation parameter matrix is defined as
   \[ \xi = \begin{bmatrix} \xi_{11} & \xi_{12} & \cdots & \xi_{1n} \\ \xi_{21} & \xi_{22} & \cdots & \xi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{m1} & \xi_{m2} & \cdots & \xi_{mn} \end{bmatrix}, \]  
   \[ \text{(A-15)} \]

where
   \[ \xi_{ij} = \frac{\Delta_{\text{min}} + \rho \Delta_{\text{max}}}{\Delta_{ij} + \rho \Delta_{\text{max}}} \]  
   \[ \text{(A-16)} \]

Therefore, \( \rho \) is set between 0 and 1, and the smaller the value of \( \Delta_{ij} \), the bigger is the value of \( \xi_{ij} \). In turn, the value of \( \Delta_{ij} \) reflects the degree of relation between the \( i \)th enterprise’s EC and the ideal referenced EC.

We define the degree of relation between \( X_i \) and \( X_0 \) as
   \[ \Gamma_i = \sum_{j=1}^{n} \omega_j \xi_{ij}; \quad i = 1, 2, \ldots, m. \]  
   \[ \text{(A-17)} \]

6. **Nonradial DEA without inputs**: Assume \( m \) decision-making units (DMUs) (i.e., the evaluated enterprises’ EC), and each unit has \( n \) outputs (indicators). Then \( y_{ij} \) is the \( r \)th output value for the \( j \)th DMU \( (j = 1, 2, \ldots, m, r = 1, 2, \ldots, n) \), and \( y_{0i} \) is the \( r \)th output value for DMU \( 0 \) under evaluation. In turn, \( 1/\delta_{0i} \) represents the relative efficiency score of DMU \( 0 \) under evaluation with respect to the \( r \)th output, and \( 1/\delta_{ij} \) represents the relative efficiency score of DMU \( i \) under evaluation. The nonradial DEA model without
7. FA-PCA: The specific operation steps of FA-PCA are:

(1) Apply factor analysis to the original indicator data and extract several factors, each of which consists of one or several indicators.

(2) Carry out principal components analysis on the data of the indicators contained in each factor and calculate each sample’s first principal component value of every factor.

(3) Synthesize every sample’s first principal component values, employing some weighting method to get the general assessment score of each sample.

We use SPSS software to calculate the values.

Appendix B

B.1. Questionnaire

Part 1: Basic information about respondent
Respondent’s name: ___ Position: ___
Telephone: ___ Fax: ___
Address: ___
Post code: ___ Email: ___

Part 2: Basic information about your enterprise
Enterprise’s name: The number of employees: ___ The number of IT staff: ___ The number of years since the inception of e-commerce: ___

Part 3: Evaluation on the importance of indicators and the actual effect of e-commerce in your enterprise

Please choose a number to indicate the importance of each e-commerce system performance indicator (1 = Very unimportant; 5 = Very important).

Please choose a number to indicate the extent to which you agree or disagree that your company’s e-commerce has provided the following benefits (1 = Strongly disagree; 5 = Strongly agree).

<table>
<thead>
<tr>
<th>E-commerce system performance indicators</th>
<th>Importance ratings</th>
<th>Evaluation ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enabled us to develop new markets and enlarge our sales area</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Improved customer service and increased customer satisfaction</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Provided more services to customers</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Reduced customer service expense</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Reduced marketing cost</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Reduced sales expense</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Increased inventory turnover</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Facilitated collaboration with our partners</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Reduced procurement cost</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Reduced procurement expense</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
<tr>
<td>Increased sales revenue</td>
<td>5 4 3 2 1</td>
<td>5 4 3 2 1</td>
</tr>
</tbody>
</table>

Part 4: Overall evaluation

The overall effect of e-commerce system: ___ (also choose from 5 to 1)

References


[20] K. Zhu, K.L. Kraemer, E-commerce metrics for net-enhance organizations: asses-

Dr. Jinghua Huang is Professor and Deputy Chair of Department of Management Science and Engineering, School of Economics and Management, Tsinghua University. She received her Ph.D. in Information Systems from Tsinghua University. She conducted research at the Sloan School of MIT and at the University of Waterloo as a visiting scholar. In 1997, she spent 1 year at the University of Illinois as a Fulbright Scholar. Her main research interests include electronic business, decision support systems, and IS analysis. Her papers have appeared in the
Communications of AIS, Computers & Industrial Engineering, the IEEE International Conference on Systems, Man and Cybernetics, the IEEE International Conference on WebIntelligence, World Multiconference on Systemics, and Cybernetics and Informatics. She is the author of nine books in Chinese.

**Mr. Ximin Jiang** is an Information Systems projects manager at the HongKong and Shanghai Banking Corporation (HSBC). He received his Bachelor’s and Master’s Degree from Tsinghua University. His main research interests and experiences include electronic business and IS analysis and development.

**Ms. Qian Tang** is a Ph.D. candidate in the department of IROM, McCombs School of Business, the University of Texas at Austin. She received her Bachelor’s and Master’s Degree from Tsinghua University. Her main research interests include electronic business and IS.