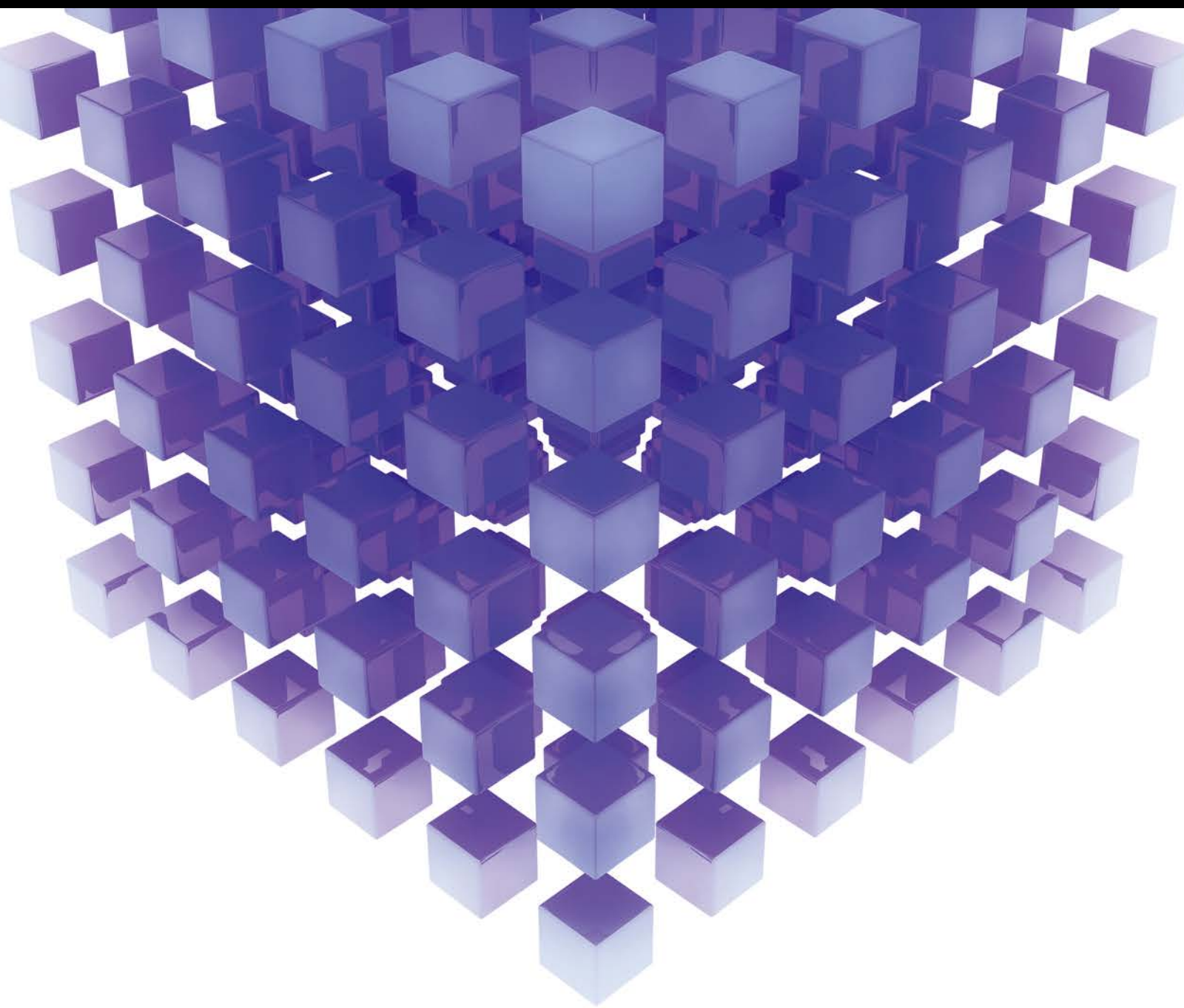
Mathematical Problems in Engineering

# Decision-Making Modeling in Service Systems

Guest Editors: Zhimin Huang, Feng Yang, Desheng D. Wu, Victor Shi, and Alireza Amirteimoori



# Decision-Making Modeling in Service Systems

Mathematical Problems in Engineering

# Decision-Making Modeling in Service Systems

Guest Editors: Zhimin Huang, Feng Yang, Desheng D. Wu, Victor Shi, and Alireza Amirteimoori

Copyright © 2017 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "Mathematical Problems in Engineering." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Editorial Board


Mohamed Abd El Aziz, Egypt
Farid Abed-Meraim, France
José Ángel Acosta, Spain
Paolo Addesso, Italy
Claudia Adduce, Italy
Ramesh Agarwal, USA
Juan C. Agüero, Australia
R Aguilar-López, Mexico
Tarek Ahmed-Ali, France
Hamid Akbarzadeh, Canada
Muhammad N. Akram, Norway
Guido Ala, Italy
Mohammad-Reza Alam, USA
Salvatore Alfonzetti, Italy
Francisco Alhama, Spain
Mohammad D. Aliyu, Canada
Juan A. Almendral, Spain
Lionel Amodeo, France
Sebastian Anita, Romania
Renata Archetti, Italy
Felice Arena, Italy
Sabri Arik, Turkey
Alessandro Arsie, USA
Edoardo Artioli, Italy
Fumihiro Ashida, Japan
Hassan Askari, Canada
Mohsen Asle Zaeem, USA
Romain Aubry, USA
Matteo Aureli, USA
Francesco Aymerich, Italy
Seungik Baek, USA
Khaled Bahlali, France
Laurent Bako, France
Stefan Balint, Romania
Alfonso Banos, Spain
Roberto Baratti, Italy
Azeddine Beghdadi, France
Denis Benasciutti, Italy
Ivano Benedetti, Italy
Elena Benvenuti, Italy
Jamal Berakdar, Germany
Michele Betti, Italy
Jean-Charles Beugnot, France
Simone Bianco, Italy
Gennaro N. Bifulco, Italy
David Bigaud, France
Antonio Bilotta, Italy
Jonathan N. Blakely, USA
Paul Bogdan, USA
Alberto Borboni, Italy
Paolo Boscariol, Italy
Daniela Boso, Italy
G. Botella-Juan, Spain
Abdel-Ouahab Boudraa, France
Fabio Bovenga, Italy
Francesco Braghin, Italy
Michael J. Brennan, UK
Maurizio Brocchini, Italy
Julien Bruchon, France
Michele Brun, Italy
Javier Buldu', Spain
Tito Busani, USA
R. Caballero-Águila, Spain
Pierfrancesco Cacciola, UK
Salvatore Caddemi, Italy
Jose E. Capilla, Spain
F. Javier Cara, Spain
Ana Carpio, Spain
Carmen Castillo, Spain
Inmaculada T. Castro, Spain
Gabriele Cazzulani, Italy
Luis Cea, Spain
Miguel E. Cerrolaza, Spain
M. Chadli, France
Gregory Chagnon, France
Ching-Ter Chang, Taiwan
Michael J. Chappell, UK
Kacem Chehdi, France
Peter N. Cheimets, USA
Chunlin Chen, China
Xinkai Chen, Japan
Francisco Chicano, Spain
Hung-Yuan Chung, Taiwan
Joaquim Ciurana, Spain
John D. Clayton, USA
Giuseppina Colicchio, Italy
Mario Cools, Belgium
Sara Coppola, Italy
Jean-Pierre Corriou, France
J.-C. Cortés, Spain
Carlo Cosentino, Italy
Paolo Crippa, Italy
Andrea Crivellini, Italy
Erik Cuevas, Mexico
Peter Dabnichki, Australia
Luca D'Acierno, Italy
Weizhong Dai, USA
Andrea Dall'Asta, Italy
Purushothaman Damodaran, USA
Farhang Daneshmand, Canada
Fabio De Angelis, Italy
Pietro De Lellis, Italy
Stefano de Miranda, Italy
Filippo de Monte, Italy
M. do Rosário de Pinho, Portugal
Xavier Delorme, France
Frédéric Demoly, France
Luca Deseri, USA
Angelo Di Egidio, Italy
Ramón I. Diego, Spain
Yannis Dimakopoulos, Greece
Zhengtao Ding, UK
M. Djemai, France
Manuel Doblaré, Spain
Alexandre B. Dolgui, France
Florent Duchaine, France
George S. Dulikravich, USA
Bogdan Dumitrescu, Romania
Horst Ecker, Austria
Ahmed El Hajjaji, France
Fouad Erchiqui, Canada
Anders Eriksson, Sweden
R. Emre Erkmen, Australia
Andrea L. Facci, Italy
Giovanni Falsone, Italy
Hua Fan, China
Yann Favennec, France
Fiorenzo A. Fazzolari, UK
Giuseppe Fedele, Italy
Roberto Fedele, Italy
Jose R. Fernandez, Spain
Jesus M. Fernandez Oro, Spain
Eric Feulvarch, France
Barak Fishbain, Israel
S. Douwe Flapper, Netherlands

Thierry Floquet, France
Eric Florentin, France
Jose M. Framinan, Spain
Francesco Franco, Italy
Elisa Francomano, Italy
Leonardo Freitas, UK
Tomonari Furukawa, USA
Mohamed Gadala, Canada
Matteo Gaeta, Italy
Mauro Gaggero, Italy
Zoran Gajic, Iraq
Erez Gal, Israel
Ugo Galvanetto, Italy
Akemi Gálvez, Spain
Rita Gamberini, Italy
Maria L. Gandarias, Spain
Arman Ganji, Canada
Xin-Lin Gao, USA
Zhong-Ke Gao, China
Giovanni Garcea, Italy
Fernando García, Spain
Jose M. Garcia-Aznar, Spain
Alessandro Gasparetto, Italy
Vincenzo Gattulli, Italy
Oleg V. Gendelman, Israel
Mergen H. Ghayesh, Australia
Agathoklis Giaralis, UK
Anna M. Gil-Lafuente, Spain
Ivan Giorgio, Italy
Alessio Gizzi, Italy
Hector Gómez, Spain
David González, Spain
Francisco Gordillo, Spain
Rama S. R. Gorla, USA
Oded Gottlieb, Israel
Nicolas Gourdain, France
Kannan Govindan, Denmark
Antoine Grall, France
Fabrizio Greco, Italy
Jason Gu, Canada
Federico Guarracino, Italy
José L. Guzmán, Spain
Quang Phuc Ha, Australia
Zhen-Lai Han, China
Thomas Hanne, Switzerland
Xiao-Qiao He, China
Sebastian Heidenreich, Germany
Luca Heltai, Italy
A. G. Hernández-Diaz, Spain
Rafael M. Herrera, Spain
M.I. Herreros, Spain
Eckhard Hitzer, Japan
Jaromir Horacek, Czech Republic
Muneo Hori, Japan
András Horváth, Italy
Gordon Huang, Canada
Nicolas Hudon, Canada
Sajid Hussain, Canada
Asier Ibeas, Spain
Orest V. Iftime, Netherlands
Giacomo Innocenti, Italy
Emilio Insfran, Spain
Nazrul Islam, USA
Benoit Iung, France
Benjamin Ivorra, Spain
Payman Jalali, Finland
Reza Jazar, Australia
Khalide Jbilou, France
Linni Jian, China
Bin Jiang, China
Zhongping Jiang, USA
Ningde Jin, China
Grand R. Joldes, Australia
Dylan F. Jones, UK
Tamas Kalmar-Nagy, Hungary
Tomasz Kapitaniak, Poland
Haranath Kar, India
Konstantinos Karamanos, Belgium
Jean-Pierre Kenne, Canada
C. M. Khalique, South Africa
Do Wan Kim, Republic of Korea
Nam-Il Kim, Republic of Korea
Oleg Kirillov, Germany
Manfred Krafczyk, Germany
Frederic Kratz, France
Petr Krysl, USA
Jurgen Kurths, Germany
Kyandoghere Kyamakya, Austria
Davide La Torre, Italy
Risto Lahdelma, Finland
Hak-Keung Lam, UK
Jimmy Lauber, France
Antonino Laudani, Italy
Aime' Lay-Ekuakille, Italy
Nicolas J. Leconte, France
Marek Lefik, Poland
Yaguo Lei, China
Stefano Lenci, Italy
Roman Lewandowski, Poland
Panos Liatsis, UAE
Anatoly Lisnianski, Israel
Peide Liu, China
Peter Liu, Taiwan
Wanquan Liu, Australia
Alessandro Lo Schiavo, Italy
Jean J. Loiseau, France
Paolo Lonetti, Italy
Sandro Longo, Italy
Sebastian López, Spain
Luis M. López-Ochoa, Spain
Vassilios C. Loukopoulos, Greece
Valentin Lychagin, Norway
Antonio Madeo, Italy
José María Maestre, Spain
Fazal M. Mahomed, South Africa
Noureddine Manamanni, France
Didier Maquin, France
Paolo Maria Mariano, Italy
Damijan Markovic, France
F. Marotti de Sciarra, Italy
Benoit Marx, France
Franck Massa, France
Paolo Massioni, France
Ge'rard A. Maugin, France
Alessandro Mauro, Italy
Michael Mazilu, UK
Driss Mehdi, France
Roderick Melnik, Canada
Pasquale Memmolo, Italy
Xiangyu Meng, Canada
Jose Merodio, Spain
Alessio Merola, Italy
Luciano Mescia, Italy
Laurent Mevel, France
Y. Vladimirovich Mikhlin, Ukraine
Aki Mikkola, Finland
Hiroyuki Mino, Japan
Pablo Mira, Spain
Vito Mocella, Italy
Roberto Montanini, Italy
Gisele Mophou, France
Marco Morandini, Italy
Simone Morganti, Italy
Aziz Moukrim, France

Emiliano Mucchi, Italy
Domenico Mundo, Italy
Jose J. Muñoz, Spain
Giuseppe Muscolino, Italy
Marco Mussetta, Italy
Hakim Naceur, France
Hassane Naji, France
Keivan Navaie, UK
Dong Ngoduy, UK
Tatsushi Nishi, Japan
Xesús Nogueira, Spain
Ben T. Nohara, Japan
Mohammed Nouari, France
Mustapha Nourelfath, Canada
Sotiris K. Ntouyas, Greece
Roger Ohayon, France
Mitsuhiro Okayasu, Japan
Calogero Orlando, Italy
Javier Ortega-Garcia, Spain
Alejandro Ortega-Moñux, Spain
Naohisa Otsuka, Japan
Erika Ottaviano, Italy
Arturo Pagano, Italy
Alkis S. Paipetis, Greece
Alessandro Palmeri, UK
Anna Pandolfi, Italy
Elena Panteley, France
Achille Paolone, Italy
Xosé M. Pardo, Spain
Manuel Pastor, Spain
Pubudu N. Pathirana, Australia
Francesco Pellicano, Italy
Marcello Pellicciari, Italy
Haipeng Peng, China
Mingshu Peng, China
Zhike Peng, China
Marzio Pennisi, Italy
Matjaz Perc, Slovenia
Francesco Pesavento, Italy
Dario Piga, Italy
Antonina Pirrotta, Italy
Marco Pizzarelli, Italy
Vicent Pla, Spain
Javier Plaza, Spain
Sébastien Poncet, Canada
Jean-Christophe Ponsart, France
Mauro Pontani, Italy
Stanislav Potapenko, Canada
Christopher Pretty, New Zealand
Carsten Proppe, Germany
Luca Pugi, Italy
Giuseppe Quaranta, Italy
Dane Quinn, USA
Vitomir Racic, Italy
Jose Ragot, France
K. Ramamani Rajagopal, USA
Gianluca Ranzi, Australia
Alain Rassineux, France
S.S. Ravindran, USA
Alessandro Reali, Italy
Oscar Reinoso, Spain
Nidhal Rezg, France
Ricardo Riaza, Spain
Gerasimos Rigatos, Greece
Francesco Ripamonti, Italy
Eugenio Roanes-Lozano, Spain
Bruno G. M. Robert, France
José Rodellar, Spain
Rosana Rodriguez-Lopez, Spain
Ignacio Rojas, Spain
Alessandra Romolo, Italy
Carla Roque, Portugal
Debasish Roy, India
Gianluigi Rozza, Italy
Rubén Ruiz García, Spain
Antonio Ruiz-Cortes, Spain
Ivan D. Rukhlenko, Australia
Mazen Saad, France
Kishin Sadarangani, Spain
Mehrdad Saif, Canada
Miguel A. Salido, Spain
Roque J. Saltarén, Spain
Francisco J. Salvador, Spain
Alessandro Salvini, Italy
Maura Sandri, Italy
Miguel A. F. Sanjuan, Spain
Juan F. San-Juan, Spain
Roberta Santoro, Italy
I. Ferreira Santos, Denmark
José A. Sanz-Herrera, Spain
Nickolas S. Sapidis, Greece
E. J. Sapountzakis, Greece
Andrey V. Savkin, Australia
Thomas Schuster, Germany
Mohammed Seaid, UK
Lotfi Senhadji, France
Joan Serra-Sagrista, Spain
Gerardo Severino, Italy
Ruben Sevilla, UK
Leonid Shaikhet, Ukraine
Hassan M. Shanechi, USA
Bo Shen, Germany
Suzanne M. Shontz, USA
Babak Shotorban, USA
Zhan Shu, UK
Dan Simon, Greece
Luciano Simoni, Italy
C. H. Skiadas, Greece
Alba Sofi, Italy
Francesco Soldovieri, Italy
Raffaele Solimene, Italy
Jussi Sopanen, Finland
Ruben Specogna, Italy
Sri Sridharan, USA
Ivanka Stamova, USA
Salvatore Strano, Italy
Yakov Strelniker, Israel
Sergey A. Suslov, Australia
Thomas Svensson, Sweden
Andrzej Swierniak, Poland
Yang Tang, Germany
Alessandro Tasora, Italy
Sergio Teggi, Italy
Alexander Timokha, Norway
Gisella Tomasini, Italy
Francesco Tornabene, Italy
Antonio Tornambe, Italy
Irina N. Trendafilova, UK
George Tsiatas, Greece
Antonios Tsourdos, UK
Vladimir Turetsky, Israel
Mustafa Tutar, Spain
Ilhan Tuzcu, USA
E. Tzirtzilakis, Greece
Filippo Ubertini, Italy
Francesco Ubertini, Italy
Hassan Ugail, UK
Giuseppe Vairo, Italy
Kuppalapalle Vajravelu, USA
Robertt A. Valente, Portugal
Eusebio Valero, Spain
Pandian Vasant, Malaysia
Marcello Vasta, Italy
M. E. Vázquez-Méndez, Spain

Josep Vehi, Spain
K. C. Veluvolu, Republic of Korea
Fons J. Verbeek, Netherlands
Franck J. Vernerey, USA
Georgios Veronis, USA
Anna Vila, Spain
R.-J. Villanueva-Micó, Spain
U. E. Vincent, UK
Mirko Viroli, Italy
Michael Vynnycky, Sweden
Shuming Wang, China
Yan-Wu Wang, China
Yongqi Wang, Germany
Roman Wendner, Austria
Desheng D. Wu, Sweden
Yuqiang Wu, China
Guangming Xie, China
Xuejun Xie, China
Gen Q. Xu, China
Hang Xu, China
Joseph J. Yame, France
Xinggang Yan, UK
Luis J. Yebra, Spain
Peng-Yeng Yin, Taiwan
Qin Yuming, China
Vittorio Zampoli, Italy
Ibrahim Zeid, USA
Huaguang Zhang, China
Qingling Zhang, China
Jian G. Zhou, UK
Quanxin Zhu, China
Mustapha Zidi, France

# Contents

**Decision-Making Modeling in Service Systems**
Zhimin Huang, Feng Yang, Desheng D. Wu, Victor Shi, and Alireza Amirteimoori
Volume 2017, Article ID 6873951, 3 pages

**Research on Methods for Discovering and Selecting Cloud Infrastructure Services Based on Feature Modeling**
Huamin Zhu, Lifa Wu, Kangyu Huang, and Zhenji Zhou
Volume 2016, Article ID 8194832, 19 pages

**Incentive Mechanism Design Aiming at Deflated Performance Manipulation in Retail Firms: Based on the Ratchet Effect and the Reputation Effect**
Biao Luo, Chengyuan Wang, and ChunYi Li
Volume 2016, Article ID 5412028, 9 pages

**Decision-Making Approach to Selecting Optimal Platform of Service Variants**
Vladimir Modrak, Slavomir Bednar, and Pavol Semanco
Volume 2016, Article ID 9840679, 8 pages

**Motivation System of Crowdsourcing Community from a Supply Chain Perspective**
Jiangang Pang and Zhiying Liu
Volume 2016, Article ID 8306950, 8 pages

**The Effects of the Emission Cost on Route Choices of International Container Ships**
Hyangsook Lee, Kang-Dae Lee, and Sangho Choo
Volume 2016, Article ID 6850308, 13 pages

**Service and Price Decisions of a Supply Chain with Optional After-Sale Service**
Xiaochen Sun, Qingshuai Zhang, and Yancong Zhou
Volume 2016, Article ID 1505817, 11 pages

**Sourcing for Quality: Cooperating with a Single Supplier or Developing Two Competing Suppliers?**
Jingxian Chen
Volume 2016, Article ID 3040343, 13 pages

**Research on Medicine Distribution Route Optimization for Community Health Service Institutions**
Zhao Li, Zhiqiang Ma, Wei Shi, and Xianggen Qian
Volume 2016, Article ID 6153898, 7 pages

**Multiple Attribute Decision Making Based on Cross-Evaluation with Uncertain Decision Parameters**
Tao Ding, Liang Liang, Min Yang, and Huaqing Wu
Volume 2016, Article ID 4313247, 10 pages

**Optimal Replenishment Policy for Weibull-Distributed Deteriorating Items with Trapezoidal Demand Rate and Partial Backlogging**
Lianxia Zhao
Volume 2016, Article ID 1490712, 10 pages

**An Analysis of Bank Service Satisfaction Based on Quantile Regression and Grey Relational Analysis**
Wen-Tsao Pan and Yungho Leu
Volume 2016, Article ID 1475148, 9 pages

**A Robust Service Selection Method Based on Uncertain QoS**
Yanping Chen, Lu Jiang, Jianke Zhang, and Xiaoxiao Dong
Volume 2016, Article ID 9480769, 10 pages

**Alternative Trade-Offs in Data Envelopment Analysis: An Application to Hydropower Plants**
Mohsen Mirzaei, Sohrab Kordrostami, Alireza Amirteimoori, and Mehrdad G. Chegini
Volume 2016, Article ID 1531282, 8 pages

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2017, Article ID 6873951, 3 pages
http://dx.doi.org/10.1155/2017/6873951

# *Editorial*
# Decision-Making Modeling in Service Systems

**Zhimin Huang,[1] Feng Yang,[2] Desheng D. Wu,[3] Victor Shi,[4] and Alireza Amirteimoori[5]**

[1]*Robert B. Willumstad School of Business, Adelphi University, Garden City, Long Island, NY 11530, USA*
[2]*School of Management, University of Science and Technology of China, No. 96 Jinzhai Road, Hefei 230026, China*
[3]*School of Economics and Management, University of Chinese Academy of Science, No. 19 Yuquan Road, Beijing 100049, China*
[4]*Lazaridis School of Business and Economics, Wilfrid Laurier University, Waterloo, ON, Canada N2L 6C1*
[5]*Department of Applied Mathematics, Islamic Azad University, Rasht, Iran*

Correspondence should be addressed to Zhimin Huang; huang@adelphi.edu

Received 17 October 2016; Accepted 17 October 2016; Published 29 January 2017

Copyright © 2017 Zhimin Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Global economic development has been exhibiting new characteristics. First, rapid developments in productivity make it possible for individuals to pursue more, as well as higher levels of, desires for novel services, diverse products, convenient communication, and innovative forms of entertainment. New service industries have emerged to fulfill such desires. Second, the innovation and development of new technologies, especially the maturity of the Internet and the appearance of IOT (Internet of Things), accelerate the growth of new service industries. Third, the pressure from environmental deterioration and depletion of resources force many countries to adjust industrial structures from manufacturing to service industries. As a result, service is playing a more significant role in modern society.

Nowadays, with the increasingly fierce market competition, compared with the same type of enterprises, it is difficult for an enterprise to get cheaper raw materials and labors, and it is also difficult to sell products at higher prices. In addition, the increasing scarcity of resource and environmental pressures force enterprises to lower resource consumption and environmental costs to provide products or services, which force enterprises to enhance the level of performance management. However, studies on operations, strategies, marketing, and financing of service systems remain scarce in the literature. Such a reality pushes us to organize a special issue aiming at discussing the decision-making methods in service systems which should be appropriate and applicable to improve the performance of service industries. It should be noticed that the performance indications of service systems are various in different industries, for example, friendliness of user interface in Internet service, profitability in banking service, timeliness in express logistics, and quality in health care.

The special issue has received 42 high-quality submissions. After a rigorous peer-review process, with an acceptance rate of about 30%, thirteen papers have been accepted for publication. These accepted papers provide a variety of applications in service systems, and the authors have contributed towards combining the innovative methodology and mathematical analysis in solving real-world problems. The authors of these papers are from China mainland, Iran, Korea, Slovakia, and Taiwan. The service systems considered in these papers include health care, information technique and Internet, transportation, supply chain, retailing, power, and banking. This selection of papers reflects the diversity of research studies applying different decision-making models to improve the operations of service systems for different areas.

In this special issue, lots of decision-making methods are developed to improve the performance of different service systems. We classify these methods into four groups, including optimization, multiattribute or multiobjective decision-making, stochastic or fuzzy decision-making, and game theory based decision-making.

Optimization is one of the most important decision-making problems. It discusses the characteristics of the best choice of the decision problem and constructs the calculation method of the optimal solution. With the rapid development of computer and the improvement of the optimization method, lots of optimization problems in real world are

solved. The optimization problems are widely found in economic planning, engineering design, production management, transportation, national defense, and other important areas. A representative study on the system optimization is to use mathematical programming to seek the best solution for a given objective function.

On the system optimization, in this special issue, H. Zhu et al. develop a method for discovering and selecting cloud infrastructure services. With the rapid development of cloud computing, more and more cloud infrastructure service providers are providing large numbers of service instances. For cloud infrastructure services, the lack of a description standard and selection methods has hindered the services' discovering and choosing for users. H. Zhu et al. use a feature model to describe the configurable properties of a cloud infrastructure service and present methods of creating the service and demand models. Based on the description of the cloud infrastructure service, a systematic discovery and selection method for cloud infrastructure services is proposed. V. Modrak et al. have contributed a very interesting problem on the decision-making approach to select optimal platform of service variants. With the trend of mass customization, services are more abstract than products, and concepts for mass customization in manufacturing cannot be used directly in service without a methodical change. V. Modrak et al. focus on the development of a methodological framework to support decisions in a selection of optimal platform of service variants when compatibility problems between service options occurred. The approach is based on mutual relations between waste and constrained design space entropy. In addition, Z. Li et al. have studied the medicine distribution route optimization problem for community health service institutions. Such a study is very important and fundamental. The optimization objective is defined as the overall cost which includes the refrigeration storage cost, vehicle fixed cost, and transportation cost. Considering the characteristics of medicine distribution for community health service institutions, Z. Li et al. set up a distribution route optimizing model to minimize the overall cost. Finally, H. Lee et al. examine the effects of the emission cost on route choices of international container ships. The paper formulates a freight network model that captures decisions and interactions of ocean carriers and port terminal operators in the maritime freight transport system. Then, the emission cost is calculated based on an activity-based approach as a component of the ocean transportation cost function.

Multiattribute or multiobjective decision-making is a very important type of decision method for service system. The essential characteristic of service operations makes the importance of multiattribute or multiobjective decision-making more prominent. Compared with manufacturing operations, service operation has more objectives. Except the operational goals in manufacturing, such as reducing costs, saving manpower, reducing the waste, controlling the pollution, increasing profits, and increasing the flexibility, the targets of service operations also need to meet other goals, like improving customer satisfaction (or utility), speeding up the service response time and the intelligent level of service, accelerating product innovation speed, and others. Different kinds of targets are difficult to achieve at the same time. Thus the service operations can be characterized as multiattribute or multiobjective decision-making problems.

On the multiattribute or multiobjective decision-making problems, M. Mirzaei et al. have contributed towards solving a key theoretic problem in the representative multiattribute decision-making technique, namely, Data Envelopment Analysis (DEA). In multidimensional input/output space, the behavior of the firms can be analyzed by using efficient frontier or supporting surfaces of production technology. The piecewise linear frontier of DEA technology is not differentiable at the extreme points. M. Mirzaei et al. show that binding efficient supporting surfaces of an efficient point may be used to define different marginal rates of substitutions, and they get different marginal rates to each frontier point. T. Ding et al. propose multiple attribute decision-making based on cross-evaluation with uncertain decision parameters. The proposed approach assumes that all attribute weights are uncertain. It can overcome the drawback in prior research that the alternatives' ranking may be determined by a single attribute with an overestimated weight. In addition, the proposed method can also balance the mean and deviation of each alternative's cross-evaluation score to guarantee the stability of evaluation.

Stochastic or fuzzy decision-making usually pays attention to the decision problems where the attribution variables are not deterministic. Stochastic variable and fuzzy variable are the main types of indeterministic variables. Service system is in the state of evolution, and the property and variables are often in an unstable situation. If the distribution of some factors has certain regularity, stochastic distribution can be used to describe the state of the factors; if the distribution of the factors has no rules, then we can only use fuzzy mathematics to describe the state of the factors.

On the stochastic or fuzzy decision-making, W.-T. Pan and Y. Leu propose an analysis of bank service satisfaction based on quantile regression and grey relational analysis. With the grey relational analysis, they compare the effects of different variables on service satisfaction. Zhao develops an optimal replenishment policy for Weibull-distributed deteriorating items with trapezoidal demand rate and partial backlogging. To minimize the total cost per unit time, L. Zhao proposes an efficient solution procedure to determine the optimal replenishment and the optimal order quantity. In addition, Y. Chen et al. propose a robust service selection method based on uncertain QoS. QoS has become an important factor used to select the most appropriate service for users. The most prominent QoS-based service selection models only take the certain attributes into account. However, in the real world, there are a large number of uncertain factors. The method of Y. Chen et al. can help obtain a robust selection under uncertain situations.

Game theory mainly studies the interaction between the incentive structures. It is a mathematical method to study the phenomenon of competition among multiple decision makers. Game theory considers the individual's forecast behavior and the actual behavior in the game and studies their optimization strategies. Game-based decision-making has been one of the most important areas in decision science.

In the area of game-based decision-making, J. Chen considers the strategic interactions under two sourcing strategies (single and dual sourcing) to improve the sourcing quality by considering a manufacturer that wants to outsource the manufacturing of two substitute products to external suppliers, which is a very innovative study. The authors find out that an appropriate proportion of quality investment sharing by the manufacturer can enable a cooperating strategy with a single supplier to be the dominant strategy. X. Sun et al. focus on the optimal service and price decisions of a supply chain with optional after-sale service in the supply chain composed of a manufacturer and a retailer. They derive the equilibrium decisions between the manufacturer and the retailer under two different cases. Crowdsourcing community is a very novel topic, and the motivation system is of great importance to develop the crowdsourcing community. J. Pang and Z. Liu have contributed a precursory study on such a topic by using the principal-agent theory. They prove that enterprises can generate innovation plans of high quality and expected utility using the crowdsourcing community. Outsourcers can encourage high-quality people to join. Similarly, B. Luo et al. also focus on the design of incentive mechanism, while they aim at deflated performance manipulation in retail firms considering the ratchet effect and the reputation effect. They model a dynamic agency setting in which both the ratchet effect and the reputation effect are related to the store manager's compensation scheme. The dynamic decision makes the study of B. Luo et al. feasible and valuable.

We hope these papers will be helpful for readers to understand the wide use of decision-making methods in service systems and inspire their interest to further the exploratory research in this area, as well as its related topics.

## Acknowledgments

We sincerely express our gratitude to the authors for their contributions and the reviewers for their time and effort in assessing the manuscripts. The support from NSFC (71322101 and 71631006) is also acknowledged.

*Zhimin Huang*
*Feng Yang*
*Desheng D. Wu*
*Victor Shi*
*Alireza Amirteimoori*

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 8194832, 19 pages
http://dx.doi.org/10.1155/2016/8194832

*Research Article*

# Research on Methods for Discovering and Selecting Cloud Infrastructure Services Based on Feature Modeling

**Huamin Zhu, Lifa Wu, Kangyu Huang, and Zhenji Zhou**

*Institute of Command Information System, PLA University of Science and Technology, Nanjing, China*

Correspondence should be addressed to Lifa Wu; wulifa@vip.163.com

Received 15 December 2015; Revised 24 July 2016; Accepted 24 August 2016

Academic Editor: Zhimin Huang

Copyright © 2016 Huamin Zhu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays more and more cloud infrastructure service providers are providing large numbers of service instances which are a combination of diversified resources, such as computing, storage, and network. However, for cloud infrastructure services, the lack of a description standard and the inadequate research of systematic discovery and selection methods have exposed difficulties in discovering and choosing services for users. First, considering the highly configurable properties of a cloud infrastructure service, the feature model method is used to describe such a service. Second, based on the description of the cloud infrastructure service, a systematic discovery and selection method for cloud infrastructure services are proposed. The automatic analysis techniques of the feature model are introduced to verify the model's validity and to perform the matching of the service and demand models. Finally, we determine the critical decision metrics and their corresponding measurement methods for cloud infrastructure services, where the subjective and objective weighting results are combined to determine the weights of the decision metrics. The best matching instances from various providers are then ranked by their comprehensive evaluations. Experimental results show that the proposed methods can effectively improve the accuracy and efficiency of cloud infrastructure service discovery and selection.

## 1. Introduction

The Infrastructure as a Service (IaaS) is an important service model in cloud computing, and it can provide users with fundamental computing resources, for example, processing, storage, and network, where users are able to deploy the desired operating systems and applications without paying attention to the management and control of the underlying cloud infrastructure [1]. IaaS is provided to users by a service instance that may be a different combination of processing, storage, and network resources that satisfies the diversified needs of the users. Cloud computing with potentially huge business opportunities has attracted a growing number of IT service providers to build infrastructures that can provide dozens, hundreds, or even tens of thousands of IaaS instances. Currently, there are over 100 public IaaS providers, and just for Amazon's Elastic Compute Cloud (EC2), the number of service instances is over 16000 [2]. How to select the best IaaS instance from these large numbers of alternatives has become a serious challenge to users.

Different IaaSs show differences in resource configuration and service quality, and thus the discovery and selection of the IaaS instances not only need to perform fine-grained matching on the resource configuration according to user needs, but also need to find the optimal instance based on comparison of the quality of service (QoS) of the matching instances. It is obviously a difficult and error-prone task to search the matching instances in the huge configuration space of an IaaS, which necessitate automatic searching methods to support it. Moreover, a set of scientific QoS metrics and an effective multicriterion decision-making method for ranking the matching instances are also needed. At present, although many scholars have done extensive research on the discovery and selection of IaaSs [3–10], the following challenges remain: (1) There exists a lack of customized description methods for IaaSs, and the prevalent natural language description cannot be used for automatic processing by computers. (2) The existing methods are unable to describe or formalize the configuration space of an IaaS completely; thus, these methods can only handle a subset of the entire configuration

spaces when discovering and selecting an IaaS, which results in a suboptimal service instance. (3) Due to the existing coarse-grained description of IaaSs, the matching between the user request and the IaaS offerings can only be done in a coarse-grained way, so that some service details, such as cost, cannot be accurately calculated. (4) Also lacking are systematic theory and method for describing, discovering, and selecting IaaSs. (5) There is no scientific and unified measurement metric system for the QoS evaluation of IaaSs. The existing IaaS decision-making methods assign metric weight mainly based on the subjective weighting method, which is likely to lead to an inaccurate evaluation of QoS.

From the above analysis, it is not difficult to see that the formalization and fine-grained description of the complete configuration space of an IaaS is the key to discovering and selecting a truly optimal IaaS instance. In general, the configuration space of an IaaS comprises a large number of service instances, and there are two ways of providing IaaS instances. The first is to provide great many standard instances from which users can choose; for example, in EC2, there are m4.large, c3.large, and r3.xlarge instances of the virtual server, and each instance has some basic options with fixed values as well as some configurable options, for example, m4.large (2.4 GHz Intel Processor, 2 vCPU, 8 GB RAM, EBS-only, 450 Mbps bandwidth) has configurable location, cost model, and operating system (OS) options [11]. The second way of providing an IaaS instance is to specify the value range of the configurable options, and any combination of all the option values within the range is an IaaS instance. There are some configurable service options in an IaaS using either approach, such as virtual machine (VM) type, RAM, OS, cost model, location, network bandwidth, and any combination of these options is an IaaS instance such that all the instances form the complete configuration space. Therefore, an IaaS can be regarded as a variability-intensive system, and it is pivotal to deal with the variability of an IaaS to fully obtain the fine-grained description of its entire configuration space.

Kang et al. [12] proposed a feature-oriented domain analysis method (FODA) in 1990. Here, features and the feature model (FM) are first introduced to capture and organize changing software requirements in special fields [13]. In a FM, features are used to represent the system functional and nonfunctional characteristics, and feature relations are used to depict the organizational structure and internal relations of these characteristics. Such an "entity-relation" description structure in a FM exhibits highly customizable properties, which has led FMs to become a very effective domain commonality/variability description method. With the continuous development of software reuse research, FMs have become the *de facto* standard for software product line commonality/variability modeling, and they can be used as a good description model for variability-intensive systems [14, 15]. Therefore, it is very appropriate to use a FM to describe an IaaS. In this paper, we design a formal description method for IaaSs by means of feature modeling and propose a systematic discovery and selection method for IaaSs based on the description method. The main contributions in this paper are as follows:

(i) A formal IaaS description method based on FMs is designed, which is able to describe the complete configuration space of an IaaS in a standard way. We define the metamodel of an IaaS using the unified modeling language (UML) on the basis of existing methods and then define the IaaS domain model based on the metamodel and the IaaS domain knowledge. Next, we present the methods for creating service and demand models based on the metamodel and the domain model. In addition, an equivalent concept ontology of IaaS FMs is also given to cope with the heterogeneous expressions in different FMs.

(ii) Based on the designed IaaS description model, we propose a method to realize the automatic process of discovery and selection of the best IaaS instance. The automatic analysis techniques of a FM are introduced to extract the service configuration information in the FM, check the validity of an IaaS FM, and perform the searching for the matching IaaS instance based on user needs.

(iii) The important nonfunctional metrics and their measurement methods for IaaSs are studied and determined, and the metrics weighting method based on the combination of the subjective and objective weights is proposed, which makes the weight-assignment process of the decision metrics more reasonable. The multicriteria decision-making method for the matching IaaS instances is also introduced.

The remainder of this paper is organized as follows. Section 2 describes related work. Section 3 presents the formal IaaS description method based on FMs. Section 4 introduces the proposed discovery and selection method of IaaSs. Section 5 describes the automatic analysis techniques and the matching method of IaaS FMs. Section 6 presents the method of ranking IaaSs. Section 7 presents an experiment to evaluate the feasibility and effectiveness of the methods introduced in this paper. Finally, the last section concludes the paper and proposes future work.

## 2. Related Work

In addition to lots of studies, there are some commercial services to support the discovery and selection of IaaSs. Amazon [16] and RackSpace [17] provide exclusive IaaS configuration platforms to support comparisons between their own different service instances. CloudScreener [18], CloudHarmony [2], Cloudorado [19], and PlanForCloud [20] are able to provide a commercial service for the configuration and comparison of IaaS instances, but they can only cover the main IaaS providers and can only support comparisons of a specified QoS metric, such as fee. Research also indicates that the recommended results from the commercial services or platforms are sometimes suboptimal [21]. In academic researches, the related methods for discovering and selecting an IaaS can be divided into the following categories.

*(1) Semantic-Technique-Based Description and Matching Methods for IaaSs.* Dastjerdi et al. [3] used the web service

modeling ontology (WSMO) as the description model of virtual appliances and virtual units in cloud computing and presented the semantic matching method of the supply and demand on IaaSs. Zhang et al. [22] created an IaaS knowledge ontology including functional and nonfunctional concepts, attributes, and relationships among them and proposed a semantic search model for heterogeneous IaaSs. Rodríguez-García et al. [4] created semantic IaaS service vectors by implementing text processing and adding semantics to the IaaS description and then sorted the IaaSs according to the results of semantic matching on IaaS service vectors and user demands. A similar method was given in [23]. Liu et al. [24] created an IaaS function ontology and presented the semantic matching method of IaaSs and the method of selecting IaaSs based on the QoS metrics. Kang and Sim [5] created a cloud service ontology and presented the semantic matching method of IaaSs from three aspects, including the concept, the object attributes of concepts, and the data attributes of concepts. The above studies described IaaSs and performed the matching of supply and demand on IaaSs by means of semantic technologies used commonly in traditional web services. Although these approaches can deal well with the heterogeneous descriptions of IaaSs, they can only describe the static IaaS information and cannot deal with the numerical service parameters and structured service information; thus, the IaaS configuration space cannot be accurately and completely described. Under this condition, the discovery method based on the semantic matching can only be match to providers and cannot be matched to specific IaaS instances, making it difficult to accurately obtain the real optimal IaaS instance. In this paper, we use a FM to describe an IaaS in a structured, complete, and accurate manner and then use the automatic analysis techniques of a FM to search the matching instances of an IaaS so that a real optimal solution can be obtained.

*(2) QoS-Evaluation-Based IaaS Selection Method*. The Cloud Service Measurement Index Consortium (CSMIC) of Carnegie Mellon University [25] presented a service measurement index for evaluating and comparing cloud services in general, including the accountability, agility, cost, performance, assurance, security, privacy, and usability. Garg et al. [6] developed 15 evaluation metrics for IaaSs by selecting and refining the CSMIC index and used these metrics to evaluate the IaaS QoS by means of the AHP. However, some metric measurement methods given by them are not feasible, and the AHP can only deal with a small amount of alternative IaaS instances. Li et al. [7] divided the QoS metrics of IaaSs into three categories, performance, economy, and security, and then designed the corresponding evaluation metrics for computing services, storage services, and comprehensive services. They provided some good metric directories for evaluating the IaaS QoS, but the security measurement was too simple. In [6, 26], the authors determined the weights of QoS metrics only according to the subjective comparison and judgment, and thus it is hard to guarantee the objectivity of evaluations of the IaaS QoS. In this paper, we evaluate the IaaS security in a systematic and objective manner and present a combination weighting method, by which the weight distribution can not only be more objective, but also reflect the subjective preference of decision makers.

*(3) FM-Based IaaS Description and Selection Methods*. Wittern et al. [8] proposed a method of selecting an IaaS based on variability modeling. However, the authors only gave a very simple domain model and did not provide a method of coping with the heterogeneous expressions of IaaSs. García-Galán et al. [9] also proposed the use of a FM to describe an IaaS. The authors took EC2 as an example to introduce the specific methods of creating an IaaS FM. Quinton et al. [14] introduced attributes to extend a Boolean FM and then used the extended FM to describe an IaaS. The authors also defined a domain knowledge ontology to cope with the heterogeneous expressions in FMs. Finally, they realized the automatic IaaS configuration. García-Galán et al. [10] proposed a method founded on software product line to configure the Amazon IaaS. The authors described the Amazon EC2, EBS, S3, and RDS services by means of FMs and used an offline solver to find the most appropriate configuration instance. Dougherty et al. [27] also described an IaaS as a FM. They used "feature" to represent the main functional units and "attribute" to represent the energy consumption of each functional unit. Their goal is to automate the IaaS configuration in the most energy-efficient and economical way. However, in [9, 10, 14, 27], the authors did not present a standardized definition for the IaaS FM, and thus the universally applicable methods of creating IaaS service and demand models could not be achieved. Wittern and Zirpins [28] only provided a general conceptual FM adapted to various web services without considering the special requirements of IaaSs. In this paper, we define the metamodel and the domain model for IaaSs in a standardized manner and provide the general method for creating the IaaS service and demand models and then propose the automatic discovery and selection methods of IaaSs based on them.

## 3. Feature-Model-Based IaaS Description Method

*3.1. Feature Model Basics.* FMs are treelike data structures, where each node represents a product feature, and the edge between two nodes represents the relationship between the corresponding features. The root of a FM represents the domain object that is described. Features are bound by means of vertical and cross-tree relationships, which define how features can be grouped in a product. The father feature and its child features are usually constrained by mandatory, optional, or cardinality vertical relationships, and the cross-tree relations mainly include REQUIRES and EXCLUDES [13, 15]. In a FM, features can represent specific functional or nonfunctional characteristics of a demand, service, component, or platform, and so forth and can also be used as abstract features that represent domain decisions, such as Linux-based features. The mandatory relationship between the father feature and its children means that the child feature must be selected when its parent feature is selected, while the optional relationship between them means that the child

feature can be selected or not selected when its father feature is selected. In addition, cardinality is primarily used to define the range of the child features that can be selected when their parent feature is selected. A FM that only contains features and feature relationships is called a Boolean FM. A Boolean FM can be extended as an attributed FM by defining feature attributes, which can represent more complex domain requirements and product models. All functional and nonfunctional property parameters of a product can be described as feature attributes. For an attributed FM, product customization can be realized by feature selection, attribute assignment, and the definition of constraint relationships among the selected features and their attributes. A correct and reasonable FM should be able to completely describe the product configuration space, and any semantic conflicts should not exist in all defined relationships.

*3.2. Feature Modeling Method for IaaSs.* The goal of IaaS feature modeling is to describe IaaSs and user demands for an IaaS effectively and accurately. We first define the metamodel of an IaaS to regulate the concepts in IaaS FMs and the relations among these concepts in a standardized manner, in order to provide the grammar fundamentals for creating IaaS FMs. Then we define the domain model of an IaaS based on the metamodel and the IaaS domain knowledge. The advanced and abstract characteristics of an IaaS are classified and unified by the domain model, which provides the conceptual basis for establishing IaaS FMs. Finally, the specific methods of creating the IaaS service and demand models are presented.

*3.2.1. Metamodel.* First, the metamodel of an IaaS is defined using UML, as illustrated in Figure 1. A metamodel is a description model of the model, which is the semantic description of the concepts and concept relationships in the model. The defined metamodel mainly includes three kinds of concepts, features, attributes, and relationships.

The main functional and nonfunctional characteristics of an IaaS instance are defined as features, such as VM, OS, location, and cost model. Features can be divided into a root feature, group features, and solitary features according to their organizational structures. Features can also be divided into a root feature, abstract features, and leaf features according to the feature abstract level. Among these, leaf features represent the most specific functional or nonfunctional characteristics. Feature group is an abstract feature used to accommodate the group features of the same type, and it can be the child of a solitary feature or another feature group. Features in the feature group are called group features, which are used to represent a series of features with common characteristics. For example, an OS feature group may contain two group features, Windows and Linux. All the features that are not included in the feature group are called solitary features, which are used to describe the relatively simple object characteristics. A solitary feature can be used as a feature group, containing a series of group features as its children. For instance, OS is a solitary feature, which contains WINDOWS, LINUX, and other group features as its children. Solitary features can be defined as mandatory or optional type, while feature cardinality is defined to constrain the range of group features that can be selected at the same time. Abstract feature and leaf feature and group feature and solitary feature are two independent feature classifications, which means that not only can abstract feature and leaf feature be defined as group feature or solitary feature, but group feature and solitary feature can also be defined as abstract feature or leaf feature.

The second kind of main concept in the metamodel is feature relationship, which is divided into two categories. The first is the refinement relationships in the vertical direction of the feature tree; that is, features are gradually refined (including decomposition, specialization, and characterization) from the root feature down to the leaf features, and the features between the root and the leaves represent the different abstract levels of characteristics. Each IaaS instance's configuration options must be refined to leaf features. The second is the horizontal cross-tree constraint relationships, which mainly contain three kinds of relationships; namely, REQUIRES, EXCLUDES, and COMPLEX CONSTRAINTS. The cross-tree constraints are generally used to describe the constraint relationships among the features of different categories and their attributes in an IaaS instance configuration; for example, price features are commonly constrained by location features (e.g., the prices of IaaS instances in Asia and North American are likely to be different). The third kind of main concept in the metamodel is feature attribute, which is usually used to describe the common properties of IaaS functions or nonfunctional characteristics, such as the numbers of vCPU, memory, and hard disk capacity. By defining the constraint relationships among the features and their attributes, we can accurately describe the configuration options of an IaaS instance. Several important definitions in the metamodel are as follows.

*Definition 1* (FEATURE GROUP CARDINALITY). Assuming that $0 \le n \le n' \le k$, where $n$, $n'$, and $k$ are nonnegative integers and $k$ is the number of all group features, we define $\langle n - n' \rangle$ as a feature group cardinality, in which $n$ is the lower bound and $n'$ is the upper bound, and $\langle n - n' \rangle$ means that when the feature group is selected, at least $n$ group features must be selected simultaneously, while at most $n'$ group features can be selected simultaneously.

*Definition 2* (REQUIRES). Assuming that features A and B are from the same FM, if feature A is selected at any time it means that feature B is selected simultaneously, and then we say that feature A REQUIRES feature B.

*Definition 3* (EXCLUDES). Assuming that features A and B are from the same FM, if feature A is selected at any time it means that feature B is not selected simultaneously, and then we say that feature A EXCLUDES feature B.

*Definition 4* (COMPLEX CONSTRAINTS). By means of arithmetic symbols such as +, −, ∗, /, and mod, and relationship symbols such as >, >=, <=, <, ==, and !=, and logical symbols such as AND, OR, NOT, IMPLIES, and IFF (if

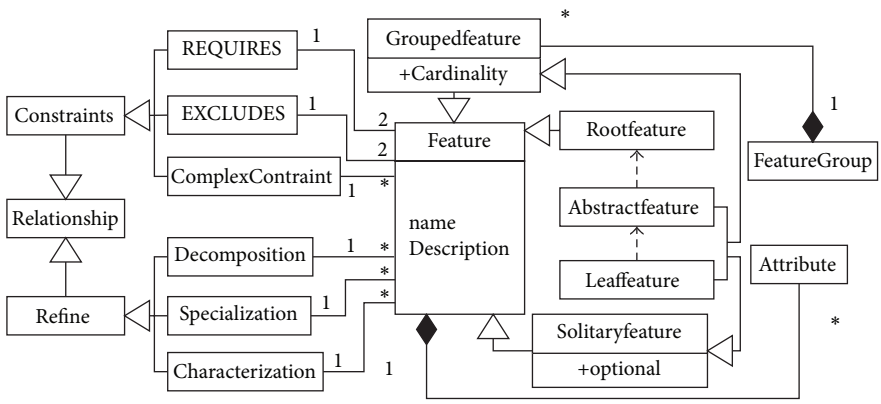



Figure 1: Metamodel of an IaaS.

only if), we can associate certain features with certain feature attributes in the FM to form a variety of logical expressions, which are used to describe a particular relationship that must be met among these features and feature attributes. We call this type of relationship among the features and feature attributes COMPLEX CONSTRAINTS.

In summary, we perform a design of the metamodel that provides the grammar supports for creating a structured IaaS FM. We give the definition of the feature, attribute, relation, and other key concepts in an IaaS FM. In addition, we introduce the roles and usage of various features and also describe the vertical refinement relations and the cross-tree constraints in an IaaS FM.

*3.2.2. Domain Model.* Because the current IaaS description has no uniform information structure and terminology, the information structure and terminology of IaaS FMs will appear random if they are created relying only on the syntax basis of the metamodel. Thus, it will be difficult to perform the matching of the IaaS service and demand models. In fact, there is a consistent functional and nonfunctional characteristics category existing in IaaSs. Therefore, an IaaS domain model is created based on the metamodel and the IaaS field knowledge so as to unify IaaS FMs' information structure and terminology. The domain model classifies and regulates the advanced and abstract features of IaaSs and defines all necessary attributes, thus providing a conceptual basis for the creation of IaaS FMs. The major structure of the domain model is illustrated in Figure 2. A summary of important attributes of the main IaaS features is shown as Table 1.

The definition of the domain model follows the principles below: first, those features and feature directories that are more definite and widely recognized should be included as completely as possible; as such, based on the domain model, the IaaS service and demand models can be created in a more unified and consistent structural form. Second, it is necessary to avoid too much refinement of the abstract features that may have a variety of representations, so as to meet the actual needs for describing IaaSs. Third, the important common attributes of the IaaS function characteristics should be defined in a uniform manner, because these attributes will be the main parameters that are evaluated and compared when performing the supply and demand matching and horizontal comparison of IaaSs. Fourth, the domain model should maintain good extensibility because of the constantly development and updating of IaaSs in cloud computing.

Table 1: Summary of important attributes of the main IaaS features.

| Feature | Attribute | Type | Remarks |
|---|---|---|---|
| IaaS | periodcost | real | Periodic cost |
| | availability | real | |
| | securitylevel | real | Security level |
| | netdelay | int | Network latency |
| VM | HDD | string | Solid state or magnetic |
| | memory | real | |
| | SPECCPU2006 | int | CPU benchmark value |
| | scaling latency | int | Scaling resource latency |
| LoadBalancer | costmonth | real | Cost per month |
| | num | int | Load balancer number |
| Network | availability | real | |
| | supportipv6 | bool | Support for IPv6 |
| | estimatetraffic | real | Estimate traffic per month |
| Storage | availability | real | |
| | persistence | real | Storage persistence |
| Freetraffic | freelimit | real | Free traffic per month |
| | extraprice | real | Extra traffic price |

*3.2.3. Service Model.* As an example, Figure 3 illustrates the major structure of the created IaaS service model from DigitalOcean [29]. Based on the metamodel and the domain model defined above, the service model of an IaaS can be created to represent the actual IaaS offerings. The key points of creating an IaaS service model are as follows: first, select features from the domain model and refine the abstract features among them according to the original IaaS description. Second, because the service model has inherited all the attributes of the related features in the domain model, it only needs to select the required feature attributes to specify

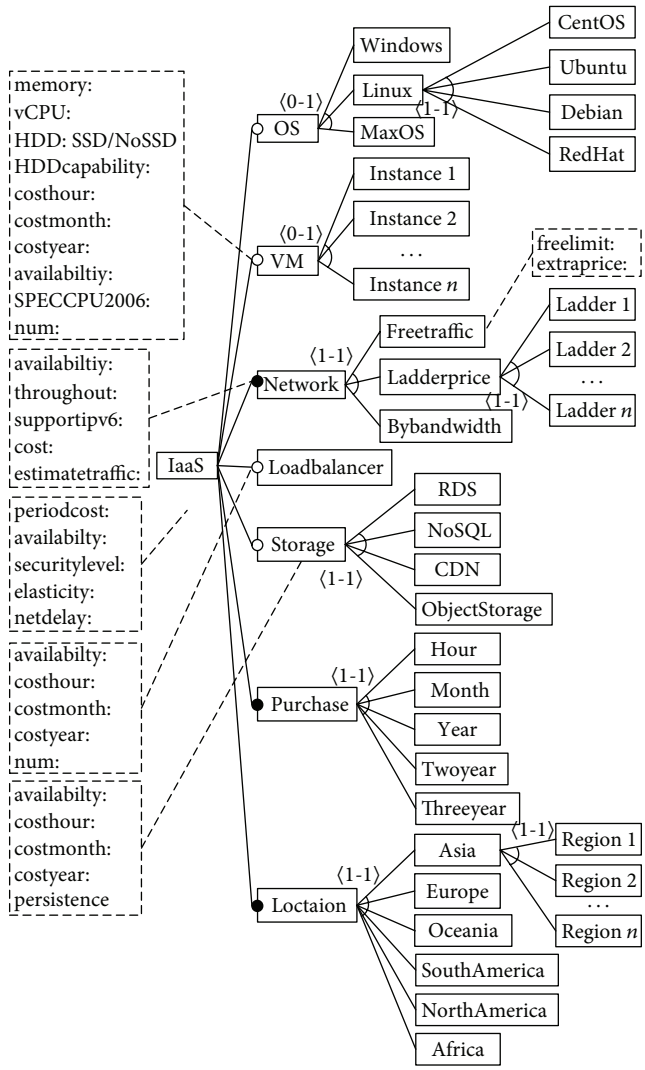



Figure 2: Major structure of the IaaS domain model.

their value ranges. Finally, the entire IaaS configuration space is described by defining many COMPLEX CONSTRAINTS and, in general, each of the COMPLEX CONSTRAINTS can represent an IaaS configuration instance. These COMPLEX CONSTRAINTS are included in the plain-text file of the IaaS service model; see Section 7.1.3 for details.

*3.2.4. Demand Model.* Figure 4 illustrates the major structure of a demand model for computing services. The demand model is used to describe user needs in an IaaS, which is created by using a similar method of creating the IaaS service model. Similarly, the first thing that should be done is to select features from the domain model and refine the abstract features among them according to user needs. The demand model has also inherited all the attributes of the related features in the domain model. It is also necessary to select the required feature attributes to specify their values or value ranges. The demand model represents a minimum of service requirements, meaning that the IaaS instances satisfying user needs usually number more than one. Therefore, when the IaaS supply and demand matching is completed, choosing the best IaaS instance is required. For example, when it is necessary to optimize service cost, we can select the IaaS instance satisfying user needs that has the minimum cost.

*3.3. Concept Ontology of IaaS FMs.* While the domain model has achieved a unified classification and specification of the advanced and abstract features in IaaS FMs, it is still necessary to refine the abstract features into the leaf features in the process of creating an IaaS FM, which may lead to the use of different terms to represent the same concept; for example, given geographical region feature, the terms "region" or "location" may be used in different IaaS FMs. In order to make all IaaS FMs mutually "understood" so that matching the IaaS supply and demand can be carried out, we have created an IaaS concept ontology to realize the mutual mapping of equivalent feature concepts; some parts of this are illustrated in Figure 5. The concepts of "VM" and "Virtualmachine," "Region" and "Location," "Windowserver_2003" and "WindowsServer2003," and "Ladderprice" and "Multilevelprice"

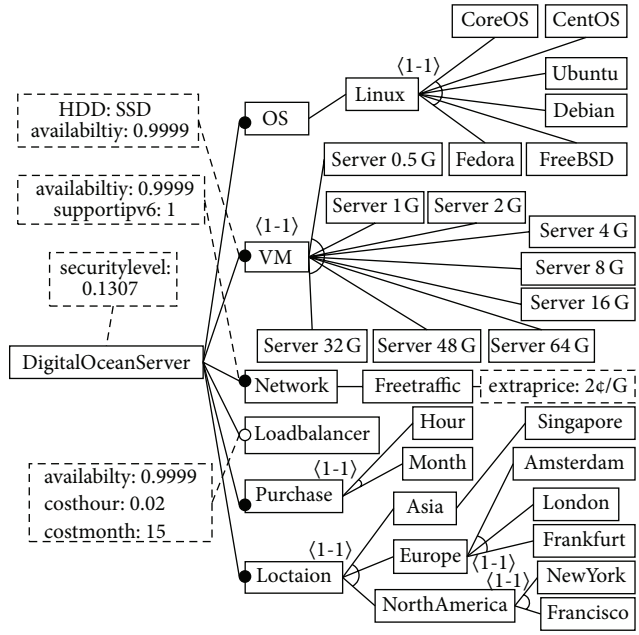



Figure 3: Major structure of the DigitalOceanServer service model.

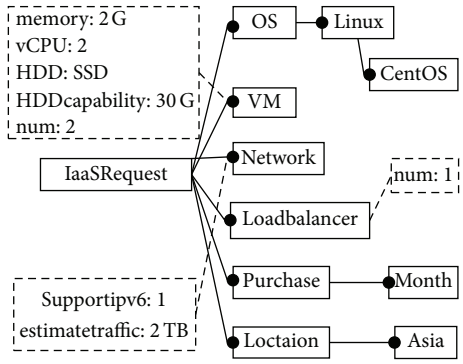



Figure 4: Major structure of an IaaS demand model.

are all equivalent. In addition, it should be noted that the domain model has unified and defined all important feature attributes considering the consistency of the IaaS feature attributes; therefore, when the IaaS FMs are created, it must be done using the unified feature attributes specified by the domain model.

*3.4. Characteristics Analysis of the Description Method.* First, it is not difficult to create and maintain an IaaS FM with the metamodel and domain model acting as the grammatical and concept bases. Graphical tools or customized machine languages can be developed to create and edit an IaaS FM. Second, our IaaS description method has good extensibility and compatibility and can continuously update the domain model to improve the conceptual repository according to the development of IaaSs. The concept ontology of IaaS FMs acts as the bridge between the different terms with the same meaning in different IaaS FMs. Finally, the metamodel provides strict grammatical norms for the creation of an IaaS FM, and this paper will check the validity of an IaaS FM based on the automatic analysis operations of a FM (see Section 5.1).

## 4. IaaS Discovery and Selection Method

Based on the formal IaaS description presented in Section 3, we propose a systematic discovery and selection method for IaaSs. An architecture of the system implementing the method is designed as Figure 6.

Four roles are managed in the architecture: User, Broker, IaaS Provider, and MTCCAs (cloud service monitoring, testing, certificating, and comparing agencies). The IaaS Provider usually publishes the IaaS service information on its portal website and sometimes chooses to perform a service certification or registration via MTCCAs. For example, the IaaS Provider may submit a security assessment report or perform a security certification in Cloud Security Alliance [30] (CSA).

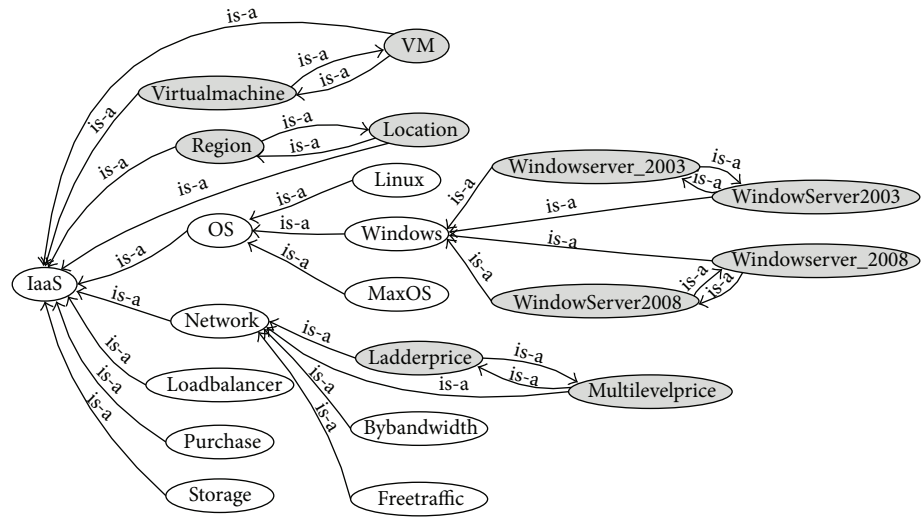



Figure 5: Concept ontology of IaaS FMs.

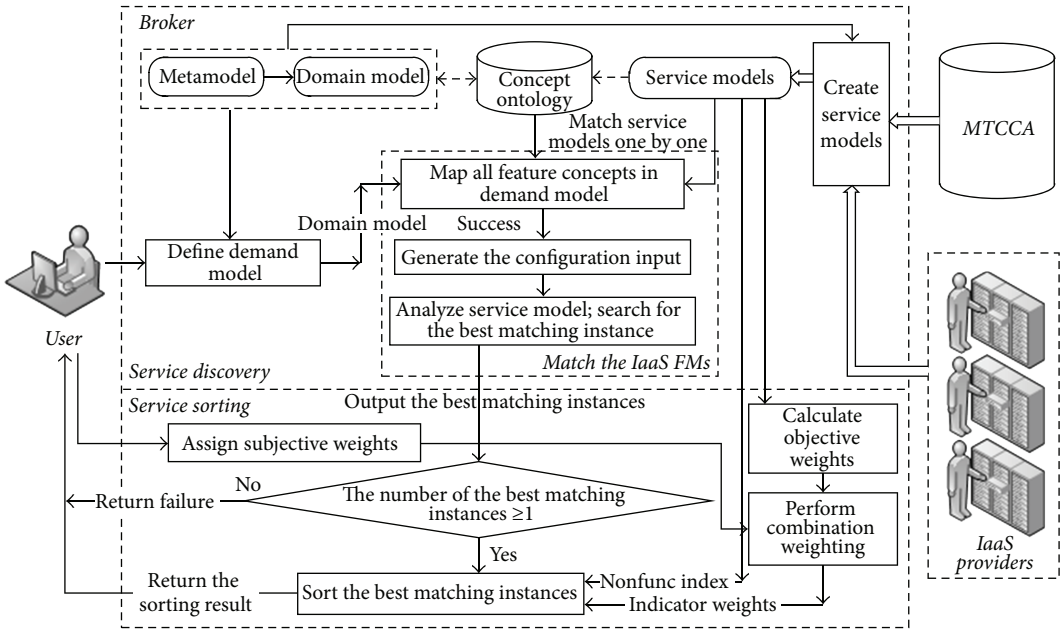



Figure 6: Architecture of the discovery and selection system of IaaSs.

MTCCAs such as CloudHarmony, CloudScreener, CSA, and Cloudorado mainly perform the monitoring, performance testing, security certification, or synthetic comparison of cloud services. The entirety of IaaS service information can be obtained by combining the IaaS information from the provider portal and the information from MTCCAs. The user needs to create and submit the IaaS demand model and perform the subjective weighting for important nonfunctional metrics. The Broker is used as the functional core of the architecture.

The Broker's main functions are divided into two parts: service discovery and service ranking. Regarding service discovery, the Broker first needs to regularly maintain the IaaS domain model based on the IaaS domain knowledge by adding new features and feature directories. Second, the Broker uses the IaaS related information from the IaaS Providers and MTCCAs to create the IaaS service models. The Broker is also responsible for the regular maintenance of the concept ontology of IaaS FMs. Finally, the Broker takes the user demand model, the concept ontology, and the IaaS service models as the inputs of the IaaS FM matching module.

The matching of supply and demand on IaaS FMs consists mainly of two steps. The first step is to map the feature concepts in the demand model into the feature concepts in the current IaaS service model. If the related concepts cannot be completely mapped successfully, the matching process will proceed to the next IaaS service model to continue. If

the mapping is successful, all the feature concepts in the demand model are converted into the feature concepts in this IaaS service FM, and then the IaaS demand configuration suitable for this service FM can be generated according to the user demand model. The second step is to search for the best matching instance by automatically analyzing the current service model. If no instances satisfying user needs are searched, the matching of supply and demand on the current IaaS service model fails. Otherwise, it needs to find the best one from the IaaS configuration instances satisfying user needs by optimizing the cost and then terminate the matching process on the current IaaS service model and proceed to the next IaaS service model to continue until the matching with all IaaS service models is completed. Finally, the IaaS FM matching module returns all the best matching instances as the output results. This service discovery method will be detailed in Section 5.

Another main function of the Broker is to rank all the best matching instances. If the number of the best matching instances returned by the IaaS FM matching module is 0, it shows that there is no IaaS instance satisfying the user demand, and then failure information will be returned. If the number is 1, then the single instance will be returned directly. When the number exceeds 1, the system will automatically start the next stage, namely, service sorting. The sorting process takes the critical nonfunctional metric set and the metric weight vectors as inputs and sorts all the best matching instances by calculating their comprehensive evaluations. Finally, the ranking list is delivered to the user. This service sorting method will be detailed in Section 6.

## 5. Automated Analysis and Matching of IaaS FMs

FMs unify the descriptions of the IaaS services and demands, which paves the way for automated IaaS discovery. IaaS discovery is the process of searching for the IaaS instances that satisfy the user demand from many IaaS offers through matching the IaaS supply and demand. To complete matching the supply and demand on IaaS FMs, it is necessary to analyze the IaaS service models automatically. Usually, there are many features, feature attributes, and COMPLEX CONSTRAINTS in an IaaS service model; moreover, the IaaS configuration space is also large. Therefore, it is impossible to use an artificial method to analyze the IaaS service model, and thus an automatic mechanism is needed. In fact, the automated analysis of feature model (AAFM) has become a special technique; it is a set of automated analysis operations that extract information from the FMs [31]. Each analysis operation can be regarded as a black-box process, with the FM and the specific operation parameters as inputs, and it can retrieve and output the corresponding results based on the different goals of analysis. These analysis operations mainly include checking whether the FM syntax is correct, judging whether there is a semantic conflict, and verifying whether the input product configuration information is valid.

### *5.1. Automated Service FM Analysis*

*5.1.1. Checking the Service FM Validity.* Before performing other analysis operations on an IaaS service model, it is necessary to determine its validity by checking its syntactical correctness and semantic conflicts. Semantic conflict checking is necessary because some semantic conflicts are very prone to be introduced when COMPLEX CONSTRAINTS on features and attributes are created in a FM. The validation operations of an IaaS service model mainly include the following:

(1) Judging the syntactical correctness of an IaaS service model text file. In Section 3, we have described the method of creating an IaaS service model in graphics. In fact, in order to perform the automated analysis on a service model, we have to choose a language to express the service model, such as XML or a custom language [15]. We just have to define the XML schema or the syntax of the language chosen before using them. We must comply with the relevant syntactical constraints of the language used when creating the service model text file.

(2) Checking whether the service FM is empty. It should be noted that any contradictions among the relations and constraints of an IaaS service model could lead to the failure of the representation of the IaaS, and we can discover the possible contradictions by analyzing the IaaS service model.

(3) Checking whether the type definitions of all features are reasonable and whether there are semantic conflicts in the cross-tree constraints. For example, when the parent feature is a mandatory type, the only child feature is naturally a mandatory type. Otherwise, an error is generated. The cross-tree constraints are also prone to semantic conflicts; for example, if a REQUIRES constraint between a mandatory feature and another optional feature is created, a semantic conflict occurs.

(4) Checking whether the feature cardinalities are reasonably set. For example, there might be a conflict between the min value of feature group cardinality and the type of feature group. When the feature group is mandatory type, the min value of its cardinality must be no less than 1. When the feature group is of optional type, the min value of its cardinality must be 0. In addition, the maximum value of group cardinality should be no more than the number of features in the feature group. Otherwise, a semantic conflict occurs.

*5.1.2. Checking the IaaS Configuration Validity.* An IaaS configuration is defined as a set of options for running an IaaS instance. For a computing instance, these options correspond to the predefined configuration points, such as OS, resource location, purchase mode, CPU size, and memory size. If at least one instance satisfying the user needs can be found in the IaaS service model according to the submitted configuration,

then the submitted configuration can be considered valid; otherwise, it is invalid. For an attributed FM, when the feature selection and attribute values assignment are completed, and the relationship constraints among the selected features and their attributes are satisfied, the service instances satisfying the configuration needs are determined.

*5.1.3. Searching for the Best IaaS Instance.* Usually, after a user configuration demand is inputted, many IaaS instances satisfying the demand will be searched from the service model; however, only the best one will be returned. The IaaS instances from the same provider are generally very similar in terms of availability, network latency, security, and other nonfunctional properties, while the cost difference is relatively large. Therefore, we choose to optimize the matching instances based on cost; that is, to search for the one with the least cost among all instances satisfying the user needs from the same provider.

*5.2. Matching between the Service and Demand Models.* The supply and demand matching operations between the service and demand models include the following:

(1) Generating the configuration input suitable for the service model from the demand model. First, all of the feature concepts in the demand model are mapped into the feature concepts in the service model based on the IaaS concept ontology such that the demand model becomes completely consistent with the service model conceptually. Second, the configuration input suitable for the service model will be generated by the mapped demand model. The configuration input represents the user needs, which mainly contain three aspects of information: features, attributes, and constraints.

(2) Searching for the configuration instance with optimal cost by analyzing the IaaS service model. This step will use the IaaS demand configuration generated in step (1) as an input and perform the automated analysis on each service model to search for the best configuration instance based on the method outlined in Section 5.1.3. If the analysis results show that the input configuration is invalid, then the matching of supply and demand on IaaS FMs fails. If the results show the input configuration is valid and many configuration instances satisfying the user needs are found, then the minimum-cost instance will be returned.

## 6. IaaS Ranking

According to our IaaS discovery and selection method, a series of best matching instances will be returned after completing the matching between the demand model and all service models (otherwise, service discovery fails). These best matching instances are the minimum-cost instances from different providers that can meet the users' functional demands. Users will choose the most satisfactory one from these instances. The usual approach is to evaluate and compare, synthetically, the important nonfunctional metrics affecting the IaaS QoS, and then make a choice. It is a multicriteria decision-making problem.

*6.1. Important Nonfunctional Metrics of IaaSs.* The establishment of a unified and effective QoS-evaluating metric system is the premise of scientific decision-making. However, the existing measurement metric systems for IaaSs are either inadequate [32, 33] or short of operational measurement methods [6, 25]. The important nonfunctional metrics affecting the IaaS QoS fall into three categories: performance, economy, and security [7]. We study and determine the main metrics affecting the IaaS QoS and their measurement methods.

*(1) Cost.* The economy of cloud computing is its prominent advantage, and it is also one of the factors the user is most concerned about. The cost of an IaaS instance depends on its resource allocation. If the configuration parameters of the IaaS instance are uncertain, then its cost cannot be obtained. If the service and demand models match successfully, the cost optimization of the matching instances will be performed. Thus the best instance and its cost can be obtained.

*(2) Availability*. This is a parameter usually promised by most providers in their service level agreement (SLA). In a certain period, the availability of the IaaS functional unit can be generally calculated in the following ways:

$$\begin{aligned}&\text{virtual machine availability} = \frac{\text{uptime}}{(\text{uptime} + \text{downtime})},\\ &\text{storage availability}\\ &= \frac{\text{number of successful requests}}{(\text{number of successful requests} + \text{number of failed requests})}.\end{aligned} \tag{1}$$

Availability can be aggregated in the following ways:

$$\begin{aligned}&\text{computing type IaaS availability}\\ &= \begin{cases}\text{virtual machine availability} \times \text{network availability} & (\text{where virtual machine number} = 1)\\ \text{load balancer availability} \times \left(\sum_{i=1}^{n}\left(\left(\frac{1}{n}\right) \times \text{virtual machine availability}\right)\right) \times \text{network availability} & (\text{where virtual machine number} > 1).\end{cases}\end{aligned} \tag{2}$$

*(3) VM Performance Metric*. The VM is the most important computing functional unit in cloud computing, and its performance has a direct impact on the quality of the computing service. Different IaaSs having the same VM

configuration usually show a significant difference in performance testing results. The Standard Performance Evaluation Corporation [34] (SPEC) is a nonprofit organization consisting of computer manufacturers, system integrators, universities, research institutions, and so forth. SPEC aims at establishing and maintaining the standards for evaluating computer systems. CPU2006SPEC is a CPU system testing suite introduced by SPEC, which is the general standard of CPU testing at present, including the SPECint and SPECfp subtesting modules. SPECint tests CPU integer processing performance, while SPECfp tests CPU float point processing performance. In this paper, we use the sum of the SPECint and SPECfp test scores as the final score of the VM performance test.

*(4) Data-Storage Persistence*. This is used to evaluate the possibility that the stored data is lost or damaged, which is represented by calculating the percentage of the lost or damaged data in a certain period: intact disk number/(intact disk number + lost or damaged disk number).

*(5) Elasticity.* The load of cloud computing applications is usually dynamic, such as e-commerce applications, which requires the IaaS resources to have prominent elasticity. Perfect elasticity enables users to adjust resources quickly according to load changes, so as to better meet the requirements of applications and avoid the purchase of redundant resources to save costs. Time is a crucial aspect in elasticity, which can be measured by the speed of response to the dynamic workload [33]. Considering that the differences in the resource release speeds of computing instances are relatively small, we have learned from the method proposed in [32], which tests VM scaling latency to measure resource elasticity. Scaling latency includes VM provisioning latency and system booting latency.

*(6) Network Latency*. This means the network delay from an IaaS resource to a target user area, which also has an important impact on the IaaS QoS. If there is an average of 20 s of network latency, the target users will face unendurable service response delay when they access the applications deployed on the IaaS resources.

*(7) Security*. Cloud computing is facing many security threats, such as data destruction, network attacks, privacy leaks, and VM attacks, so security is also a very important metric. In [35], we proposed a systematic method for measuring and evaluating cloud security (referred to as MMECS in this paper). MMECS measures and evaluates cloud security by calculating the security advantage of each cloud provider based on the Consensus Assessment Initiative Questionnaire (CAIQ) reports on cloud security, which have been submitted to the CSA by the cloud providers. CAIQ is a cloud security assessment questionnaire based on the Cloud Control Matrix (CCM), while the CCM is proposed by the CSA and defines a list of security control requirements that the cloud providers and their services should fulfill. The CCM references authority security standards and frameworks in the industry, such as SP800-53 R3, FedRAMP, ISO/IEC 27001, COBIT, and ENISA IAF. First, the evidences of security compliance provided by the CAIQ report are scored, and then the obtained score is used to measure the credibility of the CAIQ report. Second, we use the credibility of the CAIQ report to correct the original score of each security control domain in the CAIQ report. Finally, the security advantage of each IaaS Provider is calculated based on the corrected scores of control domains. The results are relatively objective and convincing, so we use the MMECS method to calculate the security metric of each IaaS Provider.

*(8) Reputation*. This is an important attribute of an IT service, and it is also a metric usually compared when users choose services, such as e-commerce services. Although many cloud service reputation calculation models have been proposed in the academic community, they are still in the theoretical research stage and not yet commercially available. Therefore, it is very difficult to obtain available reputations for IaaSs. Although the market share of an IaaS cannot substitute for its reputation, it can basically reflect the user's choice tendency and the service credibility to a certain extent, so we use the market-share data provided by CloudHarmony as a measure of the IaaS reputation.

*6.2. Combination Weighting Method for Decision Metrics.* The scientific and rational allocation of metric weights is of great importance to multicriteria decision-making problems. According to the data sources for calculating weights, the metric weighting methods fall mainly into two categories: the subjective and objective weighting methods. The subjective weighting method assigns metric weights based on a decision maker's subjective preference, such as the analytic hierarchy process (AHP) and Delphi. The objective weighting method calculates the metric weights based on their actual values, such as the maximizing deviations method, the entropy method, and principal component analysis [36]. The subjective weighting method reflects the decision maker's personal preference; however, this method is easily influenced by the knowledge and experiences of the decision maker. For the objective weighting method, there is an objective standard for the weight distribution, but it ignores the decision maker's subjective preferences; metric weights that are calculated totally based on the original metric data may deviate markedly from the subjective desires of the decision maker.

For acquiring more reasonable and scientific evaluation results, some researchers combine the two methods together to carry out the combination weighting [36–39]. The existing research mainly uses two representative ways of combining the subjective and objective weights: the multiplicative synthesis and weighted linear combination methods [36–39]. The multiplicative synthesis method generally uses the formula $q_j = a_j w_j / \sum_{j=1}^{n} a_j w_j$ (where $a_j$, $w_j$, and $q_j$ are the subjective weight, objective weight, and combination weight of the $j$th metric, resp., and $n$ is the number of metrics). However, this approach tends to make the large results even larger and the small ones even smaller, and sometimes it may lead to a very unreasonable result. For example, given

the subjective weights (0.6, 0.3, 0.1) and the objective weights (0.7, 0.2, 0.1), using the multiplicative synthesis method we can see from the combination result (0.857, 0.122, 0.02) that the first metric is obviously enlarged, while the third metric is decreased to the point where it can be ignored, which is obviously unreasonable. Research suggests that the multiplicative synthesis method is more suitable for the case in which the number of metrics is large and the weight distribution is more uniform [36, 37]. Regarding the weighted linear combination method, the general formula is $q_j = \delta a_j + (1 - \delta)w_j$ (where $a_j$, $w_j$, and $q_j$ are the subjective weight, objective weight, and combination weight of the $j$th metric, resp., and $0 \leq \delta \leq 1$). This method overcomes the shortcomings of the multiplicative synthesis method and thus is more widely used [36–39]. The key of this method is how to determine the combination coefficient of the two weighting methods. Based on the above analysis, we plan to use the weighted linear combination method to combine the results of subjective and objective weighting. The methods proposed in [40] are used to do the subjective weighting, while the entropy method is adopted to do the objective weighting.

*6.2.1. Improved AHP Weighting Method Based on Sorting.* The AHP is originally conceived as a multicriteria decision-making method and is widely used as a subjective weighting method for decision metrics. The top-down hierarchical structure is used to model the relationship among decision elements in the AHP. The top node of the hierarchical structure represents the decision goal, such as QoS, and the following layer nodes are the related metrics affecting the decision goal, such as security, availability, and cost, which can be further divided. The lowest layer nodes are the alternatives. The AHP has a unique method for determining metric weights, which defines 1–9 scales to measure the relative importance degree of two metrics by combining the qualitative and quantitative ways to measure them. The relative importance degree is defined as the ratio between the weights corresponding to two metrics related to the same goal. The relative importance degree of any two metrics is determined by paired comparisons; after that, the judgment matrix is obtained. Next, to check the consistency of the judgment matrix is unavoidable. For example, if the relative importance degree of A and B equals 2, the relative importance degree of B and C equals 2, and the relative importance degree of A and C equals 4, and then the weight distribution of A, B, and C is consistent; otherwise, it is inconsistent. When the weight distribution of metrics is not consistent, the relative importance degrees must be adjusted continuously until the judgment matrix becomes consistent. Finally, the weights of all metrics will be obtained by solving the judgment matrix. How to adjust the inconsistent judgment matrix is a difficult problem in the AHP. We have improved the AHP weighting method by referencing the method proposed by Li and Wang [40], so that the obtained judgment matrix fully meets the requirement of consistency, as follows.

*(1) Definition of the Relative Importance Degree Scale.* The scale $1.5^{\beta}$ ($\beta = 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4$) is defined to depict the relative importance degree between metrics. The relative importance degree is $1.5^0$ meaning that the two metrics are equally important. The value of 1.5 means that the former metric is slightly more important than the latter. The value of $1.5^2$ means that the former metric is more important than the latter metric. The value of $1.5^3$ means that the former is obviously more important than the latter. The value of $1.5^4$ means that the former is absolutely more important than the latter. When $\beta$ take values of 0.5, 1.5, 2.5, and 3.5, it indicates the intermediate state of the two adjacent relative importance degrees, respectively.

Table 2: Distribution of metric relative importance degrees for the cases of $\beta$ and $n$ depicted.

| $\beta$ | 0 | 0.5 | 1 | 1.5 | 2 | 2.5 | 3 | 3.5 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| $n = 1$ | $a_3$ | — | $a_2$ | — | $a_1$ | — | — | — | $a_4$ |
| $n = 2$ | — | — | $a_5$ | — | — | — | — | — | — |

*(2) Determination of the Subjective Weights.* There are three steps in calculating the subjective weights: first, the metrics are divided into several groups according to importance from low to high. Each group represents a level of importance, and the weights of metrics in the same group are all equal or have only small differences, while the weights of metrics in the different groups have significant differences. The main purpose of grouping metrics is to determine the rationality of the results of objective weighting when the combination weighting is carried out in the future. Second, the metrics in each group are sorted according to their importance ranging from low to high, while the metrics with the same importance are sorted by their numbers. Finally, the subjective weights of decision metrics are determined according to the following method. The left-hand first one of the sorted metric sequence is taken as the main reference and the left-hand adjacent one of the current metric is taken as the auxiliary reference; then, the relative importance degree between the current metric and the main reference metric and the relative importance degree between the current metric and the auxiliary reference metric are determined. According to this method, the relative importance degree of each metric in the sequence is assigned, in turn, from the left-hand to right-hand order. Finally, the subjective weight of each metric can be obtained based on all the relative importance degree values. Table 2 shows the case of the distribution of relative importance degrees, in which $\beta$ represents the relative importance degree index for each metric, and the maximum $n$ in each column represents the number of metrics with the same weight. The determination process of the relative importance degree is as follows.

First, the metric with the smallest degree of importance, namely, the left-hand first one in the sorted metric sequence, is placed in the cell where $\beta = 0$ and $n = 1$. In turn, other metrics in the sequence are compared with the left-hand first metric and its left-hand adjacent metric, and then it is placed into the corresponding cell based on the comparison results. For example, for the metric sequence $a_3$, ($a_2$, $a_5$), $a_1$, and $a_4$, the process in detail is as follows: At first, $a_3$ is placed into the cell where $\beta = 0$ and $n = 1$. Meanwhile, $a_2$ is placed

into the cell where $\beta = 1$ and $n = 1$ because $a_2$ is slightly more important than $a_3$. Considering that $a_5$ is slightly more important than $a_3$ and is as equally important as $a_2$, $a_5$ is placed into the cell where $\beta = 1$ and $n = 2$. In addition, $a_1$ is placed into the cell where $\beta = 2$ and $n = 1$ because $a_1$ is more important than $a_3$ and is slightly more important than $a_5$. Finally, $a_4$ is placed into the cell where $\beta = 4$ and $n = 1$ because $a_4$ is absolutely more important than $a_3$ and is more important than $a_1$.

It has been proved that the judgment matrixes determined by the above method are consistent [40]. The subjective weight of each metric can be calculated according to the formula $w_i = 1.5^{\beta_i} / \sum_{j=1}^{n} 1.5^{\beta_j}$, where $\beta_i$ and $\beta_j$ are the relative importance degree indexes of metrics $i$ and $j$, respectively.

*6.2.2. Calculation of the Objective Weights.* The entropy method is a commonly used objective weighting method that obtains the objective weight of each metric by calculating its entropy based on the raw metric data. Entropy is a measurement of system uncertainty; the higher the entropy value of a system, the greater its uncertainty, and the less its information, and vice versa [41]. We assume that the multicriteria decision-making problem has $m$ alternatives and $n$ decision metrics, and any metric can be expressed as $x_{ij}$ $(1 \le i \le m, 1 \le j \le n)$. The objective weights of decision metrics are then calculated according to the following steps:

The first step is to achieve the normalized decision matrix $\mathbf{P} = (p_{ij})_{m\times n}$ by normalizing the decision matrix $\mathbf{X} = (x_{ij})_{m\times n}$, so that each decision metric is the profit type and its value is between 0 and 1.

The second step is to calculate the entropy of each metric. The uncertainty of the importance degree of any decision metric can be measured by calculating its entropy according to the following formula:

$$E_j = -\sum_{i=1}^{m} p_{ij} \ln p_{ij} \quad (j = 1, 2, \ldots, n). \tag{3}$$

From the above formula, when a certain metric in all alternatives tends to be equal, the entropy of the very metric approaches its maximum value, and thus the uncertainty of the importance degree of the metric tends to be greater. The maximum entropy $E_{\max} = \ln m$, so $E_j$ can be normalized by $E_{\max}$ as follows:

$$e_j = -(\ln m)^{-1} \sum_{i=1}^{m} p_{ij} \ln p_{ij} \quad (j = 1, 2, \ldots, n). \tag{4}$$

The third step is to calculate the difference coefficient of each decision metric by the following formula:

$$d_j = 1 - e_j \quad (j = 1, 2, \ldots, n). \tag{5}$$

The smaller the metric's entropy, the greater the metric's difference, and the greater the role of the very metric in evaluating services and vice versa.

The fourth step is to calculate the objective weight of each metric as follows:

$$w_j = \frac{d_j}{\sum_{j=1}^{n} d_j}. \tag{6}$$

*6.2.3. Combination Weighting Method Based on Comprehensive Analysis.* Although the subjective weight is subjective and less accurate, it reflects the preference of the decision maker and has a superior rationality in sorting decision metrics by their importance. Therefore, the combination weighting should be mainly based on the subjective weights, and the combination coefficient of the subjective and objective weight vectors can be determined by analyzing the difference between the metric weight rankings obtained by the two weighting methods. Assuming that the subjective and objective weight vectors are identified as $\mathbf{w}_s$ and $\mathbf{w}_o$, respectively, then any linear combinations of them can be expressed as

$$\mathbf{w} = a\mathbf{w}_s + (1 - a)\mathbf{w}_o, \quad 0 \le a \le 1. \tag{7}$$

Assuming that the decision metrics are identified as $P_1, P_2, \ldots, P_n$, after subjective weighting, the decision metrics are divided into $k$ groups by importance, that is, $R_1, R_2, \ldots, R_k$, in which the group $R_2$ is more important than the group $R_1$ and the group $R_3$ is more important than the group $R_2$ and so on. The method of determining the coefficient $a$ is as follows:

(1) If there is a relatively uniform distribution and no essential difference in the weights of all decision metrics, then take $a = 0$, and the result of the objective weighting can be directly used as the result of combination weighting.

(2) If the metric weight rankings obtained by the subjective and objective weighting methods are exactly the same, in order to eliminate the influence of subjective factors, it should take only the objective weighting method to obtain the weights, which indicates that $a = 0$.

(3) If the metric weight rankings obtained by the subjective and objective weighting methods are not exactly the same, while the metric group rankings obtained by the two methods are the same, then take $a = 0.5$. For example, the metric weight ranking obtained by the subjective weighting method is $P_1$, ($P_2$, $P_3$) and ($P_4$, $P_5$) while the weight ranking of metrics obtained by the objective weighting method is $P_1$, ($P_3$, $P_2$), and ($P_5$, $P_4$). Obviously, the two orders of metric weights are not the same, but the orders of metric groups are the same, and in such circumstances, $a = 0.5$.

(4) If the weight rankings of metrics obtained by the two weighting methods are different and the orders of metric groups obtained by the two methods are also different, then take $a = 1$. This situation indicates that the results obtained by the objective weighting method deviate markedly from the subjective preferences of the decision maker, so the obtained objective weights have no reference value. The subjective weighting results will be directly used as the final metric weights.

*6.3. Ranking Alternative Instances Based on Comprehensive Evaluations.* For an alternative IaaS instance $I_i$, it is assumed

that $I_i$ has $k$ important nonfunctional metrics that are expressed as $x_{i1}, x_{i2}, \ldots, x_{ik}$, and the metric weight vector obtained by the combination weighting method is $\mathbf{w} = (w_1, w_2, \ldots, w_k)$; thus, the comprehensive evaluation $U_i$ of $I_i$ can be obtained by formula (8). The comprehensive evaluations of all the best instances returned by the matching of supply and demand on IaaS FMs can be calculated in this way. Finally, these instances are sorted based on their comprehensive evaluation. We have

$$U_i = x_{i1}w_1 + x_{i2}w_2 + \cdots + x_{ik}w_k. \quad (8)$$

## 7. Experiment and Analysis

### 7.1. Experimental Scenario, Data Sources, and Experimental Tools

*7.1.1. Experimental Scenario.* This experiment evaluates the feasibility and effectiveness of the IaaS discovery and selection method presented in this paper. Considering that a computing service acts as the main cloud infrastructure service having a wide range of applications, we specially formulate the following experimental requirements.

A user plans to buy a computing service from an IaaS Provider, and the customers are mainly located in southeastern China. The specific requirements of the computing service include the following: two virtual servers with identical configuration, a load balancer, two virtual CPUs per virtual server, 2 G memory, a solid-state hard disk, 30 G hard disk capacity, CentOS operating system, and support for IPv6; the estimated network traffic per month is 2 TB, and all service resources are located in Asia.

Three moderate-size and well-known IaaS providers, which are very representative and can provide the computing services in the Asia area, are selected as the IaaS providers in the experiment. The three providers are DigitalOcean [29], RackSpace [42], and Linode [43], which are all from the United States and have built their own cloud infrastructure centers in Europe, America, and Asia. Among them, DigitalOcean was founded in 2012, and currently has more than 15000 servers; RackSpace was founded in 1998 and has hundreds of thousands of servers; Linode was founded in 2003 and has tens of thousands of servers.

*7.1.2. Data Sources.* The service-related data of the three providers mainly derives from their portals. In addition, we have selected the following critical nonfunctional metrics for comparison according to Section 6.1. These metrics have been calculated or obtained in the following ways.

*Cost*. We obtain the best matching instances from the three IaaSs and their costs by matching the IaaS service and demand models.

*Security*. According to our measurement and evaluation method of cloud security MMECS, we first download the providers' CAIQ reports from the CSA site and then calculate the security advantage of each provider based on these reports.

*SPECCPU2006*. CloudHarmony carries out periodic performance testing on the service instances of several major cloud providers in the industry and their testing results are relatively objective and fair. We directly download the SPECCPU2006 results of related instances from the CloudHarmony site.

*Availability*. We obtain the availability of each computing service instance by formula (2).

*Network Latency*. Under an 8 M bandwidth network environment in Nanjing, by means of network testing services provided by CloudHarmony, we test the network latency of the best matching instance from each provider ten times. We remove the minimum and maximum values and then use the average value of the remaining eight tests as the final result.

*Market Share*. We directly download the market-share data from the CloudHarmony site.

*Scaling Latency*. We measure scaling latency by recording the time from when the instance is requested to when it is available to be used. We test the scaling latency for each instance ten times and remove the minimum and maximum values and then use the average value of the remaining eight tests as the final result.

*7.1.3. Editing and Analysis Tools of IaaS FMs.* We have selected the FaMa [44] framework to create and analyze the IaaS FM. The FaMa framework is an open-source automated analysis tool for FMs with four different reasoners, which support more than 20 different analysis operations. FaMa supports creating a FM containing attributes and feature group cardinality. FaMa is implemented as an Eclipse plug-in, which supports the following analysis operations: verifying a FM, product number analysis, verifying configuration, product searching, and so forth. FaMa can accept two forms of FM file formats, namely, XML and plain text. The plain-text format is the simplest and clearest among all the formats, so we select it to express the IaaS FM. The syntax of the plain-text format is defined as follows.

As illustrated in Figure 7, a FM is expressed as three sections: the "%Relationships" section, the "%Attributes" section, and the "%Constraints" section. The hierarchical features and the vertical relationships among features are defined in the "%Relationships" section. The features after the colon are the children of the feature before the colon. By default, features are of mandatory type, while optional features are presented in the form of [feature name]. The feature group is represented by cardinality and a set of features in the form: [min, max]{FeatA, FeatB, ...}. The feature attributes are defined in the "%Attributes" section; the definition of each attribute contains attribute type, range, default value, and null value. The attributes can be an enumeration type and defined in the form "FeatName.AttName:[val1, val2, ..., val$N$], defaultVal,

```
%Relationships
LinodeServer: Virtualmachine OS Network [Loadbalancer] Purchase Location;
Virtualmachine: [1,1]{Linode1G, Linode2G, Linode4G, Linode8G, Linode16G,
Linode32G, Linode48G, Linode64G}
OS: Linux;
Linux: [1,1]{Arch, Debian, centos, Ubuntu, OpenSUSE, gentoo, Slackware,
Fedora};
...
%Attributes
LinodeServer.periodcost:real[0 to 10000000],0,0; % Period cost ($)
LinodeServer.availability:real[0 to 1],0,0;
LinodeServer.Securitylevel:real[0 to 1],0.1396,0;
Virtualmachine.HDD: [SSD, NoSSD],SSD,0;
...
%Constraints
Linode1GB AND Month AND (Network.estimatetraffic > Freetraffic.freelimit)
IMPLIES
(Virtualmachine.RAM == 1 GB) AND (Virtualmachine.vCPU == 1) AND (Virtualmachine.HDDcapabiltiy == 24 GB
AND (Freetraffic.freelimit == 204 8G) AND
(Virtualmachine.costMONTh == 10$) AND
(LinodeServer.cost.month = VM.costmonth + Freetraffic.extraprice × (Network.estimatetraffic − Freetraffic.freelimit)); ...
```

Figure 7: Linode service model in plain-text format.

nullVal"; the horizontal relationships among features and the COMPLEX CONSTRAINTS among multiple features and attributes are defined in the "%Constraints" section.

*7.2. Creation and Matching of the Service and Demand Models.* According to the requirement description in this experiment and the collected related information about the three IaaS services, we have created the demand model as illustrated in Figure 4, the DigitalOcean service model as illustrated in Figure 3, the Linode service model as illustrated in Figure 8, and the plain-text format of the Linode service model as illustrated in Figure 7. Owing to space limitations, the RackSpace service model is omitted.

We take Linode as an example to demonstrate the concept mapping between the service and demand models, along with the supply and demand matching.

First, according to the concept ontology of IaaS FMs, the concept "VM" in the demand model is mapped into the concept "Virtualmachine" in the service model, "Location" is mapped into "Region," "CentOS" is mapped into "centos," and so forth. Therefore, the input configuration for the Linode service model can be easily generated by the demand model, including features such as Virtualmachine, CentOS, Network, Loadbalancer, Month, and Asia, as well as attributes such as Virtualmachine.memory = 2 G, Virtualmachine.vCPU = 2. Second, the plain-text format of the Linode service model and the input configuration are submitted to the FaMa platform; the analysis operation code is presented in Figure 9. FaMa first verify the validity of the Linode service model and the input configuration, and then the best matching instance is obtained as follows by searching the Linode service model:

Linode2GB AND LoadBalancer AND Month AND Singapore AND (Network.estimatetraffic <= Freetraffic. Freelimit × VM.num) IMPLIES (Virtualmachine.RAM == 2 GB) AND (Virtualmachine.vCPU == 2) AND (Virtualmachine.HDDcapabiltiy == 48 GB) AND (Freetraffic.freelimit == 3072 G) AND (Virtualmachine.costmonth == $20) AND (LinodeServer.periodcost = Virtualmachine.costmonth × Virtualmachine.num + LoadBalancer.monthcost × LoadBalancer.num);

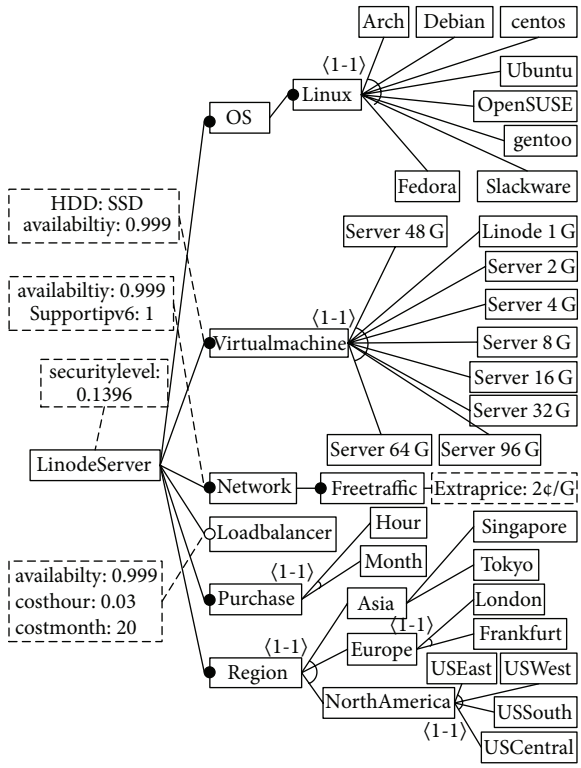



Figure 8: Major structure of the Linode service model.

LinodeServer.periodcost = 20 × 2 + 20 = $60.

We have verified the validity of the DigitalOcean service model, the RackSpace service model, and their input configurations using similar methods and obtained the best matching

```
import es.us.isa.FAMA.Reasoner.QuestionTrader; . . .
import IaaSCustomOperation.ProductSetOpt; //Matching instances are optimized in it
Public class AttributedFMoperation {
public static void main(String[] args) {
QuestionTrader qt = new QuestionTrader ();
GenericAttributedFeatureModel afm = (GenericAttributedFeatureModel)
qt.openFile ("IaaSFM/Linode.afm");
qt.setVariabilityModel (afm);
ValidQuestion vq= (ValidQuestion) qt.createQuestion ("Valid"); qt.ask (vq);
System.out.println ("The validity of the model is" + vq.isValid());
if (vq.isValid())
{ Configuration p = new Configuration();
GeneriAttributedFeature f1 = afm.searchFeatureByName ("VM"); . . .
f1.addAttribute(f1.searchAttributeByName("memory"), "2");       . . .
p.addElement (f1,1);                                                    . . .
ValidConfigurationQuestion vcq= (ValidConfigurationQuestion)
qt.createQuestion ("ValidConfiguration");
vcq.setConfiguration (p);   qt.ask (vcq);
System.out.println ("The validity of the configurationis" + vcq.isValid());
if (vcq.isValid()) { qt.addStagedConfiguration (p);
Questionq = qt.createQuestion ("Products"); qt.ask(q);
ProductsQuestion pq= (ProductQuestion) q;
if (pq.getNumberOfProcucts() == 0) system.out.println ("\n No valid product");
else
   { Iterator<? extends GenericProduct> it = pq.getAllProducts().iterator();
      if (it.hasNext()) { ProductSetOpt po = new ProductSetOpt(it);
        po.findBestProductAndprint (); }
   }
} } } }
```

Figure 9: Source code of verifications of the Linode service model and the input configuration and searching for the best instance.

Table 3: Main nonfunctional metrics of the best matching instances.

| Metric name | RackSpace | DigitalOcean | Linode |
|---|---|---|---|
| $a_1$: cost | 107.44 | 55 | 60 |
| $a_2$: security index | 0.1478 | 0.1307 | 0.1396 |
| $a_3$: SPECCPU2006 | 111 | 101 | 118 |
| $a_4$: availability | 0.9989 | 0.9980 | 0.9998 |
| $a_5$: scaling latency | 123 | 55 | 118 |
| $a_6$: network latency | 109 ms | 398 ms | 260 ms |
| $a_7$: market share | 5% | 1% | 2% |

instances. The costs of the best instances from DigitalOcean and RackSpace are \$55 and \$107.44, respectively.

*7.3. Ranking of the Best Matching Instances.* The steps in ranking the best matching instances are as follows.

*(1) Measurement of the Main Nonfunctional Metrics for the Best Matching Instances*. According to Section 7.1.2, the main nonfunctional metrics are obtained as shown in Table 3.

Accordingly, the following decision matrix $\mathbf{X}$ is created:

$$\mathbf{X} = \begin{bmatrix} 107.44 & 0.1478 & 111 & 0.9989 & 123 & 109 & 0.05 \\ 55 & 0.1307 & 101 & 0.9980 & 55 & 398 & 0.01 \\ 60 & 0.1396 & 118 & 0.9998 & 118 & 260 & 0.02 \end{bmatrix}. \quad (9)$$

In order to facilitate comparison of the above metrics of these instances, we utilize the method given in [45] to standardize all evaluation metrics through the formula "$x_{ij} = (\max x_{ij} - x_{ij}) + \min x_{ij}, i = 1 \rightarrow m$." All evaluation metrics are converted into profit metrics by the formula "$\max x_{ij} - x_{ij}$", and then a coordinate translation is carried out by "$+\min x_{ij}$," so that the values of these metrics can be restored to their original size level. Second, we normalize the converted matrix by the formula $y_{ij} = x_{ij} / \sum_{i=1}^{m} x_{ij}$, $1 \le i \le m$, $1 \le j \le n$, where $m$ is the number of alternative service instances and $n$ is the number of evaluation metrics. Finally, the standardized decision matrix $\mathbf{R}$ is obtained:

$$\mathbf{R} = \begin{bmatrix} 0.208 & 0.353 & 0.336 & 0.333 & 0.231 & 0.528 & 0.625 \\ 0.405 & 0.313 & 0.306 & 0.333 & 0.517 & 0.145 & 0.125 \\ 0.387 & 0.334 & 0.358 & 0.334 & 0.252 & 0.327 & 0.250 \end{bmatrix}. \quad (10)$$

*(2) Determination of the Subjective Weights*. According to Section 6.2.1, we first divide all evaluation metrics into four groups and sort them by ascending importance. The metric sequence $a_5$, ($a_4$, $a_6$), ($a_2$, $a_7$), and ($a_1$, $a_3$) is obtained. The importance of each metric in the sequence is greater than or equal to its left-hand closest metric. Next, the metric weight distribution is carried out, and the results are shown in Table 4.

According to the formula $w_i = 1.5^{\beta_i} / \sum_{j=1}^{7} 1.5^{\beta_j}$, $\sum_{j=1}^{7} 1.5^{\beta_j} = 15.25$, the subjective weight vector $\mathbf{w}_s$ is calculated:

$$\mathbf{w}_s = (0.221, 0.148, 0.221, 0.098, 0.066, 0.098, 0.148)^T. \quad (11)$$

Table 4: Distribution of the metric relative importance degree.

| $\beta$ | 0 | 0.5 | 1 | 1.5 | 2 | 2.5 | 3 | 3.5 | 4 |
|---|---|---|---|---|---|---|---|---|---|
| $n = 1$ | $a_5$ | — | $a_4$ | — | $a_2$ | — | $a_1$ | — | — |
| $n = 2$ | — | — | $a_6$ | — | $a_7$ | — | $a_3$ | — | — |

*(3) Calculation of the Objective Weights*. According to Section 6.2.2, the entropy of an evaluation metric is obtained by formula (4), and the results are as follows:

$$\begin{aligned} e_1 &= 0.96494, \\ e_2 &= 0.9989, \\ e_3 &= 0.99813, \\ e_4 &= 0.99999, \\ e_5 &= 0.93473, \\ e_6 &= 0.89452, \\ e_7 &= 0.81945. \end{aligned} \tag{12}$$

The metric variation coefficients are obtained by the formula $d_j = 1 - e_j$, and the results are as follows:

$$\begin{aligned} d_1 &= 0.03506, \\ d_2 &= 0.0011, \\ d_3 &= 0.00187, \\ d_4 &= 0.00001, \\ d_5 &= 0.06527, \\ d_6 &= 0.10548, \\ d_7 &= 0.18055. \end{aligned} \tag{13}$$

The objective weight vector is obtained as follows by the formula $w_j = d_j / \sum_{j=1}^{n} d_j$:

$$\begin{aligned} &\mathbf{w}_o \\ &= (0.090, 0.003, 0.005, 0.001, 0.168, 0.270, 0.463)^T. \end{aligned} \tag{14}$$

*(4) Combination Weighting*. After the completion of the subjective weighting, metrics are sorted by ascending importance as follows:

$$(a_5), (a_4, a_6), (a_2, a_7), (a_1, a_3). \tag{15}$$

From the results of the subjective weighting, $a_4$ and $a_6$, $a_2$ and $a_7$, and $a_1$ and $a_3$ are equally important; therefore, there are some equivalent results of sorting metrics:

$$\begin{aligned} &(a_5), (a_4, a_6), (a_2, a_7), (a_3, a_1); \\ &(a_5), (a_4, a_6), (a_7, a_2), (a_1, a_3); \\ &(a_5), (a_4, a_6), (a_7, a_2), (a_3, a_1); \\ &(a_5), (a_6, a_4), (a_2, a_7), (a_1, a_3); \\ &(a_5), (a_6, a_4), (a_2, a_7), (a_3, a_1); \\ &(a_5), (a_6, a_4), (a_7, a_2), (a_1, a_3); \\ &(a_5), (a_6, a_4), (a_7, a_2), (a_3, a_1). \end{aligned} \tag{16}$$

After completion of the objective weighting, metrics are sorted by ascending importance as follows.

$$a_4, a_2, a_3, a_1, a_5, a_6, a_7. \tag{17}$$

The results show that the sorting results of metrics obtained by the two weighting methods are not exactly the same; the sorting results of metric groups are also not the same. Therefore, we take the linear combination coefficient $a = 1$. According to formula (7), the combination weight vector is calculated as $\mathbf{w} = \mathbf{w}_s = (0.221, 0.148, 0.221, 0.098, 0.066, 0.098, 0.148)^T$. The result can be explained as follows: The metric weight sorting obtained by the objective weighting methods deviates markedly from the subjective preferences of the user. Therefore, the result obtained by the objective weighting method has no reference value. Based on the main nonfunctional metrics and their weights, we have obtained the comprehensive evaluations of the three best instances from the RackSpace, DigitalOcean, and Linode service models as follows:

$$\begin{aligned} U_{\text{Best-instance-from-RackSpace}} &= 0.364592, \\ U_{\text{Best-instance-from-DigitalOcean}} &= 0.302921, \\ U_{\text{Best-instance-from-Linode}} &= 0.332487. \end{aligned} \tag{18}$$

Thus, the best matching instances from the three providers are sorted as follows:

$$\begin{aligned} &U_{\text{Best-instance-from-RackSpace}} > U_{\text{Best-instance-from-Linode}} \\ &\quad > U_{\text{Best-instance-from-DigitalOcean}}. \end{aligned} \tag{19}$$

*7.4. Results Analysis.* First, the IaaS discovery process implements the matching of supply and demand on IaaSs with a configuration instance granularity, in which the search for the matching instances covers the complete configuration space for each IaaS. Therefore, the QoS metrics of all matching instances can be obtained accurately so that the QoS of these instances can be objectively evaluated and compared, which ensures that the decision results are accurate. Second, the IaaS discovery and selection process realized in this paper is systematic and automatic. After

the user submits the demand of an IaaS and the subjective weight vector of the decision metrics, the entire process will be completed automatically. Thus the efficiency of IaaS discovery and selection is greatly improved. Finally, the performance of our methods is good. The repository of IaaS service models is updated and maintained periodically by the system. Thus, the main processing part of our methods lies in the matching of supply and demand on IaaS FMs. This processing part takes out the IaaS service model from the repository sequentially and then searches the best matching instances from the IaaS configuration space represented by the service mode. Assuming there are $m$ IaaS services, each provides about $n$ configuration instances, and the COMPLEX CONSTRAINTS representation of each instance involves about $l$ conditional options; the time complexity is then $O(n \times m \times l)$. The experimental result shows that the searches for the best matching instances from the three IaaS service models are accomplished within 1 s.

## 8. Conclusions and Future Work

This paper systematically studied the description, discovery, and selection methods of IaaSs in cloud computing. First, we designed a standard description method for IaaSs that uses a fine-grained and structured way to describe the entire configuration space of an IaaS. The method overcomes the shortcomings of existing methods, which can only describe the macro information of IaaSs and provides the conditions for the automated discovery and selection of IaaSs. Second, we proposed a systematic discovery and selection method for IaaSs, which introduces the automated analysis techniques of FMs to analyze the service model and search for the best matching instance from each IaaS. The discovery efficiency is greatly improved. Moreover, due to the fine-grained matching of supply and demand on IaaSs, the matching results directly point to the specific configuration instances, so the accuracy of service discovery is effectively improved. Finally, we unified the decision metrics for service selection, and presented effective and feasible metric measurement methods. We also presented a combination weighting method for decision metrics based on the subjective and objective weighting methods. As such, the ranking result of alternative IaaS instances is more objective and reliable.

In reality, a combination of multiple IaaSs can sometimes meet the needs of some applications better, such as global e-commerce. If a combination of multiple IaaSs from different regions is used, the user requests will be effectively dispersed so that the system pressure can be greatly reduced and user experience can be improved. In addition, some applications have to be deployed on the combination of multiple IaaSs, such as a medical-related and big-data-analysis application, which needs to collect the original data from multiple countries for analysis and processing. If the laws of the relevant countries strictly prevent such original data flowing out of their territories, then users will have to buy a combination of multiple IaaSs located in the licensed areas. Therefore, there is a strong realistic significance in studying how to discover and select the best IaaS composition based on user needs.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This work was supported by the Nature Science Foundation of Jiangsu, China, under Grant no. BK20131069.

## References

[1] M. Armbrust, A. Fox, R. Griffith et al., "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[2] Cloudharmony, 2015, https://cloudharmony.com.

[3] A. V. Dastjerdi, S. G. H. Tabatabaei, and R. Buyya, "An effective architecture for automated appliance management system applying ontology-based cloud discovery," in *Proceedings of the 10th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid '10)*, pp. 104–112, May 2010.

[4] M. Á. Rodríguez-García, R. Valencia-García, F. García-Sánchez, and J. J. Samper-Zapater, "Ontology-based annotation and retrieval of services in the cloud," *Knowledge-Based Systems*, vol. 56, no. 3, pp. 15–25, 2014.

[5] J. Kang and K. M. Sim, "Ontology and search engine for cloud computing system," in *Proceedings of the International Conference on System Science and Engineering (ICSSE '11)*, pp. 276–281, Macau, China, June 2011.

[6] S. K. Garg, S. Versteeg, and R. Buyya, "A framework for ranking of cloud computing services," *Future Generation Computer Systems*, vol. 29, no. 4, pp. 1012–1023, 2013.

[7] Z. Li, L. O'Brien, H. Zhang, and R. Cai, "On a catalogue of metrics for evaluating commercial cloud services," in *Proceedings of the 13th ACM/IEEE International Conference on Grid Computing (GRID '12)*, pp. 164–173, Beijing, China, November 2012.

[8] E. Wittern, J. Kuhlenkamp, and M. Menzel, "Cloud service selection based on variability modeling," in *Service-Oriented Computing*, pp. 127–141, Springer, Berlin, Germany, 2012.

[9] J. García-Galán, P. Trinidad, O. F. Rana, and A. Ruiz-Cortés, "Automated configuration support for infrastructure migration to the cloud," *Future Generation Computer Systems*, vol. 55, pp. 200–212, 2016.

[10] J. García-Galán, O. F. Rana, P. Trinidad, and A. Ruiz-Cortés, "Migrating to the cloud: a software product line based analysis," in *Proceedings of the 3rd International Conference on Cloud Computing and Services Science (CLOSER '13)*, pp. 416–426, Aachen, Germany, May 2013.

[11] Amazon, 2015, http://aws.amazon.com.

[12] K. Kang, S. Cohen, J. Hess et al., "Feature-oriented domain analysis feasibility study," Tech. Rep. CMU/SEI-90-TR-21, Software Engineering Institute, Carnegie Mellon University, 1990.

[13] G. Shen, W. Zhang, Z. Huang et al., "Description-logic-based feature modeling and verification," *Journal of Computer Research and Development*, vol. 50, no. 7, pp. 1501–1512, 2013.

[14] C. Quinton, D. Romero, and L. Duchien, "Automated selection and configuration of cloud environments using software product lines principles," in *Proceedings of the 7th IEEE International Conference on Cloud Computing (CLOUD '14)*, pp. 144–151, Anchorage, Alaska, USA, June 2014.

[15] J. Zhou, D. Zhao, and J. Liu, "TEFL: a textual feature modeling language," *Journal of Chinese Computer Systems*, vol. 33, no. 10, pp. 2133–2140, 2012.

[16] Amazon, "AWS Total Cost of Ownership (TCO) Calculator," 2015, https://awstcocalculator.com.

[17] Rackspace, 2015, http://www.rackspace.co.uk/solutions-configurator.

[18] Cloudscreener, 2015, http://www.cloudscreener.com.

[19] Cloudorado, 2015, https://www.cloudorado.com.

[20] Planforcloud, 2015, http://www.planforcloud.com.

[21] S. Frey, F. Fittkau, and W. Hasselbring, "Search-based genetic optimization for deployment and reconfiguration of software in the cloud," in *Proceedings of the 35th International Conference on Software Engineering (ICSE '13)*, pp. 512–521, San Francisco, Calif, USA, May 2013.

[22] M. Zhang, R. Ranjan, A. Haller et al., "An ontology-based system for cloud infrastructure services' discovery," in *Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pp. 524–530, Wuhan, China, November 2012.

[23] T. Han and K. M. Sim, "An ontology-enhanced cloud service discovery system," in *Proceedings of the International Multi Conference of Engineers and Computer Scientists (IMECS '10)*, vol. 2180, HongKong, March 2010.

[24] L. Liu, X. Yao, L. Qin, and M. Zhang, "Ontology-based service matching in cloud computing," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '14)*, pp. 2544–2550, IEEE, Beijing, China, July 2014.

[25] Cloud Service Measurement Index Consortium (CSMIC), "SMI framework," 2015, http://www.csmic.org/.

[26] S. K. Garg, S. Versteeg, and R. Buyya, "SMICloud: a framework for comparing and ranking cloud services," in *Proceedings of the IEEE International Conference on Utility and Cloud Computing*, pp. 210–218, Melbourne, Australia, December 2011.

[27] B. Dougherty, J. White, and D. C. Schmidt, "Model-driven auto-scaling of green cloud computing infrastructure," *Future Generation Computer Systems*, vol. 28, no. 2, pp. 371–378, 2012.

[28] E. Wittern and C. Zirpins, "Service feature modeling: modeling and participatory ranking of service design alternatives," *Software & Systems Modeling*, vol. 15, no. 2, pp. 553–578, 2016.

[29] DigitalOcean, 2015, https://www.digitalocean.com.

[30] CSA, "CSA Security, Trust and Assurance Registry (STAR)," 2015, https://cloudsecurityalliance.org/star/.

[31] D. Benavides, S. Segura, and A. Ruiz-Cortés, "Automated analysis of feature models 20 years later: a literature review," *Information Systems*, vol. 35, no. 6, pp. 615–636, 2010.

[32] A. Li, X. Yang, S. Kandula, and M. Zhang, "CloudCmp: comparing public cloud providers," in *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement (IMC '10)*, pp. 1–14, ACM, Melbourne, Australia, November 2010.

[33] S. Islam, K. Lee, A. Fekete, and A. Liu, "How a consumer can measure elasticity for cloud platforms," in *Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering (ICPE '12)*, pp. 85–96, Boston, Mass, USA, April 2012.

[34] Spec, 2015, http://www.spec.org.

[35] H. Zhu, L. Wu, and H. Kang, "Research of cloud provider selection method based on SecLA," *Computer Science*, vol. 43, no. 5, pp. 100–108, 2016.

[36] M. Wang, "A comprehensive analysis method for determinating the weight coefficients in comprehensive evaluation of multiple indexes," *Systems Engineering*, vol. 17, no. 2, pp. 56–61, 1999.

[37] J.-X. Liu, Y.-J. Tan, and H.-P. Cai, "Study of the methods of the linear combination weighting for multiple attribute decision-making," *Journal of National University of Defense Technology*, vol. 27, no. 4, pp. 121–124, 2005.

[38] H. Chen, "Combination determining weights method for multiple attribute decision making based on maximizing deviations," *Systems Engineering and Electronics*, vol. 26, no. 2, pp. 194–197, 2004.

[39] D. Song, C. Liu, C. Shen et al., "Multiple objective and attribute decision making based on the subjective and objective weighting," *Journal of Shandong University*, vol. 45, no. 4, pp. 1–9, 2015.

[40] Y. Li and J. Wang, "Method for deriving AHP weight based on sorting," *Ordnance Industry Automation*, no. 11, pp. 42–44, 2013.

[41] W. Chen and J. Xia, "Optimal combined weighting method based on the subjective and objective weights," *Mathematics in Practice and Theory*, vol. 37, no. 1, pp. 17–22, 2007.

[42] RackSpace, 2015, http://www.rackspace.com.

[43] Linode, 2015, https://www.linode.com.

[44] ISA Research Group, "FAMA tool suite," 2015, http://www.isa.us.es/fama/.

[45] M. Li, G. Chen, and Y. Chen, "Research on the method of index standardization in comprehensive evaluation," *Chinese Journal of Management Science*, no. 12, pp. 45–48, 2004.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 5412028, 9 pages
http://dx.doi.org/10.1155/2016/5412028

*Research Article*

# Incentive Mechanism Design Aiming at Deflated Performance Manipulation in Retail Firms: Based on the Ratchet Effect and the Reputation Effect

**Biao Luo, Chengyuan Wang, and ChunYi Li**

*The School of Management, University of Science and Technology of China, Hefei, Anhui 230026, China*

Correspondence should be addressed to Chengyuan Wang; nanjing@mail.ustc.edu.cn

Received 2 March 2016; Accepted 23 August 2016

Academic Editor: Roman Wendner

Copyright © 2016 Biao Luo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Store managers in retail firms are often offered a performance-based compensation scheme accompanied with a performance target by the headquarters. The headquarters adjusts the performance target based on store managers' historical performance and therefore generates the ratchet effect. Consequently, store managers may downward manipulate performance, that is, deflated performance manipulation, so as to weasel out of target growth and smooth performance growth. However, the reputation effect that seeks fame by store managers can restrain deflated performance manipulation. We model a dynamic agency setting in which both the ratchet effect and the reputation effect are related to the store manager's compensation scheme, and the store manager has to balance her effort and deflated performance manipulation. Our findings reveal that the ratchet effect and environmental volatility jointly determine the existence of deflated performance manipulation, yet the reputation effect can restrain it with increasing environmental volatility. In addition, deflated performance manipulation is inevitable when environmental volatility is large enough, and explicit incentives may promote deflated performance manipulation.

## 1. Introduction

In retail firms, the headquarters has to motivate and supervise multiple stores. In order to enhance the performance per store, the headquarters always offers store managers a performance-based compensation scheme [1]. The compensation scheme is generally comprised of a fixed component and a revenue sharing and therefore binds the store managers' compensation to their performance [2]. In addition, the compensation scheme is often followed by a performance target which acted as a criterion to evaluate the store manager's performance. For the headquarters, such scheme seems to be an ideal arrangement to motivate store managers to make every effort. In fact, however, the scheme often generates side-effect, that is, the ratchet effect. Due to asymmetric information and lacking market insights relative to the store manager, the headquarters often sets the target on the basis of the store manger's historical performance. The high historical performance will lead to the high current target, and the circulation scrolls up by such analogy in the future fiscal years. The ratchet effect has been verified in practice by a series of empirical evidence, for example, [3–6].

Store managers can be naturally conscious of the ratchet effect, and they generally take two measures, that is, intentional reducing effort or lowering the reported performance, to evade the pressure produced by the growing performance target. Regarding managers' such adverse selections, prior agency studies mainly focused on the former, that is, investigating how to avoid effort reduction by managers (e.g., [7, 8]). Conversely, shrewd store managers often adopt the latter in practice. It is because lowering reported performance not only can lead the headquarters to set a moderate target in the next period, but also implies a portion of surplus performance that can be counted in the next period, so as to ensure a smooth performance growth. We defined such manipulative behavior that managers virtually downward reduce performance as *deflated performance manipulation*. It leaves some regulatory leeway for store managers to adjust performance in the next period, in order to maximize their interperiod incomes. Given that deflated performance manipulation is

widely common in practice, it is very necessary to investigate how store managers manipulate performance to oppose the ratchet effect.

In addition, store managers will also restrain themselves out of seeking reputation, that is, reputation effect. It is because deflated performance manipulation may not violate the law but is always accompanied by moral hazard. In addition, social reputation can be conducive to enhance store managers' bargaining power in compensation negotiation. Yet once the manipulation is ferreted out, the store manager might fall into disrepute in the eyes of the headquarters, as well as in the labor market. Fama [9] pointed out that, in a long-term incentive setting, social reputation can be an implicit incentive to restrain managers' misconduct.

Given that the ratchet effect and the reputation effect have opposite impacts on store managers, in this paper, we focus on studying how the two effects jointly influence the performance manipulation of the store manager, who is given a performance-based compensation scheme, and the optimal incentive. Due to both the ratchet effect and the reputation effect referring to long-term incentive problem, we model a dynamic agency setting to investigate how the store manager balances her effort involvement and deflated performance manipulation, so as to help the headquarters improve the incentive scheme and supervision.

This paper contributes to the agency literature in twofold. First, we bridge the research gap between the ratchet effect and deflated performance manipulation that was ignored by prior studies. We embody the relationship between the ratchet effect and deflated performance manipulation in a dynamic agency setting. Second, we jointly consider the different impacts of two effects, that is, ratchet effect and reputation effect, on the store manager's behaviors and the optimal incentive. The jointly incorporating such two effects can provide a better dialogue with prior studies.

The remainder of the paper is organized as follows. Section 2 provides a review of the literature. We construct the analysis model in Section 3. We give the cross-sectional results of the model in Section 4. Implications of our results and concluding remarks are given in Section 5.

## 2. Literature Review

To align store managers' salary with their performance, the headquarters of retail firms always offers a performance-based compensation scheme to store managers [1]. However, an unanticipated result is that, in more cases, store managers tend to manipulate the reported performance rather than make more efforts. Abundant evidence has demonstrated store managers' performance manipulation. For instance, Healy [10] found that if the bonus reward based on performance is an important part of managers' compensation, they will be motivated to handle actual profits to increase their compensation. Further researches [11–13] had also verified the relationship between the bonus plan and performance manipulation and found that store managers often manipulate performance so as to transfer wealth from the principle to them. These studies focused on a static agency setting in which managers tend to upward manipulate reported performance. In a dynamic agency setting, however, performance manipulation of store managers may be more complicated.

Recently, the incentive scheme involved with the performance manipulation has been extensively studied. Crocker and Morgan [14] investigated misreporting in a principal-agent setting. Goldman and Slezak [15] developed an agency model which shows that a stock-based compensation can not only induce managers to make productive effort but also divert valuable resources to misrepresent performance. Burns and Kedia [16] tested the effect of compensation scheme on managers' misreporting and found that stock options are associated with stronger incentives to misreport. Crocker and Slemrod [17] characterized the optimal compensation scheme in the setting where managers may take actions to hide actual performance information and demonstrated that a compensation scheme contingent on reported earnings neither promotes managers to maximize profits nor leads managers to report actual profits honestly.

Among various forces that impact store manager's performance manipulation in a dynamic agency setting, there are two important effects that can oppositely influence store managers' tendency to manipulate performance. One is the ratchet effect, which is widespread in retail firms and unexpectedly induces store managers to downward manipulate reported performance, that is, deflated performance manipulation as we defined. The ratchet effect derives from the growing performance target accompany with performance-based compensation scheme. The headquarters always sets current target based on the store manager's historical performance, and therefore the higher historical performance leads to the higher target and vice versa [5, 7, 18]. Consequently, store managers are inclined to reduce current performance to lessen the pressure of a higher target in the next period. Regarding reducing current performance, calculative managers often balance effort reduction and deflated performance manipulation. It is because the latter can help store managers use the unreported surplus performance to smooth performance in the next period, for example, earnings management [1, 19–21], so as to guarantee seemly performance growth. Goal-setting theory provides evidence to explain the reason that managers tend to take risk to attain a high target [22]. Leone and Rock [23] suggested that managers are promoted to manipulate reported earnings downward by accruals. Other scholars also suggested that managers will reduce their actual performance by various ways, for example, allowing for price discounts [24] and reducing investments [25, 26], because deflated performance manipulation is concealed and is difficult to be perceived [8].

The other one is the reputation effect, which can restrain store managers' manipulation. The reputation can enhance store managers' bargaining power in compensation negotiation with the headquarters as well as in external executive labor market and therefore will promote store managers to make more effort. The rudiment of the reputation effect derives from Fama's [9] work which proposed that the labor market appropriately uses current and past information to revise future wages. Holmstrom [27] modeled the proposition of Fama [9] and suggested that incentive problems,

at least in part, can be eased through the role of long-term reputation in a dynamic environment. Meyer and Vickers [28] analyzed the comparative performance model with introducing historical performance and showed that the role of the reputation effect is opposite with the ratchet effect. In order to maximize the total utility, the headquarters must be able to predict what the store manager will do [29] when the reputation is an effective information for prediction [30]. Accordingly, store managers can legitimately require a better compensation by the contract negotiation. Thus, the reputation effect is considered in our model and plays a role by adjusting the fixed wage in the compensation contract.

We jointly consider the ratchet effect and the reputation effect in this paper, in order to investigate how store managers under the two different effects will balance this effort involvement and the extent of performance manipulation, and help the headquarters design the optimal incentive on the basis of the participation constraint.

## 3. Model Development

We consider the following agency setting. In a retail firm, the headquarters hires a manager to operate a branch store over two periods when the first period is the start-up and the second period is the final one. The headquarters offers a performance-based compensation scheme accompany with a sales target for the manager [31–34]. The compensation scheme is a linear contract which has an explicit formula showing clearly the tradeoff between risk-sharing and incentives [35]. Holmstrom and Milgrom [36] considered the optimality of the linear contract in a natural class of dynamic moral hazard problems. Thus, we draw the following linear schemes.

*Assumption 1* (a performance-based compensation scheme). The output of period $i$ is $\pi_i$ and the corresponding performance target is $\underline{\pi}_i$. The manager's incomes derive merely from her compensation which is linear decided by her output:

$$s_i(\pi_i) = \alpha_i + \beta(\pi_i - \underline{\pi}_i) \quad (i = 1, 2), \tag{1}$$

where $\alpha_i$ and $\beta$ are parameters of the linear contract given by the headquarters. Therein, $\alpha_i$ is the fixed wage in the period $i$. As previously mentioned, the fixed wage in the second period may be changed following the manager's reputation. However, the pay-for-performance sensitivity $\beta$ is changeless and $0 \le \beta \le 1$. Thus, $\beta(\pi_i - \underline{\pi}_i)$ is the incentive bonus related to surplus performance, which is the actual performance minus performance target.

Following prior studies [27, 37, 38], we consider that the store manager's output is linear decided by her effort and talent. Since the profitability of the store is an important factor to the output, we take it into account in our model and therefore draw the following assumption.

*Assumption 2.* The actual output is given by

$$\pi_i = \lambda a_i + \eta + \varepsilon_i, \tag{2}$$

where $\lambda > 0$ is the profitability of the store, $\eta \sim N(0, \tau\sigma^2)$ is the manager's talent, and $\varepsilon_i \sim N(0, (1-\tau)\sigma^2)$ is the random variables due to exogenous environmental volatility. Though the store manager chooses the effort $a_i$, the headquarters who can just observe the report output $\widetilde{\pi}_i$ does not know the real level of $\eta$ and $\varepsilon_i$. We assume that $\mathrm{cov}(\varepsilon_1, \varepsilon_2) = 0$ and

$$\tau = \frac{\mathrm{var}(\eta)}{\mathrm{var}(\eta) + \mathrm{var}(\varepsilon_i)}, \tag{3}$$

and thus we have $\mathrm{var}(\pi_i) = \sigma^2$.

There is a cost $C(a_i)$ to exert an effort $a_i$. The effort cost is concave and therefore $C'(a_i) > 0$ and $C''(a_i) > 0$ [33, 39]. For more analytical results as prior studies did (e.g., [40, 41]), we assume the following.

*Assumption 3.* The store manager's effort cost is

$$C(a_i) = \frac{b}{2}a_i^2, \tag{4}$$

where $b > 0$ is the effort cost coefficient.

The general problem faced by the headquarters is information asymmetry, because the store manager can influence the reported performance output.

*Assumption 4.* If the manager manipulates the reported performance, which will be given by

$$\widetilde{\pi}_i = \pi_i - \Delta\pi_i = \lambda a_i + \eta + \varepsilon_i - \Delta\pi_i, \tag{5}$$

where $\Delta\pi_i$ denotes the extent of deflated performance manipulation, according to (5), in each period, the store manager has to balance her effort $a_i$ and deflated performance manipulation $\Delta\pi_i$ in order to maximize her utility across the interperiod. Hence, the expected value of $\widetilde{\pi}_i$ is given by $E(\widetilde{\pi}_i) = \lambda a_i - \Delta\pi_i$.

In retail firms, the headquarters always uses each store's annual deviation from its current sales target to set next year's sales target. The store manager has more motivation to manipulate to reduce the current performance in order to lessen the pressure of the next year higher target. We assume that the manager only chooses the deflated performance manipulation, and thus

$$\Delta\pi_i = \pi_i - \widetilde{\pi}_i \ge 0. \tag{6}$$

This part of output $\Delta\pi_i$ can be achieved later in next year by accounting means.

Considering the manager may manipulate the reported performance, the headquarters may supervise the manager. And thus, we draw the following.

*Assumption 5.* The headquarters supervises the reported performance of stores and learns whether the store manager manipulates with some probability $\gamma$ ($0 < \gamma < 1$) which indicates the effectiveness of the supervision. To bind the manager's behaviors, we consider the punishment setting

following the effort cost $C(a_i)$. The punishment $C_p$ is also concave (i.e., $C_p' > 0$ and $C_p'' > 0$). For more analytic results, the punishment for manipulation is given by

$$C_p = \frac{k}{2}\Delta\pi_i^2, \tag{7}$$

where $k > 0$ is the punishment coefficient. Thus, the expected punishment of the manipulation is $\gamma C_p$.

Hart and Holmstrom [37] argued that the manager should be risk-neutral in a dynamic setting. Accordingly, we put forward the following.

*Assumption 6.* The headquarters and the store manager are both risk-neutral.

Because the external environment and the investment of headquarters to store have no significant change, the performance target is generally based on historical performance. On the other side, the ratchet effect leads to higher target in next period if current performance is higher than the historical level and vice versa. Because the headquarters hires the manager in the first period when the talent of the manager is unknown, let $\underline{\pi}_1 = 0$ denote the performance target of the first period, and let the difference between real output and target (i.e., $\pi_1 - \underline{\pi}_1$) denote the deviation of the expected performance. Based on adaptive expectation theory, the target in next period will be adjusted according to the deviation. We assume that the target of the second period is given by

$$\underline{\pi}_2 = \underline{\pi}_1 + h\left(\pi_1 - \underline{\pi}_1\right) = h\pi_1, \tag{8}$$

where $h > 0$ is the target adaptation coefficient which reflects the adjustment of the next target according to the ratchet effect. However, the store manager can manipulate performance to achieve private interests by the role of $a_1$ and $\Delta\pi_1$ which can change $\widetilde{\pi}_1$.

Based on the reputation effect, the store manager' effort and manipulation not only affect the current compensation but also affect the future reputation. According to Assumption 2, though the manager's talent is private information, but it can be evaluated by the headquarters based on the past performance. The headquarters will evaluate the store manger's talent by the observation of historical performance while the store manager can affect such expectation by her effort $a_1$ and manipulation $\Delta\pi_1$. Suppose $\widehat{a}_1 = E(a_1)$ is the headquarters expectation about the store manager's effort. Having observed the output $\pi_1$, the headquarters will learn $\eta + \varepsilon_1 = \pi_1 - \lambda\widehat{a}_1$ while it cannot separate $\eta$ from $\varepsilon_1$. The headquarters has to infer $\eta$ according to $\pi_1$. On the basis of rational expectation theory, for a given $\pi_1$, the expected ability of the store manager $E(\eta \mid \pi_1)$ is equal to the weighted average of the priori value $E(\eta)$ and the a posteriori observed value $(\pi_1 - \lambda\widehat{a}_1)$. Correspondingly, the expected ability of the store manager is given by

$$\begin{aligned} E\left(\eta \mid \pi_1\right) &= (1-\tau)E\left(\eta\right) + \tau\left(\pi_1 - \lambda\widehat{a}_1\right) \\ &= \tau\left(\pi_1 - \lambda\widehat{a}_1\right). \end{aligned} \tag{9}$$

And if the headquarters can only observe the reported output,

$$E\left(\eta \mid \widetilde{\pi}_1\right) = \tau\left(\widetilde{\pi}_1 - \lambda\widehat{a}_1\right), \tag{10}$$

where $\widehat{a}_1$ denotes the headquarters' expected value on the store manger's effort. Based on the supervision and the manipulation, the expected ability of the store manager is given by

$$\begin{aligned} &\gamma E\left(\eta \mid \pi_1\right) + (1-\gamma)E\left(\eta \mid \widetilde{\pi}_1\right) \\ &\quad = \tau\left[\lambda a_1 + \eta + \varepsilon_1 - (1-\gamma)\Delta\pi_1 - \lambda\widehat{a}_1\right]. \end{aligned} \tag{11}$$

The reputation effect plays a role by adjusting the fixed wage in the compensation scheme; that is,

$$\alpha_2 = \alpha_1 + c\left[\gamma E\left(\eta \mid \pi_1\right) + (1-\gamma)E\left(\eta \mid \widetilde{\pi}_1\right)\right], \tag{12}$$

where $c \geq 0$ is the bargaining power coefficient of the store manager. The higher the bargaining power coefficient is, the more important the role that the reputation mechanism plays. The reputation effect and the ratchet effect are both reflected in the following formula:

$$\begin{aligned} s_2\left(\pi_2\right) &= \alpha_2 + \beta\left(\pi_2 - \underline{\pi}_2\right) \\ &= \alpha_1 + c\left[\gamma E\left(\eta \mid \pi_1\right) + (1-\gamma)E\left(\eta \mid \widetilde{\pi}_1\right)\right] \\ &\quad + \beta\left[\pi_2 - h\left(\gamma\pi_1 + (1-\gamma)\widetilde{\pi}_1\right)\right]. \end{aligned} \tag{13}$$

The second period is the final period, and hence the output of the second period will not affect the future compensation of the store manager. In the second period, the manager will maximize the current compensation. On the contrary, the managerial effort and the manipulation in the first period can affect the current and future compensation. Thus, in the first period, the manager has to consider the total revenue over two periods. The game order between the headquarters (the principle) and the store manager (the agent) is (1) the headquarters first determines the contract of the first period (i.e., the fixed wage $\alpha_1$ and the pay-for-performance sensitivity $\beta$); (2) the store manager makes an effort $a_1$ based on the contract and produces the real output $\pi_1$ which is her privacy information. However, the headquarters can discover the real output with some possibility $\gamma$ or only get the reported output $\widetilde{\pi}_1$; (3) the headquarters adjusts the contract of the second period (i.e., the fixed wage $\alpha_2$ and the target $\underline{\pi}_2$) based on historical performance of the first period; and (4) the store manager chooses her effort $a_2$ and the reported output $\widetilde{\pi}_2$. The timing of the principal-agent game is shown in Figure 1.

Utility functions of the headquarters (i.e., $U_{H1}$ and $U_{H2}$) and the store managers (i.e., $U_{M1}$ and $U_{M2}$) in each period are given by

$$U_{H1} = E\left(\pi_1 - \gamma s_1\left(\pi_1\right) - (1-\gamma)s_1\left(\widetilde{\pi}_1\right) + \frac{\gamma k}{2}\Delta\pi_1^2\right),$$

$$U_{M1} = E\left(\gamma s_1\left(\pi_1\right) + (1-\gamma)s_1\left(\widetilde{\pi}_1\right) - \frac{\gamma k}{2}\Delta\pi_1^2 - \frac{b}{2}a_1^2\right),$$

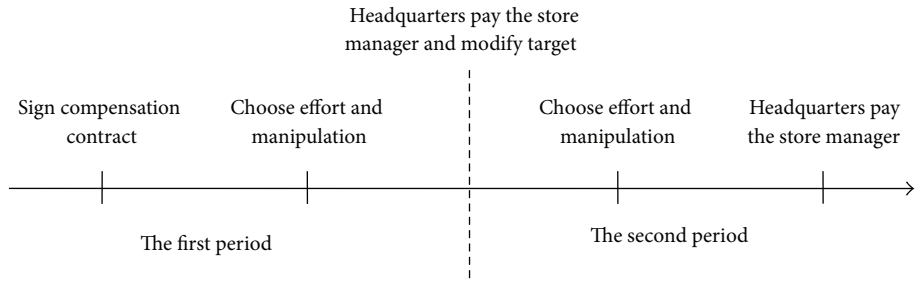



Figure 1: The dynamic (two-period) principal-agent problem.

$$U_{H2} = E\left(\pi_2 - \gamma s_2(\pi_2) - (1-\gamma)\, s_2(\tilde{\pi}_2) + \frac{\gamma k}{2}\Delta\pi_2^2 - (1-\gamma)\,\beta\Delta\pi_1\right),$$
$$U_{M2} = E\left(\gamma s_2(\pi_2) + (1-\gamma)\, s_2(\tilde{\pi}_2) - \frac{\gamma k}{2}\Delta\pi_2^2 - \frac{b}{2}a_2^2 + (1-\gamma)\,\beta\Delta\pi_1\right). \tag{14}$$

The modified agency model with manipulation is

$$\begin{aligned} \max_{\beta} \quad & U_{H1} + \delta U_{H2} \\ \text{S.t.} \quad & \text{IR } U_{M1} + \delta U_{M2} \geq \underline{U} \\ & \text{IC}_1 \ a_2^* = \arg\max U_{A2} \\ & \text{IC}_2 \ \Delta\pi_2^* = \arg\max U_{A2} \\ & \text{IC}_3 \ a_1^* = \arg\max U_{A1} + \delta U_{A2} \\ & \text{IC}_4 \ \Delta\pi_1^* = \arg\max U_{A1} + \delta U_{A2}, \end{aligned} \tag{15}$$

where IR is the participation constraint of the store manager, in which the manager's expected utility should be larger than her reservation utility $\underline{U}$. The incentive compatibility constraints (i.e., $\text{IC}_1$ and $\text{IC}_2$) show that the store manager will choose $a_2$ and $\Delta\pi_2$ to maximize her expected utility of the second period. From the incentive compatibility constraints (i.e., $\text{IC}_3$ and $\text{IC}_4$), the store manager chooses $a_1$ and $\Delta\pi_1$ to maximize her total expected utility over two periods. We also consider the time discount aims at intertemporal gains, and therefore $\delta < 1$ denotes the discount factor.

Different from traditional agency models [33], we first add incentive compatibility constraints of the performance manipulation into our model. Second, we assume that the headquarters goes after both increasing effort $a_i$ and reducing manipulation $\Delta\pi_i$. Finally, considering the long-term implicit incentive effect, the store manager will weigh against the current income and long-term gains.

From the above, the manager's optimal effort of the second period according to $\text{IC}_1$ is given by

$$a_2^* = \frac{\beta\lambda}{b}. \tag{16}$$

This result is consistent with traditional agency models [33], because the store manager does not need to consider the long-term implicit incentive in the last period. Based on $\text{IC}_2$, the manager's optimal manipulation in the second period is given by

$$\Delta\pi_2^* = 0. \tag{17}$$

In the second period, deflated performance manipulation cannot increase private interests through the ratchet effect, such that the manipulation is inadvisable in this period. According to $\text{IC}_3$ and $\text{IC}_4$, the optimal effort and the optimal manipulation in the first period are given by

$$a_1^* = \frac{\lambda\left[(1-h\delta)\,\beta + c\tau\delta\right]}{b}, \tag{18}$$
$$\Delta\pi_1^* = \frac{(1-\gamma)\left[(h\delta+\delta-1)\,\beta - c\tau\delta\right]}{\gamma k}. \tag{19}$$

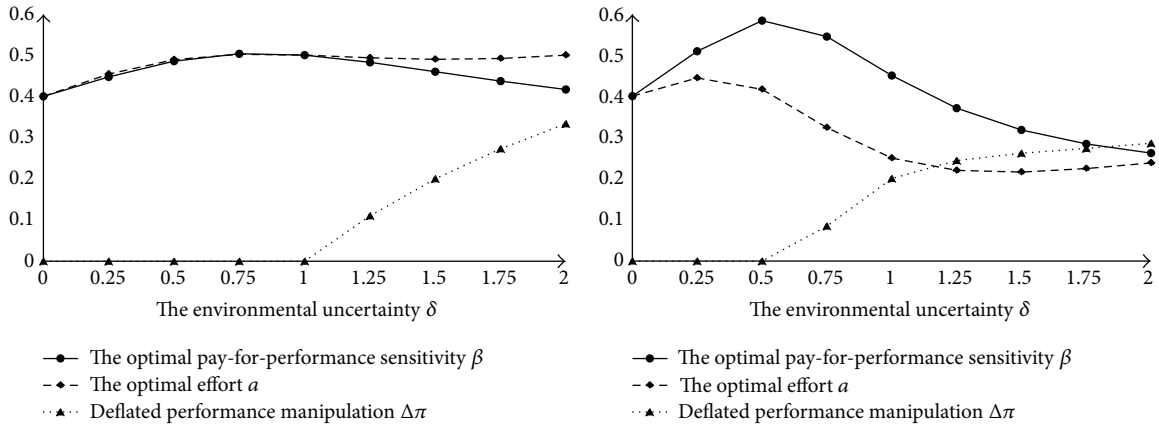
The optimal effort in the first period depends on not only the pay-for-performance sensitivity, but also the target adaptation coefficient $h$ and bargaining power coefficient $c$. In (18) and (19), $c\tau\delta$ implies the reputation effect which promotes the store manager to work hard, while $\beta h\delta$ implies the ratchet effect which leads the store manager to reduce effort.

The equilibrium of the pay-for-performance sensitivity is set so as to balance the net benefit from inducing effort and the cost of inducing manipulation. With the optimal effort and the optimal manipulation above, the optimal pay-for-performance sensitivity is given by

$$\beta^* = \frac{\lambda^2\gamma k\left[\delta + (1-h\delta)(1-2c\tau\delta)\right] + bc\tau\delta(1-\gamma)^2(\delta h+\delta-1)}{2\lambda^2\gamma k\left[(1-h\delta)^2+\delta\right] + b(1-\gamma)^2(\delta h+\delta-1)^2}. \tag{20}$$

## 4. Cross-Sectional Results

This section develops a set of cross-sectional implications that predict how the manager's manipulation and effort vary with the pay-for-performance sensitivity and observable firm characteristics which are replaced by parameters in the model. Proofs are provided in Appendix.




(a) The case that the target adaptation coefficient is equal to 0.5 ($h = 0.5$)

(b) The case that the target adaptation coefficient is equal to 1 ($h = 1$)

FIGURE 2: Numerical simulations in different levels of the target adaptation coefficient and the environmental volatility. (Note that other parameters are given by $\lambda = 1$, $\gamma = 0.5$, $\tau = 0.5$, $k = 1$, $b = 1$, and $c = 0.5$.)

### 4.1. Optimal Manipulation Implications

**Proposition 7.** *There exists a nonempty set of parameters* $\delta h + \delta - 1 < 0$ *under which the store manager will not manipulate performance no matter what the pay-for-performance sensitivity* $\beta$ *is. On the contrary, if* $\delta h + \delta - 1 > 0$ *and the pay-for-performance sensitivity satisfies*

$$\beta > \frac{c\tau\delta}{h\delta + \delta - 1}, \tag{21}$$

*the extent of manipulation* $\Delta\pi_1^* > 0$ *which is showed in (19).*

Proposition 7 suggests that the existence of deflated performance manipulation is contingent on two variables, that is, the target adaptation coefficient $h$ and environmental volatility $\delta$. According to $\delta h + \delta - 1 < 0$, the store manager will never choose to manipulate performance when $h < 1/(\delta + 1)$. In addition, the larger environmental volatility $\delta$ means the lower target adaptation coefficient $h$ that the headquarters should adopt. Accordingly, when facing a stable market circumstance (i.e., a small $\delta$), the headquarters can remain at a certain level of the target adaptation coefficient $h$ to motivate the store manager. Conversely, the store manager may inevitably manipulate performance when facing high environmental volatility. Thus, the headquarters should set the target adaptation coefficient to avoid performance manipulation on the basis of market circumstance (see Figure 2).

Furthermore, given that (21), we can find that the extent deflated performance manipulation is much more demanding when the influence of the reputation effect (reflected in $c$) increases (see Figure 3). The larger bargaining power coefficient $c$, which implies the more significant reputation effect in the compensation scheme, the deflated performance manipulation will be more hard to happen accompanied by increasing environmental volatility.

When $\Delta\pi_1^* = 0$, the optimal pay-for-performance sensitivity is given by

$$\beta^* = \frac{(1 - h\delta)(1 - 2c\tau\delta) + \delta}{2(\delta h - 1)^2 + 2\delta}. \tag{22}$$

**Proposition 8.** *With the existence of deflated performance manipulation, one has*

$$\begin{aligned} \frac{\partial \Delta\pi_1^*}{\partial \beta} &> 0, \\ \frac{\partial \Delta\pi_1^*}{\partial h} &> 0, \\ \frac{\partial \Delta\pi_1^*}{\partial c} &< 0. \end{aligned} \tag{23}$$

This proposition shows that the extent of manipulation could be enhanced by increasing the pay-for-performance sensitivity or increasing the target adaptation coefficient or decreasing the bargaining power. The target adaptation coefficient implies the influence of the ratchet effect, the bargaining power indicates the reputation effect, and the pay-for-performance sensitivity measures the relation of compensation and performance. Increasing pay-for-performance sensitivity gives a rise to the extent that the manager influences performance by manipulation and thus will also promote the manipulation. If $c\tau > \beta h$, the inhibition of the reputation effect for manipulation is superior to the motivation of the ratchet effect, and the interaction leads to lower manipulation.

### 4.2. Optimal Effort Implications

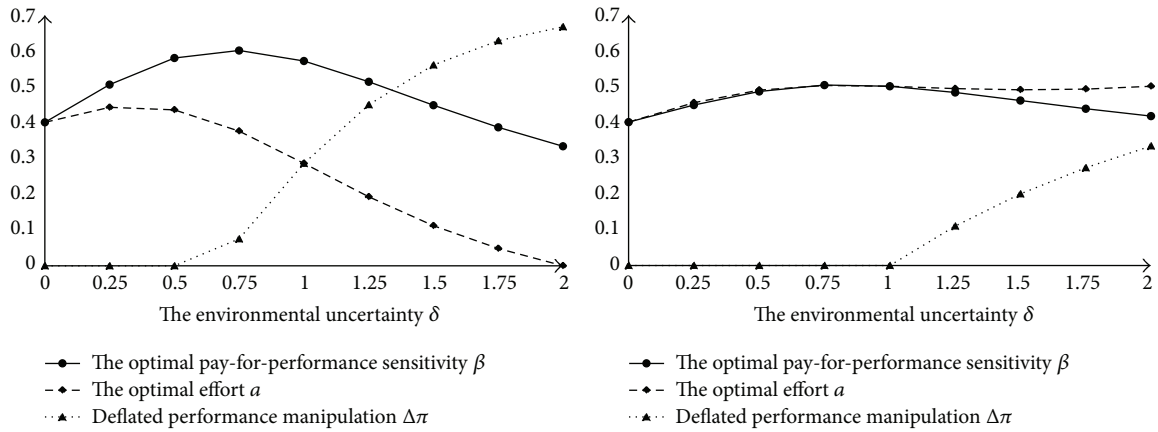
**Proposition 9.** *If* $1 - h\delta > 0$ *(i.e., h is small), the store manager will still work hard even though the pay-for-performance sensitivity* $\beta$ *is equal to zero. However, if* $1 - h\delta < 0$ *and*

$$\beta > \frac{c\tau\delta}{h\delta - 1}, \tag{24}$$

*the optimal effort is*

$$a_1^* = 0. \tag{25}$$

Traditional static agency model result in (16) shows that the store manager makes no effort only when the pay-for-performance sensitivity is equal to zero. According to




(a) The case that the bargaining power coefficient is equal to zero ($c = 0$)

(b) The case that the bargaining power coefficient is equal to 0.5 ($c = 0.5$)

Figure 3: Numerical simulations in different levels of the bargaining power coefficient and the environmental volatility. (Note that other parameters are given by $\lambda = 1$, $\gamma = 0.5$, $\tau = 0.5$, $k = 1$, $b = 1$, and $h = 0.5$.)

Proposition 9, however, we can draw that even if the pay-for-performance sensitivity is equal to zero, the reputation effect will still motivate the manager to make effort. Thus, the positive incentive effect caused by the explicit material incentive can be partly replaced by the reputation. On the other side, in the case of a large $h$ (i.e., $h\delta - 1 > 0$), even if the headquarters provides a certain amount of explicit incentives (as (24) does), the store manager instead makes no effort. This is because the managerial effort is also affected by the ratchet effect. Reducing effort leads to a lower performance target in the next period, and hence if the ratchet effect is obvious enough, potential benefits of being lazy outweigh the explicit income from being hard-working. Thus, the store manager betrays the explicit compensation and long-term implicit incentive and makes no effort.

**Proposition 10.** *From (18), there is*

$$\frac{\partial a_1^*}{\partial h} < 0, \qquad \frac{\partial a_1^*}{\partial c} > 0. \tag{26}$$

*If* $1 - h\delta < 0$*, we can draw that*

$$\frac{\partial a_1^*}{\partial \beta} < 0. \tag{27}$$

*Otherwise,*

$$\frac{\partial a_1^*}{\partial \beta} > 0. \tag{28}$$

This proposition shows that the effort could be increased by decreasing the target adaptation coefficient or enhancing the bargaining power. If $1 - h\delta < 0$, the effort decreases with increasing the pay-for-performance sensitivity. In this case, the high target adaptation coefficient will enlarge the negative effect of the pay-for-performance sensitivity. Thus, different form the traditional conclusion [35], the pay-for-performance sensitivity instead leads the store manager to reduce her effort. When $c\tau > \beta h$, the motivation of the reputation effect to the managerial effort is greater than the inhibition role of the ratchet effect, and such interactive effects will lead to the higher managerial effort.

## 5. Conclusions and Implications

This paper models a dynamic agency setting regarding the game between the headquarters and its store manager of retail firms, in which the ratchet effect and the reputation effect are jointly considered to impact the store manager's deflated performance manipulation as well as effort. Hence, we investigate how the store manager balances his effort and deflated performance manipulation under the forces of such two opposite effects and suggest the cases that the store manager is likely to manipulate performance. We find that the ratchet effect and environmental volatility interactively affect the existence of deflated performance manipulation, and the reputation effect plays a role in the extent of deflated performance manipulation.

Our findings provide several managerial implications for designing incentive scheme aiming at deflated performance manipulation. First, the headquarters should master the relationship between the ratchet effect and environmental volatility to avoid the side-effect of the former. The headquarters should rationally control the ratchet effect on the basis of the evaluation of environmental circumstances. In a stable environment, a reasonable level of ratchet effect cannot lead the store manager to manipulate performance. Conversely, when facing a high environmental volatility, the deflated performance manipulation will be inevitable, and the ratchet effect will positively affect the extent of manipulation.

Second, regarding performance target setting, the headquarters should not only be based on the store manager's historical performance, but also consider the environmental

volatility. Considering that the headquarters may be short of supervision capability as well as market insights relative to store managers, a feasible measure to evaluate a store manager's performance is by comparing her performance with both her historical and the average performance of all store managers, which can well reflect the current market circumstance. It is because that collusion across store managers seems to be impossible.

Third, the headquarters should balance the incentive power between fixed wage and pay for performance. Given that pay-for-performance sensitivity is likely to promote the store manager's performance manipulation, the headquarters can gradually enhance store managers' fixed wage to guarantee their security income in a high uncertain market environment.

Further research can be extended to study the impact of a comprehensive performance target, composed of historical and peer's performance, on the manager's performance manipulation. In addition, the store manager's social preferences, for example, loss aversion, temporal discounting, and inequity aversion, can also be incorporated into our model.

## Appendix

*Proof of Proposition 7.* If $\delta h + \delta - 1 < 0$, we can get

$$\Delta\pi_1^* = \frac{(1-\gamma)\left[(h\delta+\delta-1)\beta - c\tau\delta\right]}{\gamma k} < 0 \tag{A.1}$$

according to (19) in the text. Thus, the store manager will not manipulate no matter what the pay-for-performance sensitivity $\beta$ is.

While if $\delta h + \delta - 1 > 0$, we can get that

$$\Delta\pi_1^* = \frac{(1-\gamma)\left[(h\delta+\delta-1)\beta - c\tau\delta\right]}{\gamma k} > 0, \tag{A.2}$$

which is equivalent to

$$\beta > \frac{c\tau\delta}{h\delta+\delta-1}. \tag{A.3}$$

□

*Proof of Proposition 8.* With the existence of deflated performance manipulation, there is $\delta h + \delta - 1 > 0$ and $\beta > c\tau\delta/(h\delta + \delta - 1)$. Hence, we can get the partial derivatives as follows according to (19):

$$\begin{aligned}
\frac{\partial\Delta\pi_1^*}{\partial\beta} &= \frac{(1-\gamma)(h\delta+\delta-1)}{\gamma k},\\
\frac{\partial\Delta\pi_1^*}{\partial h} &= \frac{(1-\gamma)\delta\beta}{\gamma k},\\
\frac{\partial\Delta\pi_1^*}{\partial c} &= -\frac{(1-\gamma)\tau\delta}{\gamma k}.
\end{aligned} \tag{A.4}$$

As all of the parameters cannot be negative in practical setting and $\gamma \le 1$, thus we get $\partial\Delta\pi_1^*/\partial\beta > 0$, $\partial\Delta\pi_1^*/\partial h > 0$, and $\partial\Delta\pi_1^*/\partial c < 0$. □

*Proof of Proposition 9.* Based on (18), there is

$$a_1^* = \frac{\lambda\left[(1-h\delta)\beta + c\tau\delta\right]}{b}. \tag{A.5}$$

If $1 - h\delta > 0$, then

$$a_1^* = \frac{\lambda\left[(1-h\delta)\beta + c\tau\delta\right]}{b} > 0, \tag{A.6}$$

even if the pay-for-performance sensitivity $\beta$ is equal to zero.

While if $1 - h\delta < 0$, then

$$a_1^* = \frac{\lambda\left[(1-h\delta)\beta + c\tau\delta\right]}{b} < 0, \tag{A.7}$$

which is equivalent to

$$\beta > \frac{c\tau\delta}{h\delta-1}. \tag{A.8}$$

However, the optimal effort cannot be less than zero. Thus, there is $a_1^* = 0$ if $1 - h\delta < 0$ and $\beta > c\tau\delta/(h\delta - 1)$. □

*Proof of Proposition 10.* From (18), we can get the partial derivatives as follows:

$$\begin{aligned}
\frac{\partial a_1^*}{\partial h} &= -\frac{\lambda\delta\beta}{b} < 0,\\
\frac{\partial a_1^*}{\partial c} &= \frac{\lambda\tau\delta}{b} > 0,\\
\frac{\partial a_1^*}{\partial \beta} &= \frac{\lambda(1-h\delta)}{b}.
\end{aligned} \tag{A.9}$$

Considering the nonnegative parameters, we can get that

$$\begin{aligned}
\frac{\partial a_1^*}{\partial h} &< 0,\\
\frac{\partial a_1^*}{\partial c} &> 0.
\end{aligned} \tag{A.10}$$

And when $1 - h\delta < 0$,

$$\frac{\partial a_1^*}{\partial \beta} < 0, \tag{A.11}$$

otherwise,

$$\frac{\partial a_1^*}{\partial \beta} > 0. \tag{A.12}$$

□

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant nos. 71272064 and 71121061) and the Science and Technological Fund of Anhui Province for Outstanding Youth (Grant no. 1308085JGD07).

## References


[1] A. Beyer, I. Guttman, and I. Marinovic, "Optimal contracts with performance manipulation," *Journal of Accounting Research*, vol. 52, no. 4, pp. 817–847, 2014.

[2] G. P. Cachon and M. A. Lariviere, "Supply chain coordination with revenue-sharing contracts: strengths and limitations," *Management Science*, vol. 51, no. 1, pp. 30–44, 2005.

[3] J. E. Villegas, E. M. Gonzalez, M. P. Gonzalez, J. V. Anguita, and J. L. Vicent, "Experimental ratchet effect in superconducting films with periodic arrays of asymmetric potentials," *Physical Review B*, vol. 71, no. 2, Article ID 024519, 2005.

[4] B. W. Ickes and L. Samuelson, "Job transfers and incentives in complex organizations: thwarting the ratchet effect," *The RAND Journal of Economics*, vol. 18, no. 2, pp. 275–286, 1987.

[5] J. P. Choi and M. Thum, "The dynamics of corruption with the ratchet effect," *Journal of Public Economics*, vol. 87, no. 3-4, pp. 427–443, 2003.

[6] X. Freixas, R. Guesnerie, and J. Tirole, "Planning under incomplete information and the ratchet effect," *The Review of Economic Studies*, vol. 52, no. 2, pp. 173–191, 1985.

[7] M. L. Weitzman, "The 'ratchet principle' and performance incentives," *The Bell Journal of Economics*, vol. 11, no. 1, pp. 302–308, 1980.

[8] J. Bouwens and P. Kroos, "Target ratcheting and effort reduction," *Journal of Accounting and Economics*, vol. 51, no. 1, pp. 171–185, 2011.

[9] E. F. Fama, "Agency problems and the theory of the firm," *The Journal of Political Economy*, vol. 88, no. 2, pp. 288–307, 1980.

[10] P. M. Healy, "The effect of bonus schemes on accounting decisions," *Journal of Accounting and Economics*, vol. 7, no. 1–3, pp. 85–107, 1985.

[11] J. J. Gaver, K. M. Gaver, and J. R. Austin, "Additional evidence on bonus plans and income management," *Journal of Accounting and Economics*, vol. 19, no. 1, pp. 3–28, 1995.

[12] A. M. Goel and A. V. Thakor, "Why do firms smooth earnings?" *The Journal of Business*, vol. 76, no. 1, pp. 151–192, 2003.

[13] R. W. Holthausen, D. F. Larcker, and R. G. Sloan, "Annual bonus schemes and the manipulation of earnings," *Journal of Accounting and Economics*, vol. 19, no. 1, pp. 29–74, 1995.

[14] K. J. Crocker and J. Morgan, "Is honesty the best policy? Curtailing insurance fraud through optimal incentive contracts," *Journal of Political Economy*, vol. 106, no. 2, pp. 355–375, 1998.

[15] E. Goldman and S. L. Slezak, "An equilibrium model of incentive contracts in the presence of information manipulation," *Journal of Financial Economics*, vol. 80, no. 3, pp. 603–626, 2006.

[16] N. Burns and S. Kedia, "The impact of performance-based compensation on misreporting," *Journal of Financial Economics*, vol. 79, no. 1, pp. 35–67, 2006.

[17] K. J. Crocker and J. Slemrod, "The economics of earnings manipulation and managerial compensation," *The RAND Journal of Economics*, vol. 38, no. 3, pp. 698–713, 2007.

[18] R. J. Indjejikian, M. Matějka, K. A. Merchant, and W. A. Van Der Stede, "Earnings targets and annual bonus incentives," *The Accounting Review*, vol. 89, no. 4, pp. 1227–1258, 2014.

[19] S. Baiman, J. H. Evans III, and J. Noel, "Optimal contracts with a utility-maximizing auditor," *Journal of Accounting Research*, vol. 25, no. 2, pp. 217–244, 1987.

[20] R. A. Dye, "Earnings management in an overlapping generations model," *Journal of Accounting Research*, vol. 26, no. 2, pp. 195–235, 1988.

[21] J. S. Demski, "Performance measure manipulation," *Contemporary Accounting Research*, vol. 15, no. 3, pp. 261–285, 1998.

[22] M. D. Rablen, "Performance targets, effort and risk-taking," *Journal of Economic Psychology*, vol. 31, no. 4, pp. 687–697, 2010.

[23] A. J. Leone and S. Rock, "Empirical tests of budget ratcheting and its effect on managers' discretionary accrual choices," *Journal of Accounting and Economics*, vol. 33, no. 1, pp. 43–67, 2002.

[24] S. Roychowdhury, "Earnings management through real activities manipulation," *Journal of Accounting and Economics*, vol. 42, no. 3, pp. 335–370, 2006.

[25] L. G. Eldenburg, K. A. Gunny, K. W. Hee, and N. Soderstrom, "Earnings management using real activities: evidence from nonprofit hospitals," *The Accounting Review*, vol. 86, no. 5, pp. 1605–1630, 2011.

[26] P. H. Dybvig, "Dusenberry's ratcheting of consumption: optimal dynamic consumption and investment given intolerance for any decline in standard of living," *The Review of Economic Studies*, vol. 62, no. 2, pp. 287–313, 1995.

[27] B. Holmstrom, "Equilibrium long-term labor contracts," *The Quarterly Journal of Economics*, vol. 98, pp. 23–54, 1983.

[28] M. A. Meyer and J. Vickers, "Performance comparisons and dynamic incentives," *Journal of Political Economy*, vol. 105, no. 3, pp. 547–581, 1997.

[29] M. J. Dollinger, P. A. Golden, and T. Saxton, "The effect of reputation on the decision to joint venture," *Strategic Management Journal*, vol. 18, no. 2, pp. 127–140, 1997.

[30] K. Weigelt and C. Camerer, "Reputation and corporate strategy: a review of recent theory and applications," *Strategic Management Journal*, vol. 9, no. 5, pp. 443–454, 1988.

[31] M. Spence and R. Zeckhauser, "Insurance, information, and individual action," *The American Economic Review*, vol. 61, no. 2, pp. 380–387, 1971.

[32] J. A. Mirrlees, "The optimal structure of incentives and authority within an organization," *The Bell Journal of Economics*, vol. 7, no. 1, pp. 105–131, 1976.

[33] B. Holmstrom, "Moral hazard and observability," *The Bell Journal of Economics*, vol. 10, no. 1, pp. 74–91, 1979.

[34] S. J. Grossman and O. D. Hart, "An analysis of the principal-agent problem," *Econometrica. Journal of the Econometric Society*, vol. 51, no. 1, pp. 7–45, 1983.

[35] M. L. Weitzman, "Efficient incentive contracts," *The Quarterly Journal of Economics*, vol. 94, no. 4, pp. 719–730, 1980.

[36] B. Holmstrom and P. Milgrom, "Aggregation and linearity in the provision of intertemporal incentives," *Econometrica. Journal of the Econometric Society*, vol. 55, no. 2, pp. 303–328, 1987.

[37] O. Hart and B. Holmstrom, *The Theory of Contracts*, Department of Economics, Massachusetts Institute of Technology, 1986.

[38] B. Holmström, "Managerial incentive problems: a dynamic perspective," *The Review of Economic Studies*, vol. 66, no. 1, pp. 169–182, 1999.

[39] L. Martinez, *Reputation and Career Concerns*, Federal Reserve Bank of Richmond, 2006.

[40] D. Gill and R. Stone, "Fairness and desert in tournaments," *Games and Economic Behavior*, vol. 69, no. 2, pp. 346–364, 2010.

[41] D. Xing and T. Liu, "Sales effort free riding and coordination with price match and channel rebate," *European Journal of Operational Research*, vol. 219, no. 2, pp. 264–271, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 9840679, 8 pages
http://dx.doi.org/10.1155/2016/9840679

*Research Article*

# Decision-Making Approach to Selecting Optimal Platform of Service Variants

**Vladimir Modrak,[1] Slavomir Bednar,[1] and Pavol Semanco[2]**

[1]*Faculty of Manufacturing Technologies, Technical University of Kosice, Bayerova 1, 08001 Presov, Slovakia*
[2]*Lear Corporation Seating Slovakia s.r.o., Solivarska 1/A, 08001 Presov, Slovakia*

Correspondence should be addressed to Vladimir Modrak; vladimir.modrak@tuke.sk

Received 22 January 2016; Revised 9 May 2016; Accepted 22 May 2016

Academic Editor: Feng Yang

Copyright © 2016 Vladimir Modrak et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, it is anticipated that service sector companies will be inspired to follow mass customization trends of industrial sector. However, services are more abstract than products and therefore concepts for mass customization in manufacturing domain cannot be transformed without a methodical change. This paper is focused on the development of a methodological framework to support decisions in a selection of optimal platform of service variants when compatibility problems between service options occurred. The approach is based on mutual relations between waste and constrained design space entropy. For this purpose, software for quantification of constrained and waste design space is developed. Practicability of the methodology is presented on a realistic case.

## 1. Introduction

In general, services are defined as every kind of activity that contributes to the satisfaction of customer needs [1]. They play an ever-increasing importance in industrialized economies due to the need for the right balance between industrial and service sector. Particularly in changing business environment, there is a challenge for service providers to strengthen their position on the competitive market through innovative business strategies [2]. They also involve the advancing trend of transition from standard services to mass customized service packages [3].

However, mass customization is mostly related to tangible products and to manufacturing processes, whereas customization in services is studied only marginally [4]. On the other hand, growing mass customization tendencies induce that practically every product today has a service component in it.

Service variants in terms of mass customized environment are usually created for particular user groups. The problem with variety is not variety itself or increasing size of variety. The problem arises in the design process with an elimination of variants that are impractical due to presence of constraint(s). Such unviable variants are unadvisable for customers, because they lose time and time is too valuable for them to be wasted. Consideration of conflicts among variants causes enormous problems in the product or service design [5]. This suggests that a more integrated approach to mass customization is needed, which considers equally both the architecture design process and the subsequent implementation into configuration systems. In this context we propose by using information entropy theory a novel decision-making approach to selecting optimal platform of service variants.

## 2. Related Work

*2.1. Related Trends in Mass Customization of Services.* According to Zhang [6], recent growing trend and pressure on mass customization practice make companies expand their product and service mix better address preference of consumers. ElMaraghy et al. [7] state that this trend has several reasons including customers' demand for new product functions and features, different regional requirements, large number of market segments with different needs, and so forth. Another reason of this trend lays in the fact that

companies realized that effective way to achieve competitive edge leads not only through attractive products they produce but also through the services they provide [8]. Tien et al. [9] emphasized that a service mass customization strategy can be successfully adopted if the mass tailored services are delivered to customers at sufficiently short time scales. Liou et al. [10] focused on improving mass customization of airline services through determining dispensable ones associated with unimportant attributes. Particulalry this approach makes a connection with our research, in which we focus on unviable service variants. Unviable service variants not only are dispensable but also have negative impact on customer perception [11, 12]. In this context, authors, for example, [13, 14], argue that infeasible configurations might be hidden by using algorithm-based product configurators.

*2.2. Research on Optimal Platform of Service Variants.* There are several approaches to examining the impact of service variety on service development process; see, for example, sources [15]. As the literature shows, unviable service variants reflect higher operational complexity and decrease in efficiency of design [16, 17]. In this context, Tseng and Piller [18] proposed a generic variety structure including also configuration constraints. Therefore, the right product or service composition and their component compatibility need to be ensured [19]. As service industry is intentionally becoming more complex, a key enabler to achieve this tailoring of services to consumer needs is the application of efficient computer-aided variation technologies [20]. Such technologies usually allow service designers and decision-makers to determine the optimum level of customization and to eliminate infeasible service variants. Our paper has an ambition to contribute to the tools that can be incorporated in such computer-aided variation technologies.

## 3. Methodological Framework

*3.1. Theoretical Background of Decision-Making Approach.* The very first notion of complexity was outlined in the work of Shannon [21] where information theory was introduced. Few years later, information became a key complexity element for the description and analysis of the systems and information entropy $H_d$ for the discrete case has been defined by probability $P_i$ of $n$-state occurrence as follows [21]:

$$H_d = \sum_{i=1}^{n} P_i \log_2 P_i. \tag{1}$$

Differential information entropy of the probability density function $p(x)$ for continuous signals ($H_C$) has been expressed as

$$H_C = \int_{-\infty}^{\infty} p(x) \log_2 \left(p(x)\right) dx. \tag{2}$$

Krus [22] adopted the design information entropy for multidimensional case ($H_X$) in the following form:

$$H_X = \int_{-\infty} p(x) \log_2 \left(p(x) S\right) dx, \tag{3}$$

where $D$ as the design space within the design $x$ is defined. $S$ is the size of the design space expressed as

$$S = \int_D x\, dx. \tag{4}$$

In the case of general multivariable design space, its information entropy $H$ can be expressed as follows:

$$H = \log_2 \frac{S}{s}, \tag{5}$$

where $s$ is the region of uncertainty for the final design of validated system architecture.

According to Krus [22], each particular design "$x$" with regard to its design space has information entropy $H_X$:

$$H_X = \log_2 n_s, \tag{6}$$

where $n_s$ is number of all possible unique design alternatives (representing the so-called complete design space) as a result of a combination of service options.

There are many real cases, in which some variants are impractical due to presence of constraint(s). Then information entropy of constrained design space $H_C$ can be enumerated as

$$H_C = \log_2 n_v, \tag{7}$$

where $n_v$ is number of viable design alternatives.

Because the higher number of all possible design variants has more positive impact on consumers than smaller constrained design space, then information entropy of constrained design space in terms of mass customization environment should be maximized. In this sense, entropy of constrained design space can be considered as positive entropy.

In this context, Krus [22] proposed to express a quality of a modular design through the rest of the design space that is outside the constrained design space by the term "waste" information entropy of design space and to quantify it using the following formula:

$$H_W = H_X - H_C. \tag{8}$$

Waste information entropy of the complete design space is considered as negative entropy.

*3.2. Notation.* The notation used in the proposed methodological framework is summarized as follows:

$x$: particular design,

$n_s$: number of all possible unique design alternatives (complete design space),

$n_v$: number of viable design alternatives (restricted design space),

$H_X$: entropy of complete $D$,

$H_C$: entropy of constrained $D$,

$H_W$: waste information entropy,

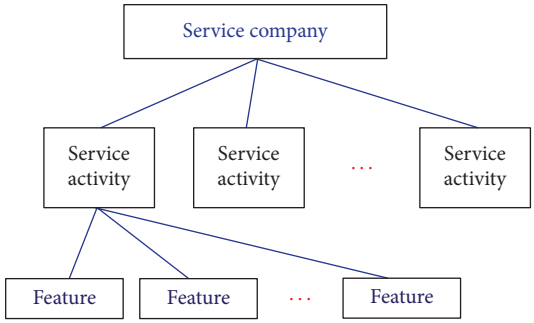



Figure 1: A higher level ontology of service company.

$D_0, D_1, \ldots, D_p$: design platforms,

$f_i, i = (1, \ldots, m)$: service features,

$A_{ij}, j = (1, \ldots, n)$: service attributes,

$O_q = \{O_1, O_2, \ldots, O_r\}$: service options,

$C_{ij}^V$: voluntary service components,

$C_{ij}^D$: delimited service components,

$C_{ij}^C$: compulsory components,

$l$: number of components chosen from a set of "$k$,"

$k$: number of single type components of $i$th feature and $j$th attribute.

*3.3. Proposed Decision-Making Method.* Based on above preconditions, we propose the following decision-making procedure to select optimal platform of service variants by using mutual relations between $H_W$ and $H_C$.

Let us have a service company providing several parallel mass customized service activities (see Figure 1.) As it is known, in order to maintain sustainability of their core activities, each of the activities might be continuously analyzed and subsequently optimized. Obviously, there are many aspects of service activities that should be analyzed. However, customers' view has strong priority. A questionnaire is not always the most effective method as some general useful findings that could be easily implemented are unnoticed. Among them is a known fact that infeasible options are unwelcome. The following method shows one possible way to reduce or eliminate them in individual service activity.

We start by taking the so-called draft design platform $D_0$ defined by $n_{s_0}$ and $n_{v_0}$. This starting platform can be either existing system of service activities or the first version of developed system of service activities, where $n_{s_0}$ presents a number of unique design alternatives as a result of a combination of service options and $n_{v_0}$ is a number of viable design alternatives.

Let us further assume that we remove from service platform $D_0$ some components, which are in conflict with other ones. Then $D_0$ can be transformed into a new state with $n_{s_1}$ and $n_{v_1}$ denoted as $D_1$. If we would continue in such a reduction of components, the design platform $D_1$ will be modified to $\{D_2, D_3, \ldots, D_p\}$, while $p$ represents a number of actions taken to reduce the system design space and waste entropy.

In order to compare two arbitrary design platforms, for example, $D_0$ and $D_1$, the following two measures are proposed:

$$\Delta H_{W_{0,1}} = \left| \frac{H_{W_1}}{H_{W_0}} - 1 \right|, \qquad \Delta H_{W_{0,1}} = \left| \frac{H_{C_1}}{H_{C_0}} - 1 \right|. \tag{9}$$

Then if $\Delta H_{W_{0,1}} > \Delta H_{C_{0,1}} \rightarrow$ design platform $D_1$ is more preferable for mass customization (MC) than $D_0$. To compare three alternative design platforms, the following subprocedure can be used. Let us suppose that design platforms $D_1$ and $D_2$ are more preferable for MC than $D_0$, based on the following criteria:

$$\Delta H_{W_{0,1}} > \Delta H_{C_{0,1}}, \qquad \Delta H_{W_{0,2}} > \Delta H_{C_{0,1}}. \tag{10}$$

Then one can select more preferable design platform between $D_1$ and $D_2$ using these three criteria:

(I) If $\Delta H_{W_{0,1}} - \Delta H_{C_{0,1}} > \Delta H_{W_{0,2}} - \Delta H_{C_{0,2}} \rightarrow$ design platform $D_1$ is more suitable than $D_2$.

(II) If $\Delta H_{W_{0,1}} - \Delta H_{C_{0,1}} < \Delta H_{W_{0,2}} - \Delta H_{C_{0,2}} \rightarrow$ design platform $D_2$ is more suitable than $D_1$.

(III) If $\Delta H_{W_{0,1}} - \Delta H_{C_{0,1}} = \Delta H_{W_{0,2}} - \Delta H_{C_{0,2}} \rightarrow$ both design platforms $D_1$ and $D_2$ are equally preferable for buyers.

Subsequently, proposed procedure for selection of optimal design platform is graphically depicted in Figure 2 and a simple software tool to select optimal design platform can be accessed online [23].

As seen in Figure 2, input parameters of the algorithm are $n_{s_0}$, number of all possible unique design alternatives, in other words alternatives that may theoretically be chosen including incompatible ones, and $n_{v_0}$, number of all viable design alternatives. Therefore, Sections 3.4 and 3.5 describe procedures for their quantification. Note that if there are no restrictions, then $n_{s_0} = n_{v_0}$ and the method is irrelevant.

*3.4. Service Variety Quantification without Restrictions.* In this section, wider description for quantification of all possible design alternatives understood as the design space $n_s$ will be treated.

This description refers to hierarchy of service ontology, as shown in Figure 3. To explain the hierarchy, the following working definitions can be useful. Service activity is here considered as a compact package of services divided into features $\{f_1, \ldots, f_m\}$. Feature consists of a set of attributes $f_i = \{A_{i1}, A_{i2}, \ldots, A_{ij}\}$.

Attributes represent characteristics of individual feature and will be denoted as $A_{ij}$, where $j = (1, \ldots, n)$.

Each attribute is represented by options $O_q = \{O_1, O_2, \ldots, O_r\}$, but the number of options depends

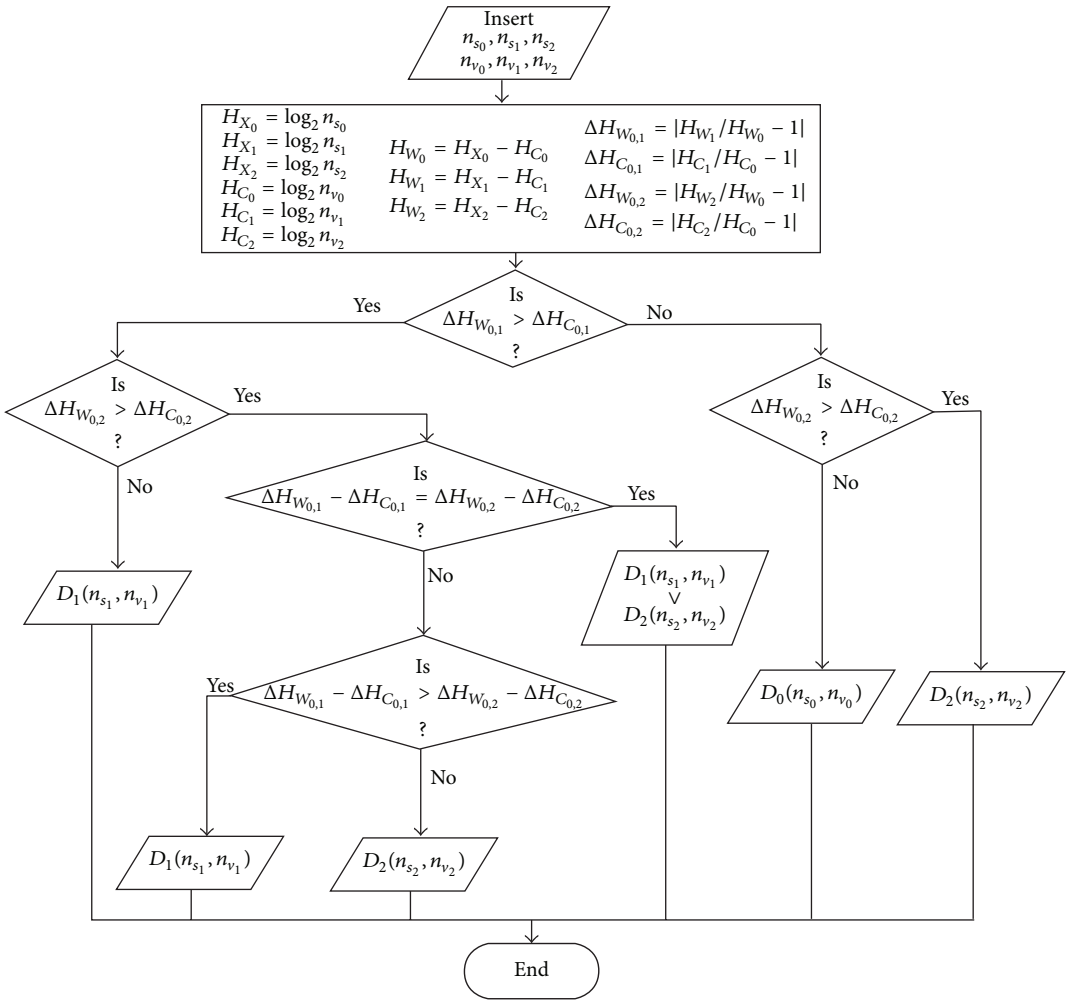



Figure 2: Procedure for the selection of optimal design platform.

on structure of attribute components which can be of different nature. We distinguish between three types of such components.

The simplest type of component is when customer can select from a portfolio of items on a voluntary basis. This means that all combinations of offered items and even no selection are possible. This type of option is denoted as voluntary option $C_{ij}^{V}$.

Delimited components $C_{ij}^{D}$ have limited selection possibilities specified by restriction rule $0 \leq l < k$, where $l$ is number of components to be chosen by customer from a set of components $C_{ij}^{D}$.

Compulsory components $C_{ij}^{C}$ have three possibilities when selecting desired service option or the so-called selection rules: $1 \leq l < k$.

Their practical importance, especially in the case of high numbers of available component, can be explained by the following example. Let us say that a structure of attribute components consists of all three component types. If we had three components $C_{ij}^{V}$ and two components $C_{ij}^{D}$, while selection of none and single component is allowed, and three compulsory components, where selection of at most two of them is allowed, then one would obtain in total 128 service options, as in Figure 4.

Enumeration of the final value of attribute options is based on the previously developed combinatorial metric of variability [24]:

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n} O_{ij} &= \left(2^{k}\right)\left(\sum_{j=1}^{n} C_{ij}^{C}\left(\frac{k!}{l!\,(k-l)!}\right)\right)\\ &\cdot\left(\sum_{j=1}^{n} C_{ij}^{D}\left(\frac{k!}{l!\,(k-l)!}\right)\right)\\ &= 2^{3}\left(\left(\frac{2!}{0!\,(2-0)!}\right)+\left(\frac{2!}{1!\,(2-1)!}\right)\right)\\ &\cdot\left(\left(\frac{3!}{1!\,(3-1)!}\right)+\left(\frac{3!}{2!\,(3-2)!}\right)\right) = 2^{3}\times(1+2)\\ &\times(3+3) = 8\times 3\times 6 = 128,\end{aligned} \tag{11}$$

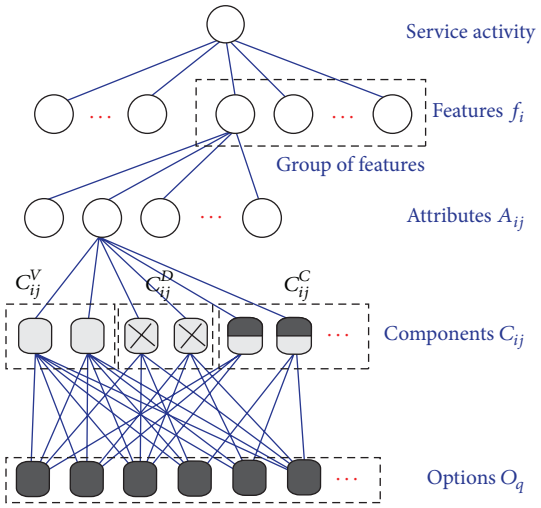



Figure 3: Hierarchy of service ontology.

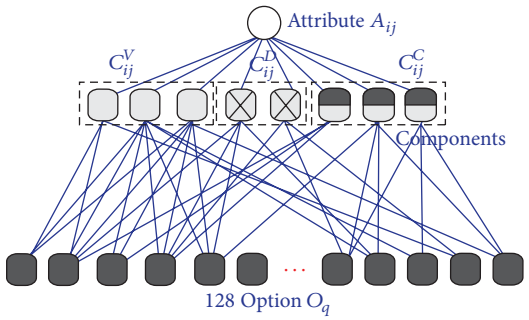



Figure 4: An example of relation between service components and resulting number of options.

where $l$ is the number of components to be chosen from a set of all components $k$ belonging to feature $f_i$, component $C_{ij}^C$ adopts the condition that $1 \leq l < k$, and component $C_{ij}^D$ adopts the condition that $0 \leq l < k$.

Current service variety in terms of MC may involve high number of features and attributes. Then, each attribute can be represented by the different types of components. Services either have fixed features with voluntary attributes or are completely customizable. The problem of such services is in the quantification of all possible design alternatives that are understood as the design space $n_s$. Then service variety designers have to cope with the extent of variety already in the design stage of service development. In the case of complex services and products, there is a lack of consistent methodological frameworks allowing modeling representations of a large number of components.

Unlike enumeration of 128 design alternatives options of a single attribute $A_{ij}$ in Figure 4, service variety designers have to be aware of the design space of service activity consisting of multiple attributes $A_{ij}$ and even multiple features $f_i$. We now use an example of feature diagram representation in Figure 5 for product consisting of exactly two features $(f_1, f_2)$ and

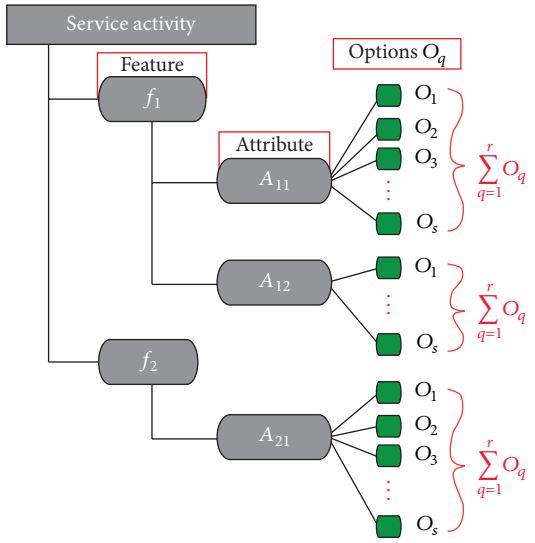



Figure 5: An example of enumeration of the design space for given activity.

three attributes with determined options. Subsequently, we can enumerate number of variants that are very useful when designing a new service, as shown in Figure 5. Summary number of variants of the service activity is calculated as follows:

$$n_s = \sum \text{Var} = \prod_{\substack{i=1 \\ j=1}}^{\substack{m \\ n}} \left( \sum_{q=1}^{r} O_q \right)_{i_j}. \tag{12}$$

This calculation of number of variants will be further used when applying proposed decision-making algorithm.

*3.5. Service Variety Quantification with Restrictions.* Real world models are seldom without any restriction; they are constrained due to functionality and compatibility issues. Therefore our aim is to propose a decision-making method to select optimal service platform where restrictions are predefined.

Constraints are specifying how two different options $O_{ij}$ cooperate within single service activity architecture. A restriction rule $R$ has two possible expressions. It can be either

(a) requiring rule; for example, $(O_1)_{ij} \rightarrow (O_{1+r})_{ij}$, then selection of option $O_1$ automatically requires selection of option $O_{1+r}$ in the same variant, or

(b) excluding rule; for example, $\neg((O_1)_{ij} \wedge (O_{1+r})_{ij})$, then selection of option $O_1$ excludes selection of option $O_{1+r}$ in the same variant.

Applying such restrictions within service architecture, system's design space $n_s$ is constrained to viable design space $n_v$.

Let us use the feature model from Figure 5 with four options on the first attribute $(\{O_1 \text{ v } O_2 \text{ v } O_3 \text{ v } O_4\} \rightarrow A_{11})$, two options at the second attribute $(\{O_1 \text{ v } O_2\} \rightarrow A_{12})$, and

```
INPUT:    S_N = Size of each subgroup, N - number of subgroups
          M = binary Service Model {1 - if i-th row links with j-th column; 0 - otherwise},
          GR = general restriction - diagonal restriction
          SR = specific restriction - off-diagonal restriction
OUTPUT:  ΣVar = Sum of variation with specified restricion
PROCEDURE Sum_Var(S_N, M, Iteration=0, Row=1, From=1)
          IF N<2 THEN RETURN ΣVar ← 0
          i = Row
          DO
                  FOR j=1,...,S_Iteration+1
                          IF Iteration=N-2 THEN ΣVar+=M[i,j]*Product(M[j]);
                          ELSE IF M[i,j]*Product(M[j])<>0 THEN
                                ΣVar+=M[i,j]*Sum_Var(S_N,M,Iteration+1,i,j);
                  i=i+1;
                  IF ITERATION<>0 THEN RETURN ΣVar;
          UNTIL i<= S_Iteration;
          RETURN ΣVar
END PROCEDURE
INITIALIZATION
Sum_Var(S_N,M)
```

Algorithm 1: Pseudocode for quantification of design spaces ($n_s$, $n_v$).

four options at the third attribute ($\{O_1 \text{ v } O_2 \text{ v } O_3 \text{ v } O_4\} \rightarrow A_{21}$), to determine the extent of both systems design and viable spaces $n_v$ and $n_s$. Three restrictions on design variants can be defined on the model for this purpose as follows:

$$\begin{aligned} &R\#1\text{: } \vdash \left((O_1)_{11} \wedge (O_1)_{12}\right), \\ &R\#2\text{: } \vdash \left((O_1)_{21} \wedge (O_2)_{12}\right), \\ &R\#3\text{: } \vdash \left((O_2)_{11} \wedge (O_1)_{21}\right). \end{aligned} \tag{13}$$

Subsequently, we may create a matrix of relations among systems options using a design matrix (Figure 6), where restrictions are assigned.

Here, relations above the diagonal express all possible collaborations within a model and "$X$" relates to the three restriction rules $R$#1–3. To resolve this restricted model, pseudocode formulation of the procedure is applied (see Algorithm 1). User tool of proprietary software developed for creation of matrix and determination of $n_s$ and $n_v$ values can be reached online [25]. The procedure begins with input variables, namely, numbers of options per each system attribute. Subsequently, code generates an interrelation matrix $\sum(O_q)_{ij} x \sum(O_q)_{ij}$. At the same time, design space of the model is determined ($n_s = 32$). Restricted model in turn provides only $n_v = 23$ viable service variants.

## 4. Realistic Case

In order to prove a practical importance of this research subject, it was decided to use realistic service problem situation to provide motivation for similar cases. This case model aims to demonstrate potential tool to decide about the optimal service design platform from customer perspective. Unlike simple example in Section 3.4, in realistic service activity, three service platforms $D_{0\text{–}2}$ are gradually identified, each with different (potentially optimal) values of $n_s$ and $n_v$ of the same service activity. This is to simulate the decision-making approach proposed in Section 3.3 to determine the most optimal service variant platform on the basis of the waste entropy concept $H_W$.

*4.1. Problem Description.* The following service model is a real representation of a hotel-booking system consisting of two steps (features $f_1$ for lodging and $f_2$ for boarding), as seen in Figure 7. Let us present a starting platform $D_0$, where each of the features has several compulsory attributes:

$$\begin{aligned} &\left(\{A_{11} \wedge A_{12} \wedge A_{13} \wedge A_{12}\} \longrightarrow f_1\right), \\ &\left(\{A_{21} \wedge A_{22}\} \longrightarrow f_2\right). \end{aligned} \tag{14}$$

Moreover, each of the attributes $A_{ij}$ has several options, as a result of component combinations. In the platform $D_0$, attribute $A_{11}$ enables customer to book, for example, one of the four room sizes ($\{O_1 \text{ v } O_2 \text{ v } O_3 \text{ v } O_4\} \rightarrow A_{11}$), or attribute $A_{12}$ allows choosing smoking and nonsmoking accommodation option: ($\{O_1 \text{ v } O_2\} \rightarrow A_{12}$).

Options of all individual attributes can be seen in Figure 7.

Complete design space of the platform $D_0$ is $n_{s_0} = 864$ variants. Platform $D_0$, however, has nine initial restrictions $R$#1–9. They are formulated in the form of exclusion; for example, 4-bed room option $(O_4)_{11}$ must not be in the same variant with superior room option $(O_2)_{13}$ $R$#1: $\vdash((O_4)_{11} \wedge (O_2)_{13})$. This rule restricts selection of service model variants and limits the design space $n_{s_0}$ to $n_{v_0}$. Then, platform $D_0$ of case service activity offers "only" $n_{v_0} = 288$ viable variants after application of all restrictions $R$#1–9.

Values of the service design spaces $n_{s_0}$ and $n_{v_0}$ have been obtained using the software tool described in Section 3.5.

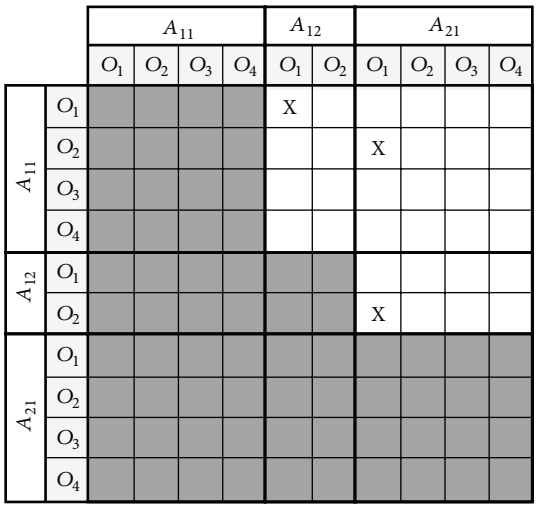



FIGURE 6: An example for design matrices for calculation of design spaces $n_s$, $n_v$.

In the second phase, alternative platform $D_1$ can be created. Let us assume that a service provider planned to optimize his service model by removing one of the attributes $A_{12}$ with two options in order to make model less differentiated but still customizable. Then the service design space has been reduced to $n_{s_1} = 432$ service variants and viable space to $n_{v_1} = 156$ service variants.

Finally, a removal of restriction of $R\#2$ has been proceeded to obtain the last platform $D_2$. After elimination of both the restriction $R\#2$ and attribute $A_{12}$, the service design space is $n_{s_2} = 432$ and $n_{v_2} = 168$ viable variants.

Then, using the software tool described in Section 3.5, we obtain values of the design spaces. Subsequently, we will apply the software tool for determination of optimal service platform, from Section 3.3, using the values of complete design spaces $n_{s_0} = 864$, $n_{s_1} = 432$, and $n_{s_2} = 432$ and viable design spaces $n_{v_0} = 288$, $n_{v_1} = 156$, and $n_{v_2} = 168$ as inputs. In the case of our service model, the most optimal platform is $D_2$.

From managerial point of view, platform $D_2$ presents more or less compromise between total reduction of infeasible variants and keeping all viable variants. Generally known, extreme solutions are not always the best. However, the method provides relevant alternative solutions with possible consequences and the final decision is up to service company management.

## 5. Conclusions

Effective and accurate execution of low-selling service components may bring relevant decrease of waste entropy as well as infeasible variants. Then more preferable platform of service variants can be identified. Naturally, presented decision-making algorithm for three platforms of service variants can be increased for unlimited number of platforms.

It was also proven in psychosocial domain (see, e.g., [26]) that any changes of long-term accepted rules in human behavior initiate disappointments or frustrations. These disappointments arise also when providers' options involve infeasible variants. To avoid such customer reaction, the proposed decision-making tool could be helpful.

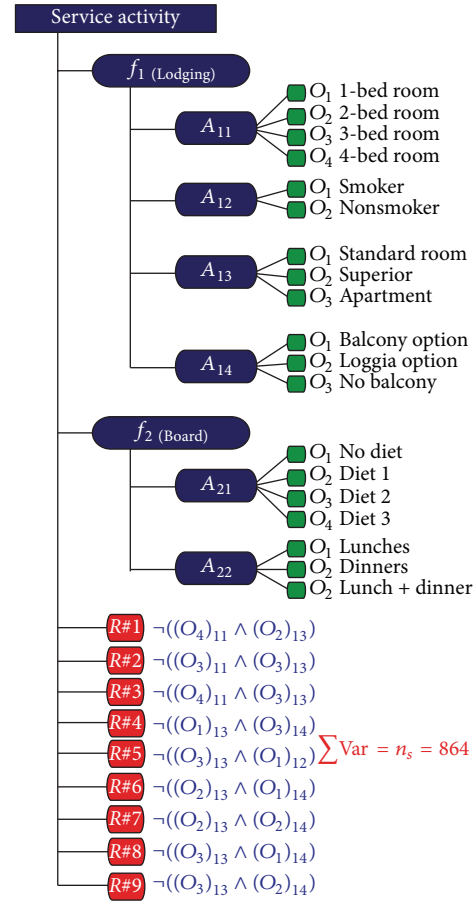



FIGURE 7: Feature diagram of the realistic case.

## Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

## References

[1] M. Hülsmann and J. Grapp, "Do service logistics need self-organization? A competence-based analysis of global supply chains in the movie production industry," in *Proceedings of the 4th International Logistics and Supply Chain Congress*, T. Baltacioglu, Ed., pp. 370–378, Izmir, Turkey, 2006.

[2] D. R. E. Thomas, "Strategy is different in service businesses," *Harvard Business Review*, vol. 56, no. 4, pp. 158–165, 1978.

[3] H. J. Bullinger and W. Schweizer, "Mass Customization in Process Industries, From Mass Production to Mass Customization," Fraunhofer Gesellschaft, Munich, 2005.

[4] G. Fragidis, D. Konstantas, and D. Paschaloudis, "A classification framework of value co-creation in electronic and mobile services," in *Exploring Services Science*, M. Snene and M.

Leonard, Eds., vol. 169 of *Lecture Notes in Business Information Processing*, pp. 40–55, Springer, Berlin, Germany, 2014.

[5] İ. Aydoğdu, A. Akın, and M. P. Saka, "Design optimization of real world steel space frames using artificial bee colony algorithm with Levy flight distribution," *Advances in Engineering Software*, vol. 92, pp. 1–14, 2016.

[6] L. L. Zhang, "Product configuration: a review of the state-of-the-art and future research," *International Journal of Production Research*, vol. 52, no. 21, pp. 6381–6398, 2014.

[7] H. ElMaraghy, G. Schuh, W. Elmaraghy et al., "Product variety management," *CIRP Annals—Manufacturing Technology*, vol. 62, no. 2, pp. 629–652, 2013.

[8] H. Gebauer, "An investigation of antecedents for the development of customer support services in manufacturing companies," *Journal of Business-to-Business Marketing*, vol. 14, no. 3, pp. 59–96, 2007.

[9] J. M. Tien, A. Krishnamurthy, and A. Yasar, "Towards real-time customized management of supply and demand chains," *Journal of Systems Science and Systems Engineering*, vol. 13, no. 3, pp. 257–278, 2004.

[10] J. J. H. Liou, L. Yen, and G.-H. Tzeng, "Using decision rules to achieve mass customization of airline services," *European Journal of Operational Research*, vol. 205, no. 3, pp. 680–686, 2010.

[11] P. Pitiot, M. Aldanondo, and E. Vareilles, "Concurrent product configuration and process planning: some optimization experimental results," *Computers in Industry*, vol. 65, no. 4, pp. 610–621, 2014.

[12] P. T. Helo, Q. L. Xu, S. J. Kyllönen, and R. J. Jiao, "Integrated vehicle configuration system-connecting the domains of mass customization," *Computers in Industry*, vol. 61, no. 1, pp. 44–52, 2010.

[13] D. Mailharro, "A classification and constraint-based framework for configuration," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 12, no. 4, pp. 383–397, 1998.

[14] M. Aldanondo, E. Vareilles, and M. Djefel, "Towards an association of product configuration with production planning," *International Journal of Mass Customisation*, vol. 3, no. 4, pp. 316–332, 2010.

[15] B. El-Haik and K. Yang, "The components of complexity in engineering design," *IIE Transactions*, vol. 31, no. 10, pp. 925–934, 1999.

[16] D. T. Matt, "Application of Axiomatic Design principles to control complexity dynamics in a mixed-model assembly system: a case analysis," *International Journal of Production Research*, vol. 50, no. 7, pp. 1850–1861, 2012.

[17] P. Ahlstrom and R. Westbrook, "Implications of mass customization for operations management: an exploratory survey," *International Journal of Operations and Production Management*, vol. 19, no. 3, pp. 262–275, 1999.

[18] M. M. Tseng and F. Piller, Eds., *The Customer Centric Enterprise: Advances in Mass Customization and Personalization*, Springer Science & Business Media, 2011.

[19] M. Bonev, L. Hvam, J. Clarkson, and A. Maier, "Formal computer-aided product family architecture design for mass customization," *Computers in Industry*, vol. 74, pp. 58–70, 2015.

[20] R. O. Reitsma, *Innovating Mass-Customized Service*, Eburon Uitgeverij, 2011.

[21] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[22] P. Krus, "Design space configuration for minimizing design information entropy," in *ICoRD'15—Research into Design Across Boundaries Volume 1: Theory, Research Methodology, Aesthetics, Human Factors and Education*, vol. 34 of *Smart Innovation, Systems and Technologies*, pp. 51–60, Springer, New Delhi, India, 2015.

[23] Platform Decision Maker, 2016, http://web.tuke.sk/fvt-mms/MMS2014/decision.php.

[24] V. Modrak and S. Bednar, "Using axiomatic design and entropy to measure complexity in mass customization," *Procedia CIRP*, vol. 34, pp. 87–92, 2015.

[25] Design Space Quantificator, 2016, http://web.tuke.sk/fvt-mms/MMS2014/calc7.php.

[26] I. Fried, "Syndrome E," *The Lancet*, vol. 350, no. 9094, pp. 1845–1847, 1997.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 8306950, 8 pages
http://dx.doi.org/10.1155/2016/8306950

*Research Article*

# Motivation System of Crowdsourcing Community from a Supply Chain Perspective

**Jiangang Pang[1,2] and Zhiying Liu[2]**

[1]*Southwest University of Science and Technology, Mianyang, Sichuan 621010, China*
[2]*University of Science and Technology of China, Hefei, Anhui 230026, China*

Correspondence should be addressed to Jiangang Pang; 392210825@qq.com

Received 7 December 2015; Accepted 9 June 2016

Academic Editor: Zhimin Huang

Copyright © 2016 J. Pang and Z. Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper uses principal-agent theory to study the issue of incentivizing crowdsourcing communities. It proves that enterprises can generate innovation plans of high quality and expected utility using the crowdsourcing community. Outsourcers can encourage high-quality people to join by adopting a linear variable compensation scheme and make the low-quality people quit by requiring them to supply more effort. The paper also shows that enterprises' participation in crowdsourcing community innovation can effectively improve their innovative ability and that it is necessary for enterprises to construct an effective and cooperative innovative system combining crowdsourcing community innovation and their own internal innovation.

## 1. Introduction

Rising competition forces enterprises to continuously explore ways to improve production. Furthermore, the development of the Internet has offered enterprises tools that encourage innovation more broadly (open innovation based on the Internet is very popular in many enterprises; e.g., IBM built its own innovation platform, Innovation Jam, in 2001; Dell built the Dell Online Community covering the functions of ideastom, directdell, dellforums, and studiodell in 2002; InnoCentive, the global third-party Internet innovative platform, built in 2001, appealed to the world's top 500 companies). In 2006, Howe defined the cooperative innovative model of mass participation based on the Internet as crowdsourcing [1], wherein a company or an organization subcontracts assignments normally done by internal staff to online contractors. These assignments can include the development of new technology, completeness of a new design, improvement of a calculation method, and analysis of mass data, which are performed by crowdsourcing communities such as Innovation Jam, Dell Online Community, InnoCentive, Starbucks, and Witkey. For such a crowdsourcing model to be successful, it is imperative that online contractors find it in their best interest to participate. Therefore, it is important to design a motivation system for the crowdsourcing community that effectively increases the initiative of participants and thereby enhance the quality of the output.

This paper analyzes incentive mechanisms on community-based crowdsourcing and studies how to best carry out the design of these incentive mechanisms. Through the analysis of the enterprise, the community-based crowdsourcing business innovation scheme can obtain higher quality and can bring about higher expected utility for enterprises. Contract-issuing party that participates in the crowdsourcing community process, using linear variable compensation incentive strategy, can get higher level network public participation and at the same time can generate general technical level network public participation but it needs to give up on putting a lot of effort on it and better enhance the innovation quality, putting forward linear crowdsourcing community innovation incentive mechanism design. Based on the conclusions, putting forward enterprises to participate in the crowdsourcing community innovation can be more effective than enhancing the innovation ability of the enterprises; at the same time, the enterprises should establish innovation system based on crowdsourcing community innovation and

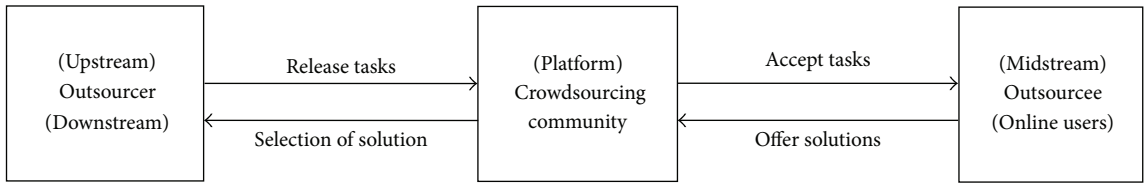



Figure 1: Crowdsourcing community supply chain.

enterprise innovation effective coordination. Before proceeding further, we define and clarify the jargon we use in this paper.

*1.1. Crowdsourcing Community Participants.* In existing crowdsourcing communities, the participants include outsourcers, the enterprises that need outsourcing services, and outsourcees, the solvers of problems (solvers of problems: in crowdsourcing community, participants mainly include outsourcer (enterprises in need), crowdsourcing community (crowdsourcing platform), and outsourcee (problem solvers, principally, are a mass participating crowdsourcing community); the solvers of problems are the main participant).

*Outsourcers.* The outsourcers are the enterprises or individuals that need the outsourcing service. There are two ways in which the outsourcer releases tasks. One way is by building a crowdsourcing community to attract outsourcees to solve problems by offering rewards. The other way is to partner with a third-party crowdsourcing community, such as help.zhubajie.com and InnoCentive.

*Outsourcees.* The outsourcee refers to crowdsourcing users who could be both professionals and nonprofessionals or teams who have an interest in and capability of studying tasks from all over the world. They can register via the third-party crowdsourcing community and accept such agreements as the selection of superior solutions, the fulfillment of offering rewards, and disposal methods of intellectual property and then serve as problem solvers who execute the activities of solving problems and submit solutions through the crowdsourcing community.

*Crowdsourcing Community.* A crowdsourcing community is the platform that connects outsourcers and outsourcees by developing either a crowdsourcing community (e.g., Amazon, the online retail giant, provides crowdsourcing service platform, Mechanical Turk) or a third-party crowdsourcing community, such as help.zhubajie.com and InnoCentive.

*1.2. Crowdsourcing Community Supply Chain.* According to the role and function of the participation body in the crowdsourcing community, solving problems is the production process of a crowdsourcing community, which builds the supply chain of the crowdsourcing community, as shown in Figure 1.

In the crowdsourcing community supply chain, the outsourcer who sits upstream subcontracts some internal tasks to online users through the crowdsourcing community and offers them rewards; meanwhile, the outsourcer downstream consumes the innovative solution produced by the midstream of the supply chain named outsourcee. The outsourcee is rewarded upon completion of tasks. The crowdsourcing community is a link in the supply chain, providing the platform that connects outsourcers and outsourcees. The value of the entire supply chain rises with the participation of users.

It is apparent that the key to guaranteeing orderly operation in the entire supply chain is keeping outsourcers and outsourcees motivated so that they take a positive role in the crowdsourcing community; therefore, it is significant and practical to explore the motivation system of crowdsourcing communities.

*1.3. Summary of Current Motivation Systems of Crowdsourcing Communities.* The current motivation systems of crowdsourcing communities fall into three categories. In the first, the outsourcer offers fixed bonuses, which is what most crowdsourcing communities do. In the second way, the outsourcers offer such awards as patents and accumulated points according to outsourcee's efforts. Participants providing innovative ideas to the crowdsourcing platform can obtain patent grade points which can be exchanged for cash. In the third way, the crowdsourcing community offers awards; http://www.chinavalue.net/ offers one such example. This platform sends half its share to registered users, who can accumulate by joining in innovative projects, such as the number of projects they participate in, the number of times of winning the bidding, and the number of published articles, to obtain shares and bonus.

## 2. Literature Review

### 2.1. Outsourcees

*2.1.1. Meeting the Psychological Needs.* The research of Teresa from the Business College at Harvard and Robert J. Stenberg from Yale University showed that creative people are good self-motivators and are willing to listen to their psychological needs. Therefore, many people participate in crowdsourcing for entertainment, self-expression, self-test, and so forth [2, 3]. Qiuyan showed that people pay the most attention to aspects such as fun, self-approval, and sense of belonging in a virtual community [4]. Most participants would like

to express their value proposition, to show their unique lifestyle, to meet their psychological needs, and to gain the conveniences the products bring them.

*2.1.2. Financial Returns.* Crowdsourcing is growing in popularity because people who offer creative solutions can get paid [5]. Lakhani and Wolf investigated the participants of SourceForge and found that one-third of interviewees indicated that receiving payment was their main motivation for joining SourceForge [6]. In addition, Lakhani and von Hippel investigated outsourcees who have solved tasks and received payment. The results of the investigation show that 65.8% of the interviewees have doctorate degrees and are engaging in scientific research and the main purpose of their participation is to get paid [7]. Organisciak tried to explain the participation motives of users through Maslow's hierarchy of needs and suggested that the purpose of participants is to meet the needs of society and growth. However, the investigation of 13 crowdsourcing websites showed that offering rewards can attract users to participate in crowdsourcing actively [8].

*2.1.3. Learning New Knowledge and Skills.* Brabham also investigated 651 participants of iStockphoto and the results showed that learning new knowledge and skills is the primary motivation for participation [9]. Through the Uses and Gratifications Theory, Brabham explained that users would like to take part in the creativity of the product because they can learn from the community and get more benefits than the manufacturer [10]. Crowdsourcing community provides a platform for like-minded people to publish their creativity freely which comprises sharing sources [11].

According to the summary and analysis of existing research, financial returns are the main motivation for crowdsourcing participants.

*2.2. Outsourcers*

*2.2.1. Cost Reduction.* Some researchers indicated that crowdsourcing released tasks online and did not require recruiting employees or setting up offices, which could reduce great costs. Crowdsourcing offers a way to reduce costs [12–14].

*2.2.2. Making Use of Collective Wisdom.* Some online users have the ability to solve specific problems. Under a similar level of knowledge and skills, the collective wisdom is better than a single wise person. A company can let volunteers from all over the world and internal staff work together to solve problems rather than forcing inner staff to learn continuously [15], which was supported by Xiao et al. who considered that, in a business economy, one problem cannot be solved by a single person but requires people with different knowledge, information, and skills. Crowdsourcing makes full use of this collective wisdom using a network of people to solve complex problems [16–18].

*2.2.3. Producing Products More in Line with Consumers' Demand.* Crowdsourcing is an example of the culture of mass participation; that is, universal participation of crowdsourcing enables many hobbyists to develop their talents during participation and interaction [19]. In many crowdsourcing models, users will meet the requirements of more audiences and spread the information needed by themselves [20]. This production can better meet consumer needs, can shorten the distance between the brand and the users, and can establish brand trust and strengthen customer loyalty. Using sports television programs as an example, Shi pointed out that sports game television programs are almost all the same. However, after adopting crowdsourcing, the programs have been diversified, which also stimulates the awareness of mass participation. Using news media as an example, Wei pointed out that users can participate in a news release that made "my news, I decide" come true [21].

The main purpose of outsourcers is to collect wisdom and find higher-quality solutions including better technology, more adaption to market, and lower costs.

From this discussion, we can see that a crowdsourcing community can exist and run effectively because the upstream, midstream, and downstream gains benefit from the supply chain.

## 3. Model Assumption

The motivation system design of the supply chain in a crowdsourcing community is the contract design between the outsourcer and the outsourcee; that is, the outsourcer entrusts outsourcees to finish innovative tasks through crowdsourcing community expecting to get innovative solutions with a higher quality while the outsourcees complete the tasks according to the requirements and get paid accordingly.

(1) Enterprises that release tasks are called outsourcer. Those enterprises are mutually independent and risk neutral. The outsourcer's expected utility is equal to expected profit.

(2) Users who accept tasks are called the outsourcee. These users are mutually independent.

(3) $w$ is the regular fee the outsourcer pays, $b$ refers to the incentive pay factor, and $w$ is the regular payment the outsourcee gets. The incentive pay factor is the sharing proportion of the benefits the outsourcee gets from the outsourcer.

(4) $i$ is the technical level of the outsourcee, where $i$ equal to 1 means the technical level is very high and $i$ equal to 2 means the technical level is low.

(5) $p_i$ is the probability the outsourcer chooses the outsourcees. $i = 1$ is the high technical level outsourcee. $i = 2$ is the low technical level outsourcee. Therefore, $p_1 + p_2 = 1$.

(6) $a$ is the quality of the innovative solution offered by the outsourcee. Because the quality of the solution has a direct impact on the product or service offered by the outsourcer, it will affect income; then income is an increasing function of quality. We assume the function of the outsourcer's income is a linear function, which can be expressed as

$$R(a) = a + \omega, \tag{1}$$

where $\omega$ is the random variable from 0 to $\delta^2$, which is the exogenous uncertainty affecting the income.

(7) The outsourcer can affect the outsourcee's decision by adjusting payments. We assume the payoff function is linear and can be expressed as

$$T(R(a)) = w + bR(a). \tag{2}$$

(8) Outsourcee's cost is a function of his/her innovation plan's quality and technology level, denoted by $C(a, i)$. For the outsourcee at the same level, that is, on the condition that the value of $i$ is definite, the higher the quality of the provided plan, the bigger the cost $C$. On the condition that the quality $a$ of the provided plan is definite, the marginal cost of outsourcee, whose $i = 1$, is lower than that of outsourcee, whose $i = 2$; according to theoretical model of supply chain analysis, set the cost function of outsourcee as $C(a, i) = \theta_i a^2$, wherein $\theta_i$ is outsourcers' cost coefficient, according to the difference of technology level of outsourcee, $\theta_1 < \theta_2$.

(9) The maximum utility is recorded as $\overline{u}$ when the outsourcee does not perform any tasks.

According to the assumption, the expected income function of the outsourcer can be expressed as

$$E(R(a)) = E(a + \omega) = a \quad \mathrm{Var}(R(a)) = \delta^2. \tag{3}$$

The expected utility is equal to the expected income minus the payment to the outsourcee; that is,

$$\begin{aligned} EU &= E(R(a) - T(R(a))) = E(a + \omega - w - bR(a)) \\ &= a - w - ba. \end{aligned} \tag{4}$$

Because the outsourcee is risk averse, the expected profit of the outsourcee is equal to the expected income minus the cost of payment and risk. Urban considered the risk cost as $(1/2)rb^2\delta^2$ [22], where $r$ is the risk measurement of the outsourcee and $r \geq 0$. The expected income is the payment by the outsourcer, given by $w + bR(a)$. The cost of payment is $\theta a^2$. The expected profit can be written as

$$\begin{aligned} CE &= E\left(T(R(a)) - C(a, i) - \frac{1}{2}rb^2\delta^2\right) \\ &= w + ba - \theta a^2 - \frac{1}{2}rb^2\delta^2. \end{aligned} \tag{5}$$

The equivalent utility of the outsourcee is equal to the expected profit. Therefore, to motivate the outsourcee to take part in the crowdsourcing community, the design of motivation system must meet the outsourcee's participation constraint and incentive compatible constraint.

The outsourcee's participation constraint is

$$w_i + b_i a_i - \theta_i {a_i}^2 - \frac{1}{2}r{b_i}^2\delta^2 \geq \overline{u} \quad i = 1, 2. \tag{6}$$

The outsourcee's incentive compatible constraint is

$$\begin{aligned} &w_1 + b_1 a_1 - \theta_1 {a_1}^2 - \frac{1}{2}r{b_1}^2\delta^2 \\ &\quad \geq w_2 + b_2 a_2 - \theta_2 {a_2}^2 - \frac{1}{2}r{b_2}^2\delta^2. \end{aligned} \tag{7}$$

And the maximum value is $\max CE$. Supposing $\partial CE/\partial a = 0$, then

$$a_i = \frac{b_i}{2\theta_i} \quad i = 1, 2. \tag{8}$$

Based on the above analysis, the model of the motivation system is

$$\begin{aligned} \max \quad & \sum_i p_i (a_i - w_i - b_i a_i) \quad i = 1, 2 \\ \text{s.t.} \quad & (IR)\ w_i + b_i a_i - \theta_i {a_i}^2 - \frac{1}{2}r{b_i}^2\delta^2 \geq \overline{u} \quad i = 1, 2 \\ & (IC)\ w_1 + b_1 a_1 - \theta_1 {a_1}^2 - \frac{1}{2}r{b_1}^2\delta^2 \\ & \geq w_2 + b_2 a_2 - \theta_2 {a_2}^2 - \frac{1}{2}r{b_2}^2\delta^2 \\ & \qquad\qquad i = 1, 2 \\ & a_i = \frac{b_i}{2\theta_i} \quad i = 1, 2. \end{aligned} \tag{9}$$

## 4. Model Analysis

*4.1. Outsourcer Adopts the Same Fixed Payment and Incentive Coefficient for All Outsourcees.* Assuming the fixed payment and incentive coefficient offered by the outsourcer are $(\widehat{w}, \widehat{b})$, then the model of motivation system should be

$$\max \quad \sum_i p_i \left(a_i - \widehat{w} - \widehat{b} a_i\right) \quad i = 1, 2 \tag{10}$$

$$\text{s.t.} \quad (IR)\ \widehat{w} + \widehat{b} a_i - \theta_i {a_i}^2 - \frac{1}{2}r\widehat{b}^2\delta^2 \geq \overline{u} \quad i = 1, 2 \tag{11}$$

$$(IC)\ \widehat{w} + \widehat{b} a_1 - \theta_1 {a_1}^2 - \frac{1}{2}r\widehat{b}^2\delta^2 \geq \widehat{w} + \widehat{b} a_2 - \theta_2 {a_2}^2 - \frac{1}{2}r\widehat{b}^2\delta^2 \tag{12}$$

$$i = 1, 2$$

$$a_i = \frac{\widehat{b}}{2\theta_i} \quad i = 1, 2. \tag{13}$$

Putting (13) into (10) and (11), the Lagrange equation can be built as

$$\begin{aligned} \Gamma = \max & \sum_i p_i \left(\frac{\widehat{b}}{2\theta_i} - \widehat{w} - \widehat{b}\frac{\widehat{b}}{2\theta_i}\right) \\ & + \sum_i \lambda_i \left(\widehat{w} + \widehat{b}\frac{\widehat{b}}{2\theta_i} - \theta_i \left(\frac{\widehat{b}}{2\theta_i}\right)^2 - \frac{1}{2}r\widehat{b}^2\delta^2 - \overline{u}\right). \end{aligned} \tag{14}$$

When $\lambda_1 = 0$, $\lambda_2 = 1$, the optimal first-order condition can be solved as

$$\widehat{b} = \frac{p_1\theta_2 + p_2\theta_1}{2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1}. \tag{15}$$

Putting $\widehat{b}$ into (13), we get

$$\begin{aligned}\widehat{a}_1 &= \frac{p_1\theta_2 + p_2\theta_1}{2\theta_1\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]} \\ \widehat{a}_2 &= \frac{p_1\theta_2 + p_2\theta_1}{2\theta_2\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]}.\end{aligned} \tag{16}$$

$\widehat{a}_1$, $\widehat{a}_2$ are the quality of innovation plan provided by outsourcee with high-level technology and outsourcee with average-level technology, when outsourcers provide the same regular compensation and incentive coefficient.

Putting the result of $\widehat{b}$, $\widehat{a}_2$ into (11), we get

$$\widehat{\omega} = \overline{u} + \frac{\left(2\theta_2 r\delta^2 - 1\right)\left(p_1\theta_2 + p_2\theta_1\right)^2}{4\theta_2\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]^2}. \tag{17}$$

Putting $\widehat{\omega}$, $\widehat{b}$ into (10) and (5), we get

$$\begin{aligned}\widehat{EU} &= \frac{\left(p_1\theta_2 + p_2\theta_1\right)^2}{4\theta_1\theta_2\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]} - \overline{u} \\ \widehat{CE}_1 &= \frac{\left(p_1\theta_2 + p_2\theta_1\right)^2\left(\theta_2 - \theta_1\right)}{4\theta_1\theta_2\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]^2} + \overline{u} \\ \widehat{CE}_2 &= \overline{u}.\end{aligned} \tag{18}$$

*4.2. Outsourcer Adopts Different Fixed Payment and Incentive Coefficient to All Outsourcees.* Assuming the fixed payment and incentive coefficient offered by the outsourcer are $(w_i, b_i)$, $i = 1, 2$, then the model of motivation system is

$$\max \quad \sum_i p_i\left(a_i - w_i - b_i a_i\right) \quad i = 1, 2 \tag{19}$$

$$\text{s.t.} \quad (IR)\ w_i + b_i a_i - \theta_i {a_i}^2 - \frac{1}{2} r {b_i}^2 \delta^2 \geq \overline{u} \quad i = 1, 2 \tag{20}$$

$$\begin{aligned}&(IC)\, w_1 + b_1 a_1 - \theta_1 {a_1}^2 - \frac{1}{2} r {b_1}^2 \delta^2 \\ &\geq w_2 + b_2 a_2 - \theta_2 {a_2}^2 - \frac{1}{2} r {b_2}^2 \delta^2 \qquad i = 1, 2\end{aligned} \tag{21}$$

$$a_i = \frac{b_i}{2\theta_i} \quad i = 1, 2. \tag{22}$$

Putting (22) into (19), (20), and (21), the Lagrange equation can be built as

$$\begin{aligned}\Gamma = \max &\sum_i p_i\left(\frac{b_i}{2\theta_i} - w_i - \frac{{b_i}^2}{2\theta_i}\right) \\ &+ \sum_i \lambda_i\left(w_i + \frac{{b_i}^2}{4\theta_i} - \frac{1}{2} r {b_i}^2 \delta^2 - \overline{u}\right) \\ &+ \tau\left[\left(w_1 + \frac{{b_1}^2}{4\theta_1} - \frac{1}{2} r {b_1}^2 \delta^2\right)\right. \\ &\left. - \left(w_2 + \frac{{b_2}^2}{2\theta_2} - \frac{\theta_1 {b_2}^2}{4{\theta_2}^2} - \frac{1}{2} r {b_2}^2 \delta^2\right)\right].\end{aligned} \tag{23}$$

The optimal first-order condition can be calculated as $\lambda_1 = 0$, $\lambda_2 = 1$, $\tau = p_1$:

$$\begin{aligned}b_1 &= \frac{1}{1 + 2\theta_1 r\delta^2} \\ b_2 &= \frac{p_2\theta_2}{2p_2{\theta_2}^2 r\delta^2 + \theta_2 - p_1\theta_1} \\ w_1 &= \overline{u} + \frac{{p_2}^2\left(\theta_2 - \theta_1\right)}{4\left(2p_2{\theta_2}^2 r\delta^2 + \theta_2 - p_1\theta_1\right)^2} \\ &\quad + \frac{2\theta_1 r\delta^2 - 1}{4\theta_1\left(1 + 2\theta_1 r\delta_2\right)^2} \\ w_2 &= \overline{u} + \frac{\theta_2 {p_2}^2\left(2\theta_2 r\delta^2 - 1\right)}{4\left(2p_2{\theta_2}^2 r\delta^2 + \theta^2 - p_1\theta_1\right)^2}.\end{aligned} \tag{24}$$

Putting $b_1$, $b_2$ into (22), we get

$$\begin{aligned}a_1 &= \frac{1}{2\theta_1\left(1 + 2\theta_1 r\delta^2\right)} \\ a_2 &= \frac{p_2}{4p_2{\theta_2}^2 r\delta_2 + 2\theta_2 - 2p_1\theta_1}.\end{aligned} \tag{25}$$

Putting $a_1$, $a_2$, $b_1$, and $b_2$ into (19) and (5), we get

$$\begin{aligned}EU &= \frac{2p_2 r\delta^2\left(p_1{\theta_2}^2 + p_2{\theta_1}^2\right) + p_1\theta_2 + p_2\theta_1 - p_1\theta_1}{4\theta_1\left(1 + 2\theta_1 r\delta^2\right)\left(2p_2{\theta_2}^2 r\delta^2 + \theta_2 - p_1\theta_1\right)} \\ &\quad - \overline{u} \\ CE_1 &= \overline{u} + \frac{\left(\theta_2 - \theta_1\right){p_2}^2}{4\left(2p_2{\theta_2}^2 r\delta^2 + \theta_2 - p_1\theta_1\right)^2} \\ CE_2 &= \overline{u}.\end{aligned} \tag{26}$$

**Proposition 1.** *On the condition that outsourcers provide the same regular compensation and incentive coefficient, outsourcee with high-level technology will provide higher-level innovation plan. That is, on the condition that outsourcers provide the single regular compensation, crowdsourcing community can select outsourcee with higher-level outsourcee for outsourcers.*

Comparing $\widehat{a}_1$ with $\widehat{a}_2$, we get $\widehat{a}_1/\widehat{a}_2 = \theta_2/\theta_1$; since $\theta_2 > \theta_1$, therefore $\widehat{a}_1 > \widehat{a}_2$.

**Proposition 2.** *When the outsourcer offers a different fixed payment and incentive coefficient, the outsourcee with a higher technical level will also offer a higher-quality solution. Therefore, the crowdsourcing community can also select the outsourcee with a higher technical level when offering various systems of payment.*

Because $\theta_2 > \theta_1$,

$$\begin{aligned} a_2 &= \frac{p_2}{4p_2{\theta_2}^2 r\delta^2 + 2\theta_2 - 2p_1\theta_1} \\ &< \frac{p_2}{4p_2{\theta_2}^2 r\delta^2 + 2\theta_2 - 2p_1\theta_2}. \end{aligned} \tag{27}$$

According to the assumption, $p_1 + p_2 = 1$. Then, we get

$$\begin{aligned} &\frac{p_2}{4p_2{\theta_2}^2 r\delta^2 + 2\theta_2 - 2p_1\theta_2} = \frac{p_2}{2\theta_2\left(2p_2\theta_2 r\delta^2 + 1 - p_1\right)} \\ &= \frac{p_2}{2\theta_2\left(2p_2\theta_2 r\delta^2 + p_2\right)} = \frac{1}{2\theta_2\left(1 + 2\theta_2 r\delta^2\right)} \\ &< \frac{1}{2\theta_1\left(1 + 2\theta_1 r\delta^2\right)} = a_1. \end{aligned} \tag{28}$$

Therefore, $a_1 > a_2$.

**Corollary 3.** *No matter what kind of incentive measurement is adopted by outsourcers, the crowdsourcing community can select outsourcee with higher-level outsourcee for outsourcers. Crowdsourcing innovation can bring innovation plan with higher-tech quality for enterprises and benefit enterprises' innovation.*

**Proposition 4.** *On the condition that outsourcers provide linear variable compensation, outsourcee with high-level technology will provide higher-level innovation plan. It is in accordance with practical state. Because of the innovation plan with higher-tech quality provided by outsourcee, the outsourcers' income as well as outsourcee's income can be increased, so outsourcees are willing to provide innovation plan with higher quality. In fact, they are obtaining a future option income.*

First, let us compare $\widehat{a}_1$ with $a_1$. Hence,

$$\begin{aligned} \widehat{a}_1 &= \frac{p_1\theta_2 + p_2\theta_1}{2\theta_1\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]} \\ a_1 &= \frac{1}{2\theta_1\left(1 + 2\theta_1 r\delta^2\right)}. \end{aligned} \tag{29}$$

The result is $\widehat{a}_1 < a_1$.

Second, let us compare $\widehat{a}_2$ with $a_2$. Hence,

$$\begin{aligned} \widehat{a}_2 &= \frac{p_1\theta_2 + p_2\theta_1}{2\theta_2\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]} \\ a_2 &= \frac{p_2}{4p_2{\theta_2}^2 r\delta^2 + 2\theta_2 - 2p_1\theta_1}. \end{aligned} \tag{30}$$

When $p_1\theta_2 - p_2\theta_2 + p_2\theta_1 > 0$, the result is $\widehat{a}_2 > a_2$.

**Corollary 5.** *When the outsourcer offers the differentiated payment, the outsourcee with a normal technical level will offer a lower-quality solution compared with the offer of a fixed payment. Because the outsourcee with a normal technical level is pessimistic about future earnings, he will not try his best to offer a high-quality solution.*

**Proposition 6.** *No matter what payment strategy is adopted, the expected utility of an outsourcee with high technical level will be higher than that of an outsourcee with a normal technical level. Therefore, rewarding those outsourcees with a higher technical level is an effective way for motivation.*

Comparing $C\widehat{E}_1, CE_1, C\widehat{E}_2, CE_2$

$$\begin{aligned} C\widehat{E}_1 &= \frac{\left(p_1\theta_2 + p_2\theta_1\right)^2\left(\theta_2 - \theta_1\right)}{4\theta_1\theta_2\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]^2} + \overline{u} \\ C\widehat{E}_2 &= \overline{u} \\ CE_1 &= \overline{u} + \frac{\left(\theta_2 - \theta_1\right){p_2}^2}{4\left(2p_2{\theta_2}^2 r\delta^2 + \theta_2 - p_1\theta_1\right)^2} \\ CE_2 &= \overline{u}. \end{aligned} \tag{31}$$

Because $(p_1\theta_2 + p_2\theta_1)^2(\theta_2 - \theta_1)/4\theta_1\theta_2[2\theta_1\theta_2 r\delta^2 + 2(p_1\theta_2 + p_2\theta_1) - \theta_1]^2 > 0$, we can get $C\widehat{E}_1 > C\widehat{E}_2$.

And because $(\theta_2 - \theta_1){p_2}^2/4(2p_2{\theta_2}^2 r\delta^2 + \theta_2 - p_1\theta_1)^2 > 0$, then $CE_1 > CE_2$.

We get $C\widehat{E}_1 < CE_1$, $C\widehat{E}_2 = CE_2$ by calculating the differences.

**Proposition 7.** *On the condition that outsourcers provide linear variable compensation, outsourcee with high-level technology will obtain higher expected utility than that on the condition of single compensation. Outsourcee with average-level technology will obtain the same expected utility under two circumstances. Linear variable compensation incentive can stimulate outsourcee with high-level technology rather than outsourcee with average-level technology. Outsourcers use linear variable compensation strategy that can attract more outsourcees with high-level technology.*

**Proposition 8.** *According to Proposition 6, the outsourcer will get a higher-quality solution, which brings about increased income when adopting the diversified payment strategy. Therefore, the outsourcer can get a higher expected utility when adopting the diversified payment strategy.*

Comparing $E\widehat{U}$, $EU$

$$\begin{aligned} E\widehat{U} &= \frac{\left(p_1\theta_2 + p_2\theta_1\right)^2}{4\theta_1\theta_2\left[2\theta_1\theta_2 r\delta^2 + 2\left(p_1\theta_2 + p_2\theta_1\right) - \theta_1\right]} - \overline{u} \\ EU &= \frac{2p_2 r\delta^2\left(p_1{\theta_2}^2 + p_2{\theta_1}^2\right) + p_1\theta_2 + p_2\theta_1 - p_1\theta_1}{4\theta_1\left(1 + 2\theta_1 r\delta^2\right)\left(2p_2{\theta_2}^2 r\delta^2 + \theta_2 - p_1\theta_1\right)} \\ &\quad - \overline{u}. \end{aligned} \tag{32}$$

Then, we can find $E\widehat{U} < EU$ through calculation.

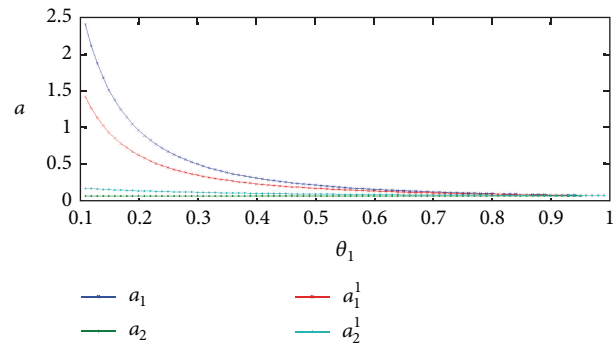



Figure 2: Trend of quality under various conditions.

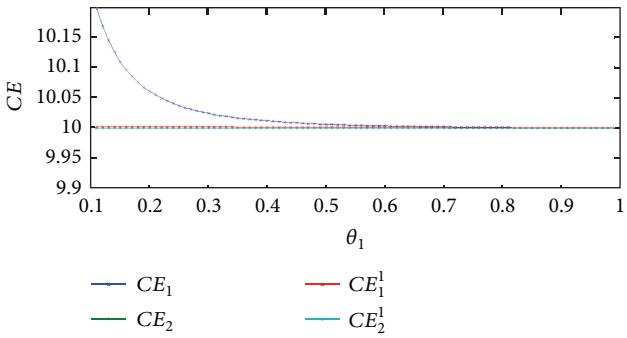



Figure 3: Trends of expected utilities.

*4.3. Numerical Analysis.* This part is to carry out example analysis in MATLAB, to concretely test the results from the models. The parameters in examples are the concrete values in the following parts. $\widehat{a}_1$ and $\widehat{a}_2$ will be represented as ${a_1}^1$ and ${a_2}^1$ in the MATLAB input. $C\widehat{E}_1$ and $C\widehat{E}_2$ will be represented as $C\widehat{E}_1^{\ 1}$ and $C\widehat{E}_2^{\ 1}$ in the MATLAB input.

When $\theta_2 = 1$, the abscissa is $\theta_1$ increasing (from 0.1 to 1), as shown in Figures 2 and 3.

It is shown in Figure 2 that (1) $\widehat{a}_1 > \widehat{a}_2$, $a_1 > a_2$. That is, no matter what kind of incentive mechanism is adopted, outsourcee with high-level technology can provide plan with higher quality. It is in accordance with the conclusions of Propositions 1 and 2 and Corollary 3. (2) One has $a_1 > \widehat{a}_1$, $\widehat{a}_2 > a_2$. That is, when linear variable compensation is used, outsourcee with high-level technology can provide plan with higher quality, and outsourcee with average-level technology can provide plan with lower quality. It is in accordance with the conclusions of Proposition 4 and Corollary 5.

It is shown in Figure 3 that (1) $C\widehat{E}_1 > C\widehat{E}_2$, $CE_1 > CE_2$. No matter what kind of incentive mechanism is used, the outsourcee with high-level technology will have higher expected utility, in accordance with the conclusion of Proposition 6. (2) One has $C\widehat{E}_1 < CE_1$, $C\widehat{E}_2 = CE_2$; that is, under the linear variable compensation strategy, outsourcee with high-level technology will obtain higher expected utility than that on the condition of single compensation, and outsourcee with average-level technology will obtain the same expected utility under two circumstances. It is in accordance with the conclusion of Proposition 7.

## 5. Results and Discussion

The model results have strong practical significance for crowdsourcing community design.

Firstly, no matter which kind of incentive mechanism is adopted, outsourcers can select the best innovation plan to improve innovation efficiency. The development of the Internet helped to develop open innovation and mass production and to release collective intelligence. Meanwhile, customers' needs become more various. Only when products sufficiently meet customers' need can they be sold on the market. Users innovation becomes the best way for enterprises to merge customers' needs. Outsourcees, as the online mass in crowdsourcing community, are the designers of innovation plans and customers. In fact, they are the users' innovation. They have a better idea about customers' needs, so the innovation plans are more attuned to customers' willingness. Therefore, crowdsourcing community can select the optimal innovation plan for enterprises. Meanwhile, because of the low cost of operating on the Internet, the selecting process is of low cost. Therefore, it is a beneficial strategy for enterprises to participate in crowdsourcing. In reality, it is the reason why the global top 500 companies actively take part in crowdsourcing.

Secondly, linear variable compensation strategy is outsourcers' better bid strategy choice. This conclusion provides basis for outsourcers' bid. In reality, there is a "lemon effect" (lemon effect: the most famous Akeerluofu thought is the so-called lemon market theory; ripe lemon is sweet, but a man says the green lemon is sweet, because he only has green lemon; this is a kind of negative psychological defense; there is a serious problem of information asymmetry between enterprises and IT enterprises; in an IT game between enterprises and enterprise customers, the final result is often a suboptimal choice, that is, the so-called "lemon effect") in crowdsourcing community innovation. Single regular compensation strategy makes online mass of outsourcees select simple crowdsourcing projects, and they are unwilling to invest more spare time in crowdsourcing projects. At the same time, because of simplifying tasks, outsourcers are unwilling to pay more, so the "lemon effect" comes into being. We need to extend the contract relationship between the outsourcer and the outsourcee. One strategy is that outsourcers provide longer-term compensation to stimulate outsourcee to participate in innovation continually. Just like continually upgrading software, outsourcers make outsourcees optimize their project design again and again.

## 6. Conclusions

This paper analyzed the motivation system of a crowdsourcing community from the perspective of the supply chain and designed a motivation system. The analysis showed that the outsourcer of an enterprise will get a higher-quality solution through the crowdsourcing community. During participation in a crowdsourcing community, the outsourcer will encourage more outsourcees with high technical level to join and force some of the outsourcees with normal technical level to quit in order to improve the quality of the solution

as long as the outsourcer adopts a diversified payment strategy.

However, this work has at least the following limitation, which will be studied in the future. In this motivation system design, we assume the users (the outsourcees) are mutually independent. However, users influence each other by mutual learning and interaction, which will be the focus of the next study.

## Competing Interests

The authors declare that there are no competing interests regarding the publication of this paper.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 71472172), Social Science Foundation of Sichuan (no. SC13E012), Soft Science Foundation of Sichuan (no. 2015ZR0167), and Foundation of Sichuan Library and Intelligence Research Center (no. SCTQ2016YB15).

## References

[1] J. Howe, "The rise of crowdsourcing," *Wired Magazine*, vol. 14, no. 6, pp. 176–183, 2006.

[2] J. Howe, *Crowdsourcing: Why the Power of Online Masses Drives the Future of Business*, CITIC Publishing House, Beijing, China, 2009.

[3] L. Changjiang, X. Jing, and S. Fan, "Realization mechanism of mutual trust among the enterprises in supply chain nodes," *Industrial Engineering and Management*, no. 2, pp. 89–91, 2005.

[4] Z. Qiuyan, "Empirical research on continued participation behaviors of users in crowdsourcing community," *Journal of Dalian University of Technology*, no. 3, pp. 1–6, 2011.

[5] D. Ding and J. Chen, "Coordinating a three level supply chain with flexible return policies," *Omega*, vol. 36, no. 5, pp. 865–876, 2008.

[6] K. R. Lakhani and R. G. Wolf, "Why hackers do what they do: understanding motivation and effort in free/open source software projects," in *Perspectives on Free and Open Source Software*, J. Feller, B. Fitzgerald, S. Hissam, and K. R. Lakhani, Eds., no. 3, pp. 3–21, MIT Press, Cambridge, Mass, USA, 2005.

[7] K. R. Lakhani and E. von Hippel, "How open source software works: 'free' user-to-user assistance," *Research Policy*, vol. 32, no. 6, pp. 923–943, 2003.

[8] P. Organisciak, *Why bother examining the motivations of users in large-scale crowd-powered online initiatives [M.S. thesis]*, University of Alberta, Master of Arts, 2010.

[9] D. Brabham, *Crowdsourcing*, MIT Press, 2013.

[10] D. C. Brabham, "Moving the crowd at threadless: motivations for participation in a crowdsourcing application," *Information Communication & Society*, vol. 13, no. 8, pp. 1122–1145, 2010.

[11] E. Simpson and S. Roberts, "Bayesian methods for intelligent task assignment in crowdsourcing systems," in *Decision Making: Uncertainty, Imperfection, Deliberation and Scalability*, vol. 538, pp. 1–32, Springer, 2015.

[12] F. Liu, "The application of knowledge management in the internet witkey mode in China," *Journal of Knowledge and System Sciences*, vol. 4, no. 4, pp. 32–41, 2007.

[13] Q. Liping, "The path of achieving innovation coordination selected by the Collaborative Innovation Network Organization," *Science & Technology Progress and Policy*, no. 5, pp. 16–19, 2011.

[14] Z. Peng and L. Ruoyu, "Research on crowdsourcing innovation motivation mechanism—based on the principal-agent theory," *Technology Economics and Management Research,*, no. 6, pp. 45–49, 2012.

[15] L. Liang and X. Binglei, "The supervision and incentive of supply chain," *Systems Engineering—Theory Methodology Applications*, no. 3, pp. 21–23, 2002.

[16] X. Lan and G. Changchun, "Crowdsourcing—Changing enterprises' innovation model," *Economic Research of Shanghai*, no. 3, pp. 35–41, 2010.

[17] J. Prpić, P. P. Shukla, J. H. Kietzmann, and I. P. McCarthy, "How to work a crowd: developing crowd capital through crowdsourcing," *Business Horizons*, vol. 58, no. 1, pp. 77–85, 2015.

[18] C. Changbin and C. Gongyu, "The relationship of the supply chain on the formation and development," *Industrial Technology & Economy*, vol. 25, no. 11, pp. 24–28, 2006.

[19] Z. Libin, "Research overview on problems of crowdsourcing," *Science & Technology Progress and Policy*, no. 3, pp. 154–158, 2012.

[20] S. Yingjie, "The prisoners dilemma analysis of cooperation innovation in supply chain," *Technology Economics and Management Research*, no. 5, pp. 25–33, 2009.

[21] W. Shuancheng, "The generation and development of crowdsourcing and the path observed when building the business model of crowdsourcing," *Shanghai Management Science*, no. 2, pp. 55–57, 2010.

[22] G. L. Urban and E. Von Hippel, "Lead user analyses for the development of new industrial products," *Management Science*, vol. 34, no. 5, pp. 569–582, 1988.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 6850308, 13 pages
http://dx.doi.org/10.1155/2016/6850308

*Research Article*

# The Effects of the Emission Cost on Route Choices of International Container Ships

**Hyangsook Lee,[1] Kang-Dae Lee,[2] and Sangho Choo[3]**

[1]*Graduate School of Logistics, Incheon National University, 12-1 Songdo-dong, Yeonsu-gu, Incheon 406-772, Republic of Korea*
[2]*Department of Packaging, Yonsei University, 234 Maeji, Heungup, Wonju, Kangwon-do 220-710, Republic of Korea*
[3]*Department of Urban Design & Planning, Hongik University, 72-1 Sangsu-dong, Mapo-gu, Seoul 121-791, Republic of Korea*

Correspondence should be addressed to Sangho Choo; shchoo@hongik.ac.kr

Received 18 December 2015; Accepted 5 May 2016

Academic Editor: Zhimin Huang

Copyright © 2016 Hyangsook Lee et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Maritime freight shipping has increased significantly and air pollution from international ships has grown accordingly, having serious environmental effects all over the world. This paper analyzes the effects of the emission cost on ocean route choices, focusing on international container ships. First, the paper formulates a freight network model that captures decisions and interactions of ocean carriers and port terminal operators in the maritime freight transport system. Then, the emission cost is calculated based on an activity-based approach as a component of the ocean transportation cost function. A case study is examined to find if the emission cost affects ocean route choices. The results indicate that the optimal ocean route and transportation cost are changed distinctively due to the emission cost. The research discusses how the emission cost plays a role in route changes and why ocean carriers have to consider these costs in their routing decisions.

## 1. Introduction

International freight shipping has been increasing significantly due to the growth of global trades and the following increment in maritime freight transportation. Up to 2008, the growth of world container traffic was a continuous process, with a change from 28.7 million TEUs in 1990 to 152.0 million TEUs in 2008, showing an increase of 430% (Bureau of Transportation Statistics). Slow economy of 2009-2010 had an impact on container flow, experiencing a drop of 49 million TEUs. However, the container movements have been rising again since 2010.

Nowadays, ocean carriers are regarded as one of the biggest sources of air pollution. International ships are usually powered by diesel engine system which emits major air pollutants such as nitrogen oxides ($NO_x$), carbon dioxides ($CO_2$), sulfur dioxide ($SO_2$), and particulate matters (PM). These are known to have a harmful influence on ship workers and local residents, causing asthma, respiratory disease, cardiac disorder, and so forth. Therefore, as maritime freight shipping has grown rapidly, the air pollution problem has become a serious and urgent issue all over the world.

There has been global effort to regulate air pollution. For example, the Kyoto Protocol that controls global warming gases took effect in 2005. International Maritime Organization (IMO) activated the regulations for air pollution from ships in the same year. Also, advanced countries like Japan, USA, and EU set the emission standard and tightened the control of air pollution caused from big ships and ports. In addition, some dominant ocean shipping companies started to consider the environment cost when transporting goods in aspects of environment-friendly company image and social contribution.

Carriers are transportation service providers in the freight transportation network. A set of carriers includes ocean carriers, land carriers, and port terminal operators. In general, particularly in the common case of a landlord port, a Port Authority leases marine terminals to private terminal operators. Port terminal operators offer the transportation service in port complex as a special type of carrier. Ocean

carriers typically choose departure and arrival marine ports in a cost-efficient manner. Therefore, ocean carriers are regarded as leaders and port terminal operators are followers in the maritime freight shipping market.

Carriers are also private transportation companies that attempt to maximize their profits; therefore they determine delivery routes (service processes) by minimizing the transportation cost that contains the operating cost, the travel time, and so forth. If the environmental cost is considered in the transportation cost function, their optimal route choices may be changed. This paper examines the effects of the emission cost on the routing pattern, concentrating on the emissions of container ships. For the emission cost estimation of ocean-going vessels, an activity-based approach is employed.

The structure of the paper is as follows. The next section presents an up-to-date literature review of the existing freight network models and relevant studies on the environmental cost. Section 3 defines carrier network structures and modeling approaches. Section 4 formulates a bilevel model that captures hierarchical interactions between ocean carriers and port terminal operators. Section 5 develops an algorithm to solve the bilevel model. Next, a numerical example is deployed to demonstrate the validity of the model and to analyze changes caused by the emission cost in optimal ocean routes and port operations.

## 2. Literature Review

Lots of works capturing decisions and relationships of key stakeholders (i.e., producers, consumers, shippers, carriers, and governments) involved in land transportation have been done. Regarding the seaborne freight network problem, relatively a small amount of literature exists. Zan [1], Kuroda et al. [2], and Zhang et al. [3] considered interactions between shippers and ocean carriers or port terminal operators, assuming decisions of the other carriers are given. Lee et al. [4, 5] presented hierarchical relationships among three types of carriers focusing on the different carrier problem in both competitive and cooperative markets. Lee et al. [6] captured interactions between shippers and three types of carriers by considering one more key stakeholder. Wang [7] and Min and Guo [8] developed more complicated mathematical models adding another major decision maker, Port Authority. These studies used equilibrium concepts (e.g., user equilibrium, Nash equilibrium, spatial price equilibrium, or compensation principle) to analyze the behavior of individual stakeholder groups. For multilevel games, Stackelberg game or multi-leader-follower game frequently was applied. Zhang et al. [9] explored competitive, cooperative, and Stackelberg congestion pricing for local regions. Practical pricing schemes between different administrative regions were presented employing both Nash game and Stackelberg game. On the other hand, G. S. Kim and T. S. Kim [10] presented cooperative development of logistics systems through the functional networking of port cities. The paper showed the allocation of logistics function and cooperative networking in logistics systematically and horizontally.

Several works have focused on environmental issues in the freight network problems. Jiang and Kronbak [11] presented a good summary of major research on estimating the external costs of maritime transportation and defined a generalized form of the voyage-based model. Berechman and Tseng [12] studied the emission cost of ships and trucks for eight main air pollutants in the Port of Kaohsiung, Taiwan. A bottom-up methodology was employed defining the engine load, load factor, and emission factor. Over 123 million dollars per year was estimated, indicating that tankers, container ships, bulk ships, and trucks are major contributors causing environmental problems. Yau et al. [13] developed a detailed maritime emission inventory for ocean-going vessels in Hong Kong. A bottom-up activity-based approach was adopted to estimate the emissions from main and auxiliary engines and boilers. Containers ships contributed over 70% of the total emissions of $NO_x$, $SO_2$, and $PM_{10}$. Saraçoğlu et al. [14] investigated the effects of ship sourced emission in Izmir Port, Turkey. Exhaust gas emissions from ships in the port were calculated by adopting a ship activity-based methodology. Total emissions of $NO_x$, $SO_2$, $CO_2$, HC, and PM, classified according to operation modes and ship types, were estimated and compared with relevant study results. Deniz et al. [15] studied maritime shipping emissions in Candarli Gulf, Turkey. Various emissions were calculated based on the activity-based emission model in Aliaga Bay and Nemrut Bay. Corbett et al. [16] discussed the effectiveness of the speed reduction on emissions for international shipping. By applying a profit-maximizing function that corporates the opportunity cost to estimate route-specific, economically efficient speeds, policy impacts of a fuel tax and a speed reduction mandate on $CO_2$ emissions were explored. On the other hand, Ahn and Rakha [17] investigated the effects of route choice decisions on vehicle energy consumption and emission rate for different types of vehicles using some emission estimation tools. Research results demonstrated that the faster route choice is not always the best from an environmental or energy consumption perspective. Park et al. [18] estimated trade-off relationships among logistics cost, time, and $CO_2$ emissions of the freight transportation system in Korea. The trade-off relationships between attributes were clearly shown and the desired level of modal split of intermodal freight transportation system was identified. Shukla and Alam [19] focused on the interaction between the vehicle emissions like $NO_x$, HC, CO, and $CO_2$ and traffic control measures. Results showed that high emission rates occurred during acceleration and delay events, while vehicle emissions were relatively low when the vehicle was idling.

In summary, most former studies focused on interactions between key stakeholders in the freight network problem. Some works estimated emissions and emission costs of trucks, vessels, or port sites, respectively. On the other hand, this research combined two issues, freight network problem and environmental problem. The paper formulates a freight network model capturing ocean carriers' optimal routing choices in determining marine ports using bilevel optimization programming. Here, the environmental cost is considered as a factor of the ocean transportation cost

function to demonstrate the effects of the emission cost on routing decisions by employing an activity-based approach.

## 3. Problem Definition

*3.1. Network Structure.* The carrier network is a physically detailed multimodal network that shows transportation routes and modes. Three types of networks, the marine transport network, the port network, and the land transport network, are defined for the maritime freight transportation problem. Links in the ocean and land transport networks represent alternative delivery routes and links in the port network express different port service processes such as loading, unloading, moving, and storing. The ocean transport network includes port-entrance links to access port terminals and the land transport network contains port-exit links to egress from port terminals. Figure 1 depicts a general carrier network and Figure 2 shows samples of port networks.

*3.2. Assumptions.* Basic assumptions for the definiteness and simplicity of modeling are presented as follows:

(1) Freight is shipped from a departure port terminal via an arrival port terminal to a land destination, assuming that commodities are produced near the departure port terminal.

(2) Ocean carriers choose port terminals of the alternative ones located in competitive regions.

(3) Multiple ocean carriers and port terminal operators have independent operations without cooperation in the competitive market environment.

*3.3. Modeling Approach.* Ocean carrier objective function is a cost minimization function. Ocean carriers departing a port terminal choose one of the arrival port terminals by considering the operating cost and the travel time on the transportation network as well as other determinants such as the port location and the port service charge. In particular, this research considers the emission cost in order to analyze the environmental effect on making ocean route choices. The port terminal operator objective function is a profit maximization function. Port terminal operators suggest competitive and reasonable service charges to attract more ocean liners and they simultaneously aim to minimize port service costs in the port complex. Therefore, ocean carriers are the leaders and port terminal operators are the followers.

At the first level, the ocean carrier problem is formulated using user equilibrium. User equilibrium finds the optimal ocean routes for which each ocean carrier obtains the lowest transportation in the competitive market environment. At the second level, the port terminal operator problem is formulated using Nash equilibrium. Nash equilibrium finds the optimal service charge and port service pattern for which each port terminal operator obtains the greatest profit in the competitive market environment.

Hierarchical interactions between ocean carriers and port terminal operators are captured employing a bilevel optimization programming. The concept of multi-leader-follower game is applied to a bilevel game, assuming the multiple and competitive leaders. Port service charges affect ocean carriers' routes, while port throughputs are influenced by ocean carriers' routing decisions.

*3.4. Emission Cost.* In this study, an activity-based maritime emission is employed. The emission cost for sailing between marine ports is calculated based on the vessel speed and sailing time. The total amount of emissions is estimated first and then transformed into money value by applying the unit emission cost according to the type of air pollutant.

For the main engine, emissions were estimated by using the following generic equation [20]:

$$\mathrm{E} = P \times \mathrm{LF} \times A \times \mathrm{EF}, \tag{1}$$

where E is emission from engine (g), $P$ is engine power (kW), LF is engine load factor, $A$ is ship activity (h), and EF is emission factor (g/kW-h).

Data on engine power ratios, engine load power, and emission factor were derived from literature. Table 1 illustrates main and auxiliary engine power ratios for six types of ships, ranging from 0.191 to 2.789. The main engine power ratio of the container is 0.220, showing a relatively small number. Load factors for four types of ship operations are defined in Table 1.

Emission factors are different according to the air pollutant and sailing speed. Emission factors for three speed categories (high, medium, and low) and three major air pollutants ($NO_x$, $SO_2$, and $PM_{10}$) are shown in Table 2. The emission of $NO_x$ is the greatest, while $PM_{10}$ is the smallest, indicating a significant difference. The speed change influences both $NO_x$ and $PM_{10}$, although the direction of the correlation is opposite.

Meanwhile, activity time implies the sailing duration for ships on major routes. Activity time data can be created by obtaining vessel speed profiles and distances between marine ports.

## 4. Mathematical Formulations

*4.1. Ocean Carrier Problem.* Ocean carriers provide transportation services between marine ports via waterways. They compete with others to choose marine terminals in a cost-efficient manner.

*4.1.1. Equilibrium Condition.* User equilibrium is used to find the optimal routing pattern for which each ocean carrier obtains the lowest transportation cost. Each user noncooperatively attempts to minimize his cost of transportation. No ocean carrier may lower his transportation cost by having unilateral action in equilibrium:

$$\mathrm{TC}_w^o\left(f_w^{o*}\right) \le \mathrm{TC}_w^o\left(f_w^o\right) \quad \forall o \in O,\ w \in W. \tag{2}$$

Equation (2) implies that if ocean carrier $o$ for $O$-$D$ pair $w$ changes his routing choices from optimal to any other feasible decisions when other ocean carriers keep their routes constant, the transportation cost will increase.

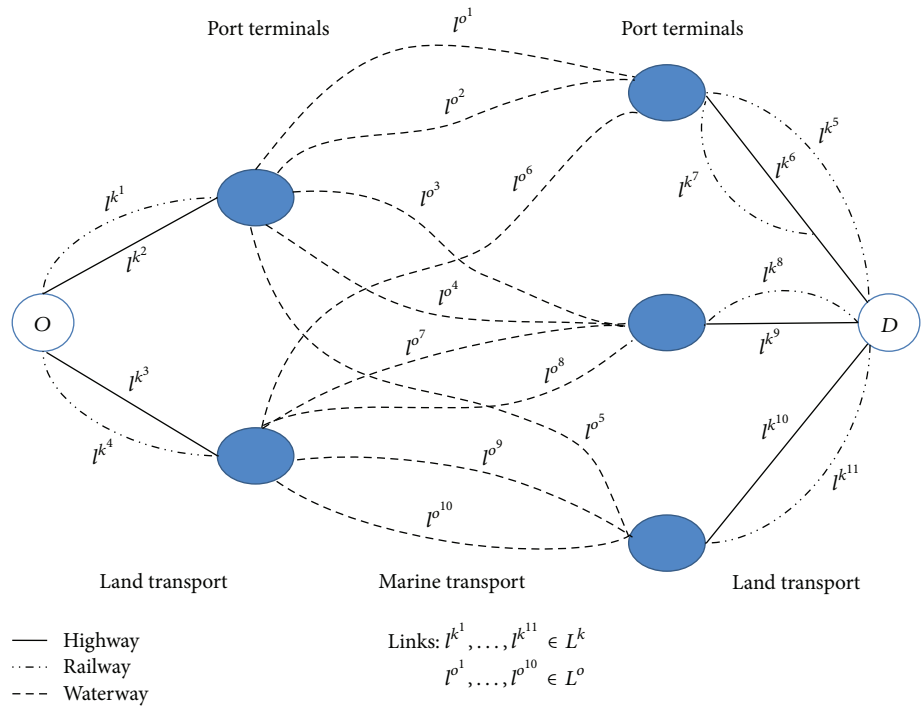



Figure 1: Carrier network.

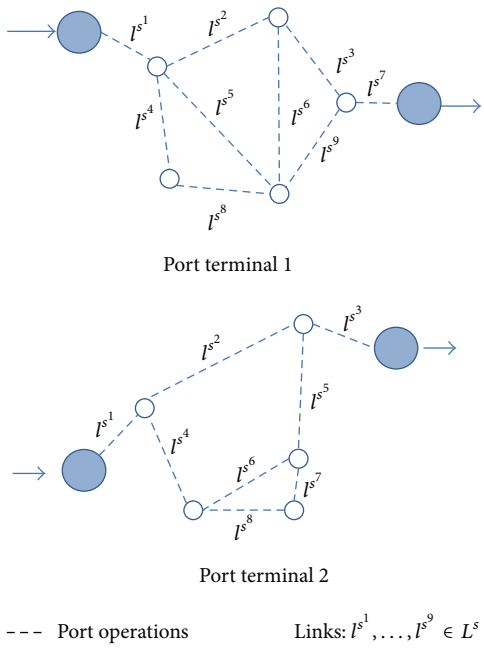



Figure 2: Port network (samples).

*4.1.2. Objective Function.* The objective function is a cost minimization function. Each ocean carrier aims to minimize the transportation cost by maximizing the revenue and minimizing the transportation cost:

$$\min \quad \mathrm{TC}_w^o = \sum_{p^o}^{P^o} \sum_{p^o}^{P^o} \mathrm{TC}_w^{p^o}\left(f_w^{p^o}\right) f_w^{p^o} \quad \forall o \in O,\ w \in W. \tag{3}$$

Equation (3) represents the transportation cost function decided by the transportation cost function depending on the flow on all used ocean paths.

*4.1.3. Assumptions for the Objective Function.* The ocean path transportation cost function is a linear combination of the following three attributes: (a) the sum of ocean link transportation cost functions if the ocean link is on the path, (b) the sum of emission costs, and (c) the sum of port service charges if ocean carrier $o$ chooses port terminal operator $s$. The emission cost on the ocean path is calculated based on the vessel speed (Tables 1 and 2) and sailing time (obtained from vessel speed profiles and distances marine ports) according to the type of air pollutant:

$$\begin{aligned} \mathrm{TC}_w^{p^o}\left(f_w^{p^o}\right) = {} & \sum_{l^o}^{L^o} \mu_w^{l^o,p^o} \mathrm{TC}^{l^o}\left(f^{l^o}\right) + \mathrm{EC}_w^{p^o} f_w^{p^o} \\ & + \pi^{o,s} c^s f_w^{p^o} \quad \forall w \in W,\ p^o \in P_w^o,\ s \in S. \end{aligned} \tag{4}$$

If the ocean link is on the ocean path for $O$-$D$ pair $w$, $\mu_w^{l^o,p^o}$ is 1; otherwise, it is 0. If ocean carrier $o$ chooses port terminal operator $s$, $\pi^{o,s}$ is 1; otherwise, it is 0. The ocean link transportation cost function comprises (a) the average ocean link operating cost function and (b) the average ocean link

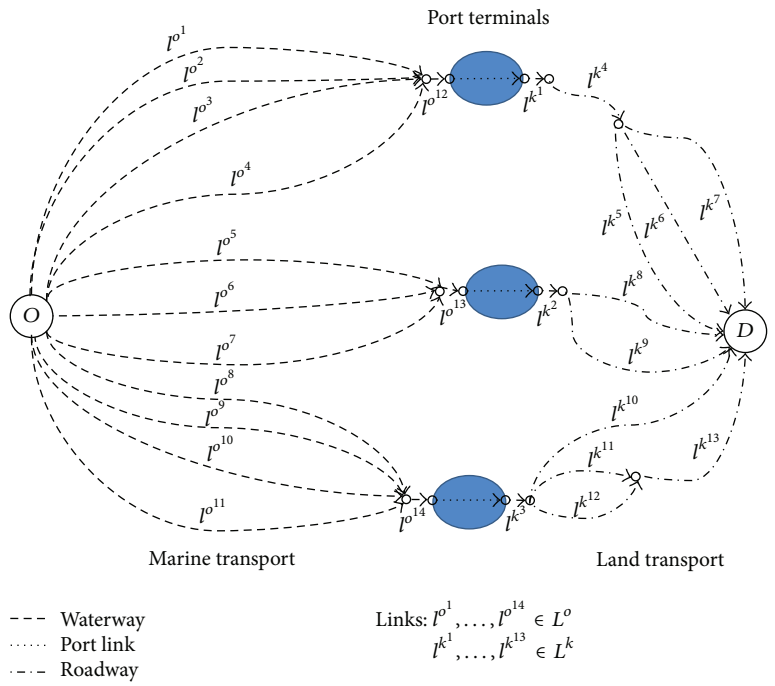



Figure 3: Carrier network in the example.

travel (shift) time function. The travel (shift) time is changed to money using value of time:

$$\mathrm{TC}^{l^o}\left(f^{l^o}\right) = \mathrm{OC}^{l^o}\left(f^{l^o}\right) f^{l^o} + \mathrm{vot}^o \mathrm{TT}^{l^o}\left(f^{l^o}\right) f^{l^o} \quad \forall l^o \in L^o, \tag{5}$$

where

$$\mathrm{OC}^{l^o}\left(f^{l^o}\right) = \omega_o + \omega_1 \left[\frac{f^{l^o}}{\mathrm{cap}^{l^o}}\right] + \omega_2 \left[\frac{f^{l^o}}{\mathrm{cap}^{l^o}}\right]^2 \quad \forall l^o \in L^o, \tag{6}$$

$$\mathrm{TT}^{l^o}\left(f^{l^o}\right) = \mathrm{tt}^{l^o} + \mathrm{st}^{l^o}\left[1 + \varepsilon_o\left(\frac{f^{l^o}}{\mathrm{cap}^{l^o}}\right)\right]^{\varepsilon_1} \quad \forall l^o \in L^o.$$

The average operating cost and travel (shift) time functions are assumed to be continuous and strictly monotone increasing in the flow.

*4.1.4. Feasible Region of the Objective Function.* The feasible region of the ocean carrier objective function (OFR) is defined by linear equality and nonnegativity constraints:

$$\sum_{o}^{O} d_w^o = \sum_{p^o}^{P^o} f_w^{p^o} \quad \forall w \in W, \tag{7}$$

$$f_w^{p^o} \geq 0 \quad \forall o \in O,\ w \in W. \tag{8}$$

Equation (7) ensures that the sum of individual ocean carrier service demands for *O-D* pair $w$ is equivalent to the sum of ocean carrier flows on all used ocean paths. These linear equality constraints define a closed and convex feasible region. Equation (8) states nonnegativity of ocean container flow.

*4.1.5. Mathematical Formulation.* For the ocean carrier problem, the research adopts a container transport network equilibrium model proposed by Zhang et al. [3]. As a modeling methodology, variational inequality by Nagurney [21] was adopted:

$$\mathrm{TC}_w^{p^o *} - \mathrm{MC}_w^* = \begin{cases} = 0 & f_w^{p^o *} > 0, \\ \geq 0 & f_w^{p^o *} = 0 \end{cases} \quad \forall p^o \in P_w^o,\ \forall w \in W, \tag{9}$$

where $\mathrm{MC}_w^*$ denotes the minimum path cost between *O-D* pair $\omega$ in equilibrium, $f_w^{p^o *}$ represents the container volume of ocean path $p$ in equilibrium, and $\mathrm{TC}_w^{p^o *}$ represents the total transportation cost in equilibrium.

**Theorem 1.** *The necessary and sufficient condition of $f^*$ as the equilibrium solution of (9) is that $f^*$ is the solution of the following variation inequality:*

$$\sum_{w}^{W} \sum_{p^o}^{P_w^o} \left(TC_w^{p^o *} - MC_w^*\right)\left(f_w^{p^o} - f_w^{p^o *}\right) \geq 0. \tag{10}$$

*Proof.* First, verify sufficiency.

$\forall w \in W$, $p^o \in P_w^o$, if $f_w^{p^o *} = 0$, then, based on (9), we have the following: $\forall f_w^{p^o} \geq 0$, $(\mathrm{TC}_w^{p^o *} - \mathrm{MC}_w^*)(f_w^{p^o} - f_w^{p^o *}) =$

$(\mathrm{TC}_w^{p^o*} - \mathrm{MC}_w^*) f_w^{p^o} \geq 0$; if $f_w^{p^o*} > 0$, according to (9), we have $\mathrm{TC}_w^{p^o*} - \mathrm{MC}_w^* = 0$; $\forall f_w^{p^o} \geq 0$, we have $(\mathrm{TC}_w^{p^o*} - \mathrm{MC}_w^*)(f_w^{p^o} - f_w^{p^o*}) = 0$. So $(\mathrm{TC}_w^{p^o*} - \mathrm{MC}_w^*)(f_w^{p^o} - f_w^{p^o*}) \geq 0$. At the same time, $\forall w \in W$, $p^o \in P_w^o$, $(\mathrm{TC}_w^{p^o*} - \mathrm{MC}_w^*)(f_w^{p^o} - f_w^{p^o*}) \geq 0$, so $\sum_w^W \sum_{p^o}^{P_w^o} (\mathrm{TC}_w^{p^o*} - \mathrm{MC}_w^*)(f_w^{p^o} - f_w^{p^o*}) \geq 0$.

Second, verify necessity.

Taking $k \in W$, $d \in P_k$ and $\forall w \in W \setminus k$, $p^o \in P_w^o$, $p^o \in p_k^o \backslash d$, let $f_w^{p^o} - f_w^{p^o*}$; then (10) is changed to $(\mathrm{TC}_k^{d*} - \mathrm{MC}_k^*)(f_k^d - f_k^{d*}) \geq 0 \forall f_k^d \geq 0$, so, for all $k \in W$, $p \in P_\omega$, if $f_k^{d*} = 0$, then $\mathrm{TC}_k^{d*} - \mathrm{MC}_k^* \geq 0$; if $f_k^{d*} > 0$, then $\mathrm{TC}_k^{d*} - \mathrm{MC}_k^* = 0$. Obviously, (10) is proved.

Equation (9) can be transformed into the following formula:

$$\sum_w^W \sum_{p^o}^{P_w^o} \mathrm{TC}_w^{p^o*} \left(f_w^{p^o} - f_w^{p^o*}\right) - \sum_w^W \sum_{p^o}^{P_w^o} \mathrm{MC}_w^* \left(f_w^{p^o} - f_w^{p^o*}\right) \geq 0. \tag{11}$$

According to flow conservation law, the above formula is changed to

$$\sum_w^W \sum_{p^o}^{P_w^o} \mathrm{TC}_w^{p^o*} \left(f_w^{p^o} - f_w^{p^o*}\right) - \sum_w^W \sum_{p^o}^{P_w^o} \mathrm{MC}_w^* \left(d_w^{p^o} - d_w^{p^o*}\right) \geq 0. \tag{12}$$

The vector form is

$$\mathrm{TC}\left(f^{o*}\right)^T \left(f^o - f^{o*}\right) - M^{*T}\left(g - g^*\right) \geq 0. \tag{13}$$

Assuming that the demand of *O-D* is fixed, (13) becomes $\mathrm{TC}(f^{o*})^T(f^o - f^{o*}) \geq 0$.

According to the relation between transportation arc and transportation path, the above variational inequality is expressed as

$$\mathrm{TC}\left(f^{o*}\right)^T \left(f^o - f^{o*}\right) \geq 0. \tag{14}$$

□

*4.2. Port Terminal Operator Problem.* Port terminal operators provide the transportation service for freight arriving in the port complex. The port service charge affects ocean carriers' routes, while port throughputs are influenced by ocean carriers' routing decisions.

*4.2.1. Equilibrium Condition.* Nash equilibrium is used to find the optimal port service charge and pattern for which each port terminal operator obtains the greatest profit. No port terminal operator can be better off by changing its decisions unilaterally in equilibrium. Consider

$$U^s\left(c^{s*}, c^{-s*}, f^{s*}\right) \geq U^s\left(c^s, c^{-s*}, f^s\right) \quad \forall s \in S. \tag{15}$$

*4.2.2. Objective Function.* The objective function is a profit maximization function. Each port terminal operator aims to maximize the profit by maximizing the port revenue and minimizing the port service cost. Consider

$$\max \quad U^s = \underbrace{G^s\left(c^s, \overline{c^{-s}}\right) c^s}_{R^s} - \underbrace{\sum_v^V \sum_{p^s}^{P^s} \mathrm{SC}_v^{p^s}\left(f_v^{p^s}\right) f_v^{p^s}}_{\mathrm{SC}^s} \quad \forall s \in S. \tag{16}$$

Equation (16) expresses that the profit function of port terminal operator $s$ is composed of the revenue function and the port service cost function. The port terminal operator $s$ is determined by the port throughput depending on the service charge, while other terminal operators keep their current level of service charges constant. The port service cost is decided by the port service cost function depending on the flow on all used port paths.

*4.2.3. Assumptions for the Objective Function.* The port throughput function of port terminal operator $s$ is assumed to be linear and strictly monotone decreasing in the port service charge:

$$R^s = \beta_0 - \beta_1 c^s \quad \forall s \in S. \tag{17}$$

The port path service cost function is the sum of port link service cost functions if the port link is on the path:

$$\mathrm{SC}_v^{p^s}\left(f_v^{p^s}\right) = \sum_{l^s}^{L^s} \varphi_v^{l^s, p^s} \mathrm{SC}^{l^s}\left(f^{l^s}\right) \quad \forall v \in V,\ p^s \in P^s. \tag{18}$$

The port link service cost function includes the following two attributes: (a) the average port link operating cost function and (b) the average port link service time function. These functions are assumed to be continuous and strictly monotone increasing in the flow. Consider

$$\mathrm{SC}^{l^s}\left(l^s\right) = \mathrm{SC}^{l^s}\left(f^{l^s}\right) f^{l^s} + \mathrm{vot}^s \mathrm{ST}^{l^s}\left(f^{l^s}\right) f^{l^s} \quad \forall l^s \in L^s. \tag{19}$$

Hessian matrix (second derivatives of the objective function) is negative definite and symmetric. Hence, the objective function is strictly concave in $(c^{s*}, f_v^{p^{s*}})$.

*4.2.4. Feasible Region of the Objective Function.* The feasible region of the port terminal operator objective function (PFR) is defined by linear equality and nonnegativity constraints. Consider

$$\sum_w^W \sum_{p^o}^{P_w^o} \pi_w^{p^o, s} f_w^{p^o} = G^s = \sum_v^V \sum_{p^s}^{P^s} f_v^{p^s} \quad \forall s \in S, \tag{20}$$

$$Sc^s < c^s < Bc^s \quad \forall s \in S, \tag{21}$$

$$v^s, f_v^{p^s} \geq 0 \quad \forall s \in S,\ v \in V,\ p^s \in P^s. \tag{22}$$

Equation (20) ensures that the total amount of freight transported via ocean paths is equivalent to the throughput of the port terminal operator $s$ when ocean paths are connected to the port terminal. Also, the throughput is equivalent to the sum of port flows on all used paths in a port complex. Equation (21) ensures that the port service charge ranges from a small number to a large number. Equation (22) states nonnegativity of the port service charge and container flow.

*4.2.5. Mathematical Formulation.* Based upon the objective function properties and feasible region, a port terminal operator model is formulated by a variational inequality problem. In the competitive game, port terminal operator $s$ finds the optimal service charge and port service pattern to obtain the maximum profit. For a new port service charge, the port revenue and the port service cost are updated to compare the profit. Each port terminal operator attempts to minimize the total port service cost. Thus, port terminal operator $s$ exhibits a system equilibrium-like behavior with the marginal transportation cost function. A port terminal operator model is expressed as follows:

$$-\sum_{s}^{S}\left(G^{s}\left(c^{s*},\overline{c^{-s*}}\right)c^{s*}-\mathrm{SC}^{s}\right)\left(c^{s}-c^{s*}\right)\geq 0,$$
$$\text{where }\sum_{v}^{V}\sum_{P^{s}}^{P^{s}}\nabla_{f}\mathrm{SC}_{v}^{p^{s}}\left(f_{v}^{p^{s}*}\right)\left(f_{v}^{p^{s}}-f_{v}^{p^{s}*}\right)\geq 0, \tag{23}$$
$$\forall\left(c^{s},f^{s}\right)\in \mathrm{PFR}.$$

The vector form is

$$-\left(G^{s}\left(c^{s*},\overline{c^{-s*}}\right)c^{s}-\mathrm{SC}^{s}\right)^{T}\left(c^{s}-c^{s*}\right)\geq 0,$$
$$\text{where }\nabla\mathrm{SC}\left(f^{s}\right)^{T}\left(f^{s}-f^{s*}\right)\geq 0. \tag{24}$$

*4.2.6. Existence and Uniqueness of the Solution.* The formulations above are continuous and linear equality constraints of PFR define a closed convex feasible region. Hence, the existence of the solution is proved. The port terminal operator game has a solution when $c^{s*}$, $f^{s*}$ satisfy $\|c^{s*}, f^{s*}\| < R$. The service charge $(c^{s})$ is bounded in (21); therefore the port throughput $(G^{s})$ is also bounded. With the bounded $G^{s}$, the flow $(f^{s})$ is bounded due to the flow conservation constraint in (20). The formulations are strictly monotone in $(c^{s}, f^{s})$ from the concave properties of the objective function. Therefore, the solution is unique.

*4.3. Bilevel Model.* A bilevel model is formulated to capture hierarchical interactions of ocean carriers and port terminal operators when ocean carriers are the leaders and port terminal operators are the followers in the maritime freight transport system. Equation (25) shows a vector form of the bilevel model as follows:

$$\mathrm{TC}\left(f^{o*}\right)^{T}\left(f^{o}-f^{o*}\right)$$
$$\text{s.t. }-\left(G^{s}\left(c^{s*},\overline{c^{-s*}}\right)c^{s}-\mathrm{SC}^{s}\right)^{T}\left(c^{s}-c^{s*}\right)\geq 0, \tag{25}$$
$$\text{where }\nabla\mathrm{SC}\left(f^{s}\right)^{T}\left(f^{s}-f^{s*}\right)\geq 0,\ \forall\left(c^{s},f^{s}\right)\in \mathrm{PFR}.$$

## 5. Solution Algorithms

Algorithms are developed to solve the bilevel model using the extra gradient algorithm, an improved version of the projection algorithm, which resolves a variational inequality problem into a sequence of subproblems by the general iterative scheme. Ocean carriers examine the reactions of port terminal operators for feasible scenarios and finally choose the best one which gives the minimum transportation cost. Individual carrier problems are solved in the following steps.

*5.1. Ocean Carrier Problem*

*Step 0.* Define the initial ocean path flow $f_{w}^{p^{o0}} = f_{w}^{p1^{o}} + f_{w}^{p2^{o}} + f_{w}^{p3^{o}} + \cdots + f_{w}^{pm^{o}}$ for *O-D* pair $w$. Let $u$ denote the order of ocean path $f_{w}^{p^{o}}$. Let $z$ denote the order of the iterations. Set $u := 1$, $z := 1$, and $\sigma_1$ is a preset tolerance.

*Step 1.* Determine $f_{w}^{p^{o}}$ for $u$th ocean path. Then, decide the next ocean path flow by setting $u := u + 1$ until $u < m$. If $|f_{w}^{p1,z} - f_{w}^{p1,z-1}| < \sigma_1$, $|f_{w}^{p2,z} - f_{w}^{p2,z-1}| < \sigma_1, \ldots, |f_{w}^{p3,z} - f_{w}^{p3,z-1}| < \sigma_1$, stop and denote the solution as $(f_{w}^{pm^{o}*})$.

*5.2. Port Terminal Operator Problem*

*Step 0.* Define the initial port terminal operator service charge $c^{s0} = c^{s1}, c^{s2}, \ldots, c^{sn}$ and flow $f_{w}^{p^{s}} = f_{w}^{p1^{s1}} + f_{v}^{p1^{s2}} + f_{v}^{p1^{s3}} + \cdots + f_{v}^{pm^{sn}}$. Let $g$ denote the order of port terminal operator $s$ and let $u$ denote the order of port path $f_{v}^{p^{s}}$. Let $x$ and $z$ denote the order of the iterations, respectively. Set $g := 1$, $u := 1$, $x := 1$, and $z := 1$. $\sigma_2$ and $\sigma_3$ are preset tolerances.

*Step 1.* Determine $c^{s}$ for $g$th port terminal operator with $\overline{c^{-s}} = c^{-s,x-1}$ for the port service demand determined by ocean carriers' routing decisions.

*Step 2.* Determine $f_{v}^{p^{s}}$ for $u$th port path of $g$th port terminal operator. Then, decide the next port path flow by setting $u := u + 1$ until $u < m$. If $|f_{v}^{p1,z} - f_{v}^{p1,z-1}| < \sigma_2$, $|f_{v}^{p1,z} - f_{v}^{p1,z-1}| < \sigma_2, \ldots, |f_{v}^{pm,z} - f_{v}^{p1,z-1}| < \sigma_2$, stop and go to Step 3; otherwise, $z := z + 1$ and repeat Step 2.

*Step 3.* Solve the next port terminal operator problem via Steps 1 and 2, by setting $g := g + 1$ until $g < n$.

*Step 4*. If profit differences of port terminal operators are smaller than $\sigma_3$, stop and denote solutions as $c^{s*}$, $f_v^{p^{s*}}$; otherwise, set $x := x + 1$ and turn to Step 1.

## 6. Numerical Example

A multimodal carrier network is tested using MATLAB R2010a (Pentium IV 3.00 GHz CPU with 8.00 GB of RAM). The carrier network in the example has 45 links including 14 ocean links ($l^{o^1} \sim l^{o^{14}}$), 18 port links ($l^{s^1} \sim l^{s^{18}}$), and 13 land links ($l^{k^1} \sim l^{k^{13}}$). Individual carrier groups provide transportation services at different parts of the carrier network. Freight is shipped from a departure port terminal via an arrival port terminal to a destination, assuming that commodities are produced near the departure port terminal. Ocean carriers that depart a port terminal choose arrival port terminals near a final inland destination by considering the transportation cost, the port location, and the port service cost. Three alternative *O-D* pairs between marine ports exist in this example. Three port terminal operators located in a competitive region provide port transportation services in the port complex separately. Links in the port network express different port service processes. Land carriers deliver freight from three port terminals to a destination via trucks and rails. Land carrier problem is beyond the study scope. Each port terminal has its own subnetwork according to the port layout and service process. The carrier network is shown in Figure 3.

The total maritime transportation demand is assumed to be 150,500 TEUs during a time period. Multiple ocean carrier companies aim to minimize their transportation costs. In particular, the research analyzes the effects of the emission cost on carrier route choices as seriousness of air pollution and the global concern in the environmental problem increase rapidly. The emissions on ocean path were estimated by multiplying emission components like engine, engine power, engine load factor, ship activity, and emission factor for the three main pollutants of $NO_x$, $SO_2$, and $PM_{10}$ and transformed into money value by applying the unit emission cost. Input data were obtained from various references. In order to estimate the emission cost of each container, the maximum ship capacity is assumed to be 15,000 TEUs. Emission costs are calculated to be \$232.8~\$241.6 on ocean links. Table 3 shows parameters in the ocean link operating cost and travel (shift) time functions.

Port terminal operators compete with each other to attract more ocean carriers and attempt to minimize the port service cost with a system equilibrium-like behavior. Port service charge ranges from \$300 to \$400 per a unit of freight due to ocean shipping market circumstances. Table 4 shows parameters in the port link operating cost and service time functions. The land transportation cost is assumed to be given in the example.

Table 5 shows container flow changes in ocean links without/with considering the emission costs in the ocean transportation cost. The result indicates how the optimal ocean route decisions are changed by adding the emission costs. The transportation cost at any ocean path is equivalent at \$609 and \$848, respectively.

Port throughputs are determined by ocean carriers' routing decisions. Based on the relationships, port profits were also changed by the emission costs. Port terminal operator 2 experienced profit growth, while opposite patterns occur for port terminals 1 and 3. Table 6 shows profit changes of port terminal operators.

The container flow is influenced by port throughputs since port terminal operators attempt to minimize the total port service cost in the current port infrastructure. Table 7 illustrates container flow changes in port links at each port terminal. The marginal transportation cost at any ocean path is equivalent to \$206 and \$205 at port terminal 1, \$217 and \$221 at port terminal 2, and \$210 and \$205 at port terminal 3, respectively.

## 7. Conclusions

The paper formulated a freight network model that captures decisions and interactions of ocean carriers and port terminal operators in the maritime freight transport system. Their hierarchical interactions were captured using a bilevel modeling approach. Numerical examples demonstrated the validity and applicability of the bilevel model.

Today, air pollution from international ships is a serious and urgent problem from perspectives of environmental destruction and health harm factors. Therefore, this study analyzed the effects of the environmental cost on ocean route choices by considering the emission cost of international container ships in addition to the general ocean transportation cost components. As a result, the optimal ocean routes and transportation costs were changed distinctively due to the different equilibrium conditions caused by the emission cost, compared with the case that does not consider the cost. Based on the relationships between ocean carriers and port terminal operators, port throughputs and operations were also changed accordingly. The results indicate that the emissions cost affects ocean routes directly and ocean shipping companies have to reflect this cost when making routing decisions to improve the current and future environment problems. The companies or relevant stakeholders can use the emission cost proposed in (1) and regard the emission cost as a key cost for calculating the transportation cost. The simulation results considering this new cost can provide new optimal strategies and reacting responses of competitors.

This study contributes to a freight network modeling approach considering the environmental cost, showing how the emission cost plays a role in ocean route changes. The research may be extended to a case with a different sum of the container flow and a more comprehensive study that finds the behavior of other stakeholders such as shippers and public bodies (i.e., Port Authorities) at the same time. Although this paper concentrated on the emissions of container ships, other types of ships like general cargo ships, barges, tankers, and so forth can be considered in future research. Also, the emission cost at marine ports and inland areas can be combined to analyze the impacts of emission costs extensively.

Table 1: Engine power ratios and load factor assumptions.

| Ship type | AE to ME power ratio | Load factor | | | |
|---|---|---|---|---|---|
| | | Cruise | RSZ | Maneuver | Hotel |
| Bulker/general cargo | 0.222 or 0.191 | 0.17 | 0.27 | 0.45 | 0.22 |
| Container | 0.220 | 0.13 | 0.25 | 0.50 | 0.17 |
| Passenger/ferry | 0.278 | 0.80 | 0.80 | 0.80 | 0.64 |
| Ro-Ro | 0.259 | 0.15 | 0.30 | 0.45 | 0.30 |
| Tanker | 0.211 | 0.13 | 0.27 | 0.45 | 0.67 |
| Miscellaneous | 0.257 | 0.17 | 0.27 | 0.45 | 0.22 |

Sources: Work Shipping Encyclopedia, USEPA [20], and Yau et al. [22].

Table 2: Emission factors by sailing speed.

| Engine type | Year built | $NO_x$ | $SO_2$ | $PM_{10}$ |
|---|---|---|---|---|
| | | Main engine (g/kW h) | | |
| SSD | Before 2000 | 18.1 | 10.16 | 1.05 |
| SSD | After 2000 | 17 | 10.16 | 1.05 |
| MSD | Before 2000 | 14 | 10.16 | 1.11 |
| MSD | After 2000 | 13 | 10.16 | 1.11 |
| HSD | | 12.7 | 10.16 | 1.11 |

Note: SSD: slow speed diesel; MSD: medium speed diesel; HSD: high speed diesel.
Sources: Entec [23], Lloyd's Register [24], USEPA [20], and Yau et al. [22].

## Appendix

See Tables 3 and 4.

## Notations

*Indices*

*(i) Ocean Carrier*

$O$: Set of ocean carriers
$W$: Set of ocean carrier $O$-$D$ pairs
$L^o$: Set of ocean links
$P_w^o$: Set of ocean paths for $O$-$D$ pair $w$
$o$: Ocean carrier ($o \in O$)
$w$: Ocean carrier $O$-$D$ pair ($w \in W$)
$l^o$: Ocean link ($l^o \in L^o$)
$p^o$: Ocean path ($p^o \in P_w^o$)

*(ii) Port Terminal Operator*

$S$: Set of port terminal operators
$V$: Set of port terminal $O$-$D$ pairs
$L^s$: Set of port links
$P_v^s$: Set of port paths
$s$: Port terminal operator ($s \in S$)
$v$: Port terminal $O$-$D$ pair ($v \in V$)
$l^s$: Port link ($l^s \in L^s$)
$p^s$: Port path ($p^s \in P_v^s$)

*Data and Parameters*

$d_w^o$: Transportation demand for $O$-$D$ pair $w$
$EC_w^{p^o}$: Emission cost on the ocean path $p^o$ connecting $O$-$D$ pair $w$
$vot^o$: Value of time of ocean carrier $o$
$vot^s$: Value of time of port terminal operator $s$
$tt^{l^o}$: Average travel time on the ocean link $l^o$
$st^{l^o}$: Average shift time on the ocean link $l^o$
$st^{l^s}$: Average service time on the port link $l^s$
$cap^{l^o}$: Capacity of the ocean link $l^o$
$cap^{l^s}$: Capacity of the port link $l^s$

*Variables*

*(i) Ocean Carrier*

Table 3: Parameters in the transpiration cost functions.

| Link | $\omega_o$ | $\omega_1$ | $\omega_2$ | tt | st | $\varepsilon_o$ | $\varepsilon_1$ | Capacity |
|---|---|---|---|---|---|---|---|---|
| Ocean link 1 ($l^{o^1}$) | 180 | 2.2 | — | 17.8 | — | — | — | 30,000 |
| Ocean link 2 ($l^{o^2}$) | 185 | 2.2 | — | 17.0 | — | — | — | 30,000 |
| Ocean link 3 ($l^{o^3}$) | 184 | 2.3 | — | 16.8 | — | — | — | 30,000 |
| Ocean link 4 ($l^{o^4}$) | 188 | 2.8 | — | 16.2 | — | — | — | 30,000 |
| Ocean link 5 ($l^{o^5}$) | 179 | 2.9 | — | 15.5 | — | — | — | 30,000 |
| Ocean link 6 ($l^{o^6}$) | 175 | 3.1 | — | 15.9 | — | — | — | 30,000 |
| Ocean link 7 ($l^{o^7}$) | 181 | 2.9 | — | 15.2 | — | — | — | 30,000 |
| Ocean link 8 ($l^{o^8}$) | 183 | 3.2 | — | 16.3 | — | — | — | 30,000 |
| Ocean link 9 ($l^{o^9}$) | 186 | 2.6 | — | 16.2 | — | — | — | 30,000 |
| Ocean link 10 ($l^{o^{10}}$) | 182 | 2.4 | — | 17.1 | — | — | — | 30,000 |
| Ocean link 11 ($l^{o^{11}}$) | 184 | 2.2 | — | 16.8 | — | — | — | 30,000 |
| Ocean link 12 ($l^{o^{12}}$) | 28 | 3.5 | 2.5 | — | 0.3 | 3.8 | 2 | 60,000 |
| Ocean link 13 ($l^{o^{13}}$) | 36 | 3.7 | 2.75 | — | 0.8 | 4.2 | 2 | 60,000 |
| Ocean link 14 ($l^{o^{14}}$) | 32 | 3.3 | 2.4 | — | 0.2 | 3.8 | 2 | 60,000 |

Table 4: Parameters in the average port service cost function.

| Link | $\gamma_o$ | $\gamma_1$ | $\gamma_2$ | st | $\lambda_o$ | $\lambda_2$ | Capacity |
|---|---|---|---|---|---|---|---|
| Port link 1 ($l^{o^1}$) | 78 | 2.7 | 0.85 | 4.5 | 1.2 | 2 | 100,000 |
| Port link 2 ($l^{o^2}$) | 102 | 2.1 | 0.78 | 3.8 | 2.2 | 2 | 100,000 |
| Port link 3 ($l^{o^3}$) | 68 | 3.1 | 0.65 | 4.2 | 2.1 | 2 | 100,000 |
| Port link 4 ($l^{o^4}$) | 26 | 2.3 | 0.55 | 8.0 | 1.7 | 2 | 100,000 |
| Port link 5 ($l^{o^5}$) | 56 | 1.4 | 0.58 | 3.8 | 2.3 | 2 | 100,000 |
| Port link 6 ($l^{o^6}$) | 52 | 1.7 | 0.74 | 2.8 | 1.6 | 2 | 100,000 |
| Port link 7 ($l^{o^7}$) | 112 | 2.2 | 0.75 | 5.6 | 2.2 | 2 | 100,000 |
| Port link 8 ($l^{o^8}$) | 42 | 2.6 | 0.62 | 3.8 | 2.4 | 2 | 100,000 |
| Port link 9 ($l^{o^9}$) | 32 | 3.2 | 0.63 | 3.2 | 2.6 | 2 | 100,000 |
| Port link 10 ($l^{o^{10}}$) | 29 | 2.2 | 0.82 | 2.3 | 1.2 | 2 | 100,000 |
| Port link 11 ($l^{o^{11}}$) | 74 | 1.9 | 0.53 | 3.5 | 1.8 | 2 | 100,000 |
| Port link 12 ($l^{o^{12}}$) | 120 | 2.2 | 0.87 | 7.2 | 3.2 | 2 | 100,000 |
| Port link 13 ($l^{o^{13}}$) | 63 | 1.6 | 0.47 | 6.2 | 3.4 | 2 | 100,000 |
| Port link 14 ($l^{o^{14}}$) | 88 | 2.6 | 0.85 | 4.9 | 3.6 | 2 | 100,000 |
| Port link 15 ($l^{o^{15}}$) | 95 | 1.8 | 0.72 | 6.3 | 2.2 | 2 | 100,000 |
| Port link 16 ($l^{o^{16}}$) | 43 | 0.9 | 0.62 | 3.4 | 2.8 | 2 | 100,000 |
| Port link 17 ($l^{o^{17}}$) | 59 | 1.2 | 0.58 | 4.6 | 3.1 | 2 | 100,000 |
| Port link 18 ($l^{o^{18}}$) | 85 | 2.1 | 0.77 | 5.3 | 3.5 | 2 | 100,000 |

$f_w^o$: Container volume of ocean carrier $o$ for $O$-$D$ pair $w$

$f_w^{p^o}$: Container volume on the ocean path $p^o$ connecting $O$-$D$ pair $w$

$f^{l^o}$: Container volume on the ocean link $l^o$

$\mu^{l^o,p^o}$: 1, if the ocean link $l^o$ is on the ocean path $p^o$; otherwise 0

$\pi^{o,p}$: 1, if the ocean carrier $o$ chooses port terminal operator $p$; otherwise 0

*(ii) Port Terminal Operator*

$c^s$: The service charge of port terminal operator $s$

$c^{-s}$: The service charges of other port terminal operators except port terminal operator $s$

$f_v^s$: Container volume of port terminal operators connecting $O$-$D$ pair $w$

$f_v^{p^s}$: Container volume on the port path $p^s$ connecting $O$-$D$ pair $v$

$f^{l^s}$: Container volume on the port link $l^s$

$\varphi^{l^s,p^s}$: 1, if the port link $l^s$ is on the port path $p^s$; otherwise 0

Table 5: Container flow changes in ocean links.

| Ocean link | Without emission cost | With emission cost | Difference ($A - B$) |
|---|---|---|---|
| Ocean link 1 ($l^{o^1}$) | 13,506 | 1,506 | 12,000 |
| Ocean link 2 ($l^{o^2}$) | 4,740 | 1,740 | 3,000 |
| Ocean link 3 ($l^{o^3}$) | 2,680 | 32,680 | −10,000 |
| Ocean link 4 ($l^{o^4}$) | 11,724 | 12,724 | −1,000 |
| Ocean link 5 ($l^{o^5}$) | 6,810 | 6,310 | 500 |
| Ocean link 6 ($l^{o^6}$) | 29,670 | 40,070 | −10,400 |
| Ocean link 7 ($l^{o^7}$) | 4,319 | 1,019 | 3,300 |
| Ocean link 8 ($l^{o^8}$) | 19,277 | 28,977 | −9,700 |
| Ocean link 9 ($l^{o^9}$) | 13,902 | 14,902 | −1,000 |
| Ocean link 10 ($l^{o^{10}}$) | 12,536 | 8,536 | 4,000 |
| Ocean link 11 ($l^{o^{11}}$) | 11,336 | 2,036 | 9,300 |
| Ocean link 12 ($l^{o^{12}}$) | 52,650 | 48,650 | 4,000 |
| Ocean link 13 ($l^{o^{13}}$) | 40,799 | 47,399 | −6,600 |
| Ocean link 14 ($l^{o^{14}}$) | 57,051 | 54,451 | 2,600 |
| Total | 150,500 | 150,500 | — |

Table 6: Profit changes of port terminal operators.

| Port terminal operator | Without | With | Difference ($A - B$) |
|---|---|---|---|
| Port terminal operator 1 | 15,795,000 | 14,595,000 | −1,200,000 |
| Port terminal operator 2 | 12,239,700 | 14,219,700 | 1,980,000 |
| Port terminal operator 3 | 17,115,300 | 16,335,300 | −780,000 |
| Total | 45,150,000 | 45,150,000 | — |

Table 7: Container flow changes in port links.

| | Ocean link | Without | With | Difference ($A - B$) |
|---|---|---|---|---|
| Port terminal 1 | Port link 1 ($l^{o^1}$) | 39,810 | 37,110 | 2,700 |
| | Port link 2 ($l^{o^2}$) | 14,260 | 12,560 | 1,700 |
| | Port link 3 ($l^{o^3}$) | 12,840 | 11,340 | 1,500 |
| | Port link 4 ($l^{o^4}$) | 25,550 | 24,550 | 1,000 |
| | Port link 5 ($l^{o^5}$) | 12,840 | 11,340 | 1,500 |
| | Port link 6 ($l^{o^6}$) | 38,390 | 35,890 | 2,500 |
| | *Total* | *52,650* | *48,650* | *4,000* |
| Port terminal 2 | Port link 7 ($l^{o^7}$) | 22,530 | 26,035 | −3,505 |
| | Port link 8 ($l^{o^8}$) | 18,269 | 21,364 | −3,095 |
| | Port link 9 ($l^{o^9}$) | 18,269 | 21,364 | −3,095 |
| | Port link 10 ($l^{o^{10}}$) | 18,269 | 21,364 | −3,095 |
| | Port link 11 ($l^{o^{11}}$) | 40,799 | 47,399 | −6,600 |
| | *Total* | *40,799* | *47,399* | *−6,600* |
| Port terminal 3 | Port link 12 ($l^{o^{12}}$) | 27,371 | 27,356 | 15 |
| | Port link 13 ($l^{o^{13}}$) | 27,371 | 27,356 | 15 |
| | Port link 14 ($l^{o^{14}}$) | 18,440 | 15,425 | 3,015 |
| | Port link 15 ($l^{o^{15}}$) | 18,440 | 15,425 | 3,015 |
| | Port link 16 ($l^{o^{16}}$) | 11,240 | 11,670 | −430 |
| | Port link 17 ($l^{o^{17}}$) | 11,240 | 11,670 | −430 |
| | Port link 18 ($l^{o^{18}}$) | 11,240 | 11,670 | −430 |
| | *Total* | *57,051* | *54,451* | *2,600* |
| Total | | 150,500 | 150,500 | — |

*Functions*

*(i) Ocean Carrier*

$TC_w^o$: Transportation cost function of ocean carrier for *O-D* pair $w$
$TC_w^{p^o}$: Transportation cost function of ocean path $p^o$ connecting *O-D* pair $w$
$EC_w^{p^o}$: Emission cost function of ocean path $p^o$ connecting *O-D* pair $w$
$TC^{l^o}$: Transportation cost function of ocean link $l^o$
$OC^{l^o}$: Operating cost function of ocean link $l^o$
$TT^{l^o}$: Travel (shift) time function of ocean link $l^o$

*(ii) Port Terminal Operator*

$U^s$: Profit function of port terminal operator $s$
$G^s$: Port throughput function of port terminal operator $s$

$SC^s$: Port service cost function of port terminal operator $s$
$SC_v^{p^s}$: Service cost function of the port path $p^s$ connecting *O-D* pair $v$
$SC^{l^s}$: Port service cost function of the port link $l^s$
$ST^{l^s}$: Port service time function of the port link $l^s$.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This work was supported by Incheon National University Research Grant in 2015.

## References

[1] Y. Zan, "Analysis of container port policy by the reaction of an equilibrium shipping market," *Maritime Policy and Management*, vol. 26, no. 4, pp. 369–381, 1999.

[2] K. Kuroda, M. Takebayashi, and T. Tsuji, "International container transportation network analysis considering post-panamax class container ships," *Research in Transportation Economics*, vol. 13, pp. 369–391, 2005.

[3] T. Zhang, Q. Zhao, and W. Wu, *Bi-Level Programming Model of Container Port Game in the Container Transport Supernetwork*, Korean Society for Computational and Applied Mathematics, 2008.

[4] H. Lee, M. Boile, and S. Theofanis, "Modeling carrier interactions in an international freight transport system," *International Journal of Information Systems and Supply Chain Management*, vol. 7, no. 1, pp. 15–39, 2014.

[5] H. Lee, M. Boile, S. Theofanis, and S. Choo, "Game theoretical models of the cooperative carrier behavior," *KSCE Journal of Civil Engineering*, vol. 18, no. 5, pp. 1528–1538, 2014.

[6] H. Lee, Y. Song, S. Choo, K.-Y. Chung, and K.-D. Lee, "Bi-level optimization programming for the shipper-carrier network problem," *Cluster Computing*, vol. 17, no. 3, pp. 805–816, 2014.

[7] Y. Wang, *A bi-level programming approach for the shipper-carrier network problem [Ph.D. thesis]*, New Jersey Institute of Technology, 2001.

[8] H. Min and Z. Guo, "Developing bi-level equilibrium models for the global container transportation network from the perspectives of multiple stakeholders," *International Journal of Logistics Systems and Management*, vol. 6, no. 4, pp. 362–379, 2010.

[9] X. Zhang, H. M. Zhang, H.-J. Huang, L. J. Sun, and T.-Q. Tang, "Competitive, cooperative and stackelberg congestion pricing for multiple regions in transportation networks," *Transportmetrica*, vol. 7, no. 4, pp. 297–320, 2011.

[10] G. S. Kim and T. S. Kim, "Functional networking of logistics port cities in Northeast Asia," *International Journal of Urban Sciences*, vol. 14, no. 1, pp. 73–85, 2011.

[11] L. Jiang and J. Kronbak, *The Model of Maritime External Costs', Work Package*, University of Southern Denmark, 2012.

[12] J. Berechman and P.-H. Tseng, "Estimating the environmental costs of port related emissions: the case of Kaohsiung," *Transportation Research Part D: Transport and Environment*, vol. 17, no. 1, pp. 35–38, 2012.

[13] P. S. Yau, S. C. Lee, Y. Cheng, Y. Huang, S. C. Lai, and X. H. Xu, "Contribution of ship emissions to the fine particulate in the community near an international port in Hong Kong," *Atmospheric Research*, vol. 124, pp. 61–72, 2013.

[14] H. Saraçoğlu, C. Deniz, and A. Kılıç, "An investigation on the effects of ship sourced emissions in Izmir port, Turkey," *The Scientific World Journal*, vol. 2013, Article ID 218324, 8 pages, 2013.

[15] C. Deniz, A. Kilic, and G. Civkaroglu, "Estimation of shipping emissions in Candarli Gulf, Turkey," *Environmental Monitoring and Assessment*, vol. 171, no. 1, pp. 219–228, 2010.

[16] J. J. Corbett, H. Wang, and J. J. Winebrake, "The effectiveness and costs of speed reductions on emissions from international shipping," *Transportation Research Part D: Transport and Environment*, vol. 14, no. 8, pp. 593–598, 2009.

[17] K. Ahn and H. Rakha, "The effects of route choice decisions on vehicle energy consumption and emissions," *Transportation Research Part D: Transport and Environment*, vol. 13, no. 3, pp. 151–167, 2008.

[18] D. Park, N. S. Kim, H. Park, and K. Kim, "Estimating trade-off among logistics cost, $CO_2$ and time: a case study of container transportation systems in Korea," *International Journal of Urban Sciences*, vol. 16, no. 1, pp. 85–98, 2012.

[19] A. Shukla and M. Alam, "Assessment of real world on-road vehicle emissions under dynamic urban traffic conditions in Delhi," *International Journal of Urban Sciences*, vol. 14, no. 2, pp. 207–220, 2011.

[20] USEPA, *Current Methodologies and Best Practices in Preparing Port Emission Inventories*, United States Environmental Protection Agency, 2006.

[21] A. Nagurney, *Network Economics: A Variational Inequality Approach*, Kluwer Academic Publishers, Boston, Mass, USA, 1999.

[22] P. S. Yau, S. C. Lee, J. J. Corbett, C. Wang, Y. Cheng, and K. F. Ho, "Estimation of exhaust emission from ocean-going vessels in Hong Kong," *Science of the Total Environment*, vol. 431, pp. 299–306, 2012.

[23] Entec, "Qualification of emissions from ships associated with ship movements between ports in the European community," Final Report, Entec, Newcastle upon Tyne, UK, 2002.

[24] Lloyd's Register, *Marine Exhaust Emissions Research Programme*, Lloyd's Register, 1995.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 1505817, 11 pages
http://dx.doi.org/10.1155/2016/1505817

*Research Article*

# Service and Price Decisions of a Supply Chain with Optional After-Sale Service

**Xiaochen Sun,[1] Qingshuai Zhang,[1] and Yancong Zhou[2]**

[1]*Department of Mathematics, Tianjin University, Tianjin 300072, China*
[2]*School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China*

Correspondence should be addressed to Xiaochen Sun; sxc722@163.com

Received 4 December 2015; Revised 22 February 2016; Accepted 10 March 2016

Academic Editor: Zhimin Huang

Copyright © 2016 Xiaochen Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For durable products, the high quality after-sales service has been playing an increasingly important role in consumers' purchase behaviors. We mainly study a supply chain composed of a manufacturer and a retailer. In a process of products sales, the manufacturer will provide a basic free quality assurance service. On this basis, the retailer provides paid optional quality assurance service to consumers to promote sales. Users are divided into two categories in this paper: users with no optional service and users with optional services. We derive the equilibrium decisions between the manufacturer and the retailer under the following two cases: (i) the optional after-sales service level and the wholesale price determined by the manufacturer and the retail price determined by the retailer; (ii) the wholesale price determined by the manufacturer and the optional after-sales service level and the retail price determined by the retailer.

## 1. Introduction

In recent years, as the living standard of people is improving with the rapid development of economy, people become more and more sensitive to nonprice factors they could enjoy rather than a price attribute. In 2007 the global consumer electronics products consumer research showed that service has become the second important factor that affects consumer's purchase behaviors. More and more enterprises have realized this point in time. Relying solely on price advantage, it is difficult to maintain a lasting competitive edge, such as Lenovo, WAL-MART, SONY, General Electric, and DELL. These enterprises also have to provide better services to customers so that they can maintain their market shares.

For the same product, providing different services in daily life is also becoming more common. When we buy a computer, we will encounter this kind of situation. Basic quality assurance services are to be purchased, but not all consumers are satisfied with this basic service; there is a part of the consumers who want to get more and better service. Therefore, for retailers and manufacturers, it is necessary to set up a reasonable and optional service policy to meet more consumers.

Kameshwaran et al. [1] considered the pricing problem in the following three cases: (i) only selling product, not providing service, (ii) offering product and service independently, and (iii) offering product and service bundled. Lu et al. [2] studied the supply chain decision-making problem of two competing manufacturers and one retailer. It emphasizes the importance of service provided by the manufacturer when customers are more sensitive to the price and service level of the product. Xiao and Yang [3] constructed a model that considers both price and service competition, to study the optimal decision of each member when the demand is uncertain.

In the model of Cohen and Whang [4], a customer can obtain after-sales service only from either the manufacturer or an independent service shop. Li et al. [5] considered a supply chain comprising a manufacturer and a retailer. The manufacturer supplies a product to the retailer, while the retailer sells the product bundled with after-sales service to consumers in a fully competitive market. He finds that

the manufacturer's sharing of the cost with the retailer to build service capacity improves the profits of both parties. In the model of Kurata and Nam [6], a customer can obtain simultaneously two after-sales services, one from the manufacturer and the other from the retailer. And formulating five analytical models, they find that the service level maximizing profits is not the service level that can satisfy consumers the most. As an extension of Kurata and Nam [6], Kurata and Nam [7] explored the effect of uncertainty on after-sales service decisions by comparing several information structures in a two-stage supply chain.

In fact, the optional service to a certain extent to meet the needs of customers is a stimulating factor, so we assume that the demand is subject to the impact of optional services. This is consistent with the study of Xu et al. [8]. In this case, all users are divided into two categories: users with no optional service and users with an optional service. We will focus on the equilibrium decisions between the manufacturer and the retailer with two categories of users, in which a basic service is provided by the manufacturer and an optional service is provided by the retailer or the manufacturer.

Our problem is most relevant to Kurata and Nam [6]. However, we consider a model with an optional after-sale service, in which the customers need to pay for selecting the optional service, and the price of the optional service is a decision variable. And we study the decision problems under the following two cases: (i) the optional after-sales service level and the wholesale price determined by the manufacturer and the retail price determined by the retailer; (ii) the wholesale price determined by the manufacturer and the optional after-sales service level and the retail price determined by the retailer.

The remainder of this paper is organized as follows. The second part is the problem description and modeling. In the third part, we establish the model where the optional service level is decided by the manufacturer. The fourth part is the model analysis and the solution. In the fifth part, we establish the model where the optional service level is decided by the retailer. The sixth part is the model analysis and the solution. The seventh part is the comparisons of the two kinds of equilibrium decisions by numerical results. Finally, conclusions and the future research directions are addressed in Section 8.

## 2. Problem Description and Models

We consider a supply chain composed of a manufacturer and a retailer, in which the manufacturer sells the product to the retailer at the wholesale price $w$, and the retailer sells the product to the customer at the price $p$.

There are two categories of users in the market. The first category of users only use basic service items, such as the basic quality assurance services; the second category of users are not satisfied with the basic services but also need enterprises to provide more optional services, such as extended warranty service. We use $a_1$ to represent the base-case potential market size for the first category of users. We use $a_2$ to represent the base-case potential market size for the second category of users. We assume that the basic service and the retail price are bound; that is, after the payment of the cost of the sales, the basic service does not require an additional payment, so the two categories of users are similarly sensitive to the basic service. But the optional service will attract a part of the first category of users and other product users to transfer; therefore, the optional service is negative for the first category of users' demand and is positive for the second category of users' demand, and the impact of the second category of users is more. Let $d_1$ and $d_2$, respectively, represent the first and second categories of user's demand function for the product. As popularly used in other literatures such as [9–11] on price and service competition, here $d_1$ and $d_2$ are thought of as linear functions about retail price and optional service level. Thus, the demand functions could be described by price and optional service as follows:

$$\begin{aligned} d_1(p, y) &= a_1 - \beta p - \mu_1 y, \\ d_2(p, y) &= a_2 - \beta p + \mu_2 y, \end{aligned} \tag{1}$$

where $y$ indicates the optional service level and $\beta$ indicates the price sensitive coefficient. $\mu_1$ measures the responsiveness of market demand for the first category of users to the optional service level. When the optional service is high enough, the customer insensitive to the service level will change the attitude, so the first category of customers will be less. Therefore, the optional service level has a negative effect on $d_1$, where $\mu_1 > 0$. Corresponding to the first category of customers, the second category of customers are sensitive to the optional service level, so the optional service level has a positive effect on $d_2$. $\mu_2$ measures the responsiveness of market demand for the second category of users to the optional service level, so $\mu_2 > 0$. With the optional service level increasing, the total demand of two categories should be increased, so there should be $\mu_1 < \mu_2$, meaning that, with the optional service level improving, the overall demand for the product will increase $(\mu_2 - \mu_1)y$.

If one user chooses the optional service, she/he needs to pay a certain service charge, which we call the optional service price. Of course, the optional service price changes with the optional service level. Let $y$ indicate the optional service level; then the optional service price is $r(y)$, and it is a concave function. For convenience, make $r(y) = \alpha y$.

In order to maximize their own profits, the manufacturer and the retailer need to make decisions based on their own pieces of information. We assume that the basic service level is known as well as the optional service price given by the retailer. We consider the decision-making problem of supply chain members under the condition of determined environment.

## 3. The Optional Service Level Decided by the Manufacturer

In this part, the optional service level $y$ is determined by the manufacturer, the manufacturer also determines the wholesale price $w$, and the retail price $p$ is determined by the retailer.

Hence, the manufacturer's profit function is

$$\Pi_M(w, y) = (w - c)(d_1(p, y) + d_2(p, y)) - \frac{\eta y^2}{2}. \quad (2)$$

In the above equation, the first item is the sales revenue, and the second is the cost of the optional service, as Tsay and Agrawal [10] does.

Then, the retailer's profit function can be formulated by the following:

$$\Pi_R(p) = (p - w) d_1(p, y) + (p + \alpha y - w) d_2(p, y). \quad (3)$$

Based on (1)–(3), the general optimization model can be described as follows:

$$\max \Pi_M(w, y) = (w - c)(a_1 + a_2 - 2\beta p - \mu_1 y + \mu_2 y) - \frac{\eta y^2}{2} \quad (4)$$

subject to

$$\max \Pi_R(p) = (p - w)(a_1 - \beta p - \mu_1 y) + (p + \alpha y - w)(a_2 - \beta p + \mu_2 y). \quad (5)$$

We can see that, based on the analysis stated above, in this model the manufacturer determines the wholesale price $w$ and the optional service level $y$, and the retailer determines the retail price $p$, respectively.

## 4. Results and Discussions for the Manufacturer Case

We now derive chain members' optimal strategies by maximizing their profits in the above general optimization model.

**Theorem 1.** *For any given wholesale price $w$ and optional service level $y$, the profit function $\Pi_R(p)$ is concave with respect to the retail price $p$.*

*Proof.* Consider the following derivatives:

$$\frac{\partial \Pi_R(p)}{\partial p} = a_1 + a_2 - \mu_1 y + \mu_2 y + 2\beta w - 4\beta p - \alpha\beta y,$$
$$\frac{\partial^2 \Pi_R(p)}{\partial p^2} = -4\beta. \quad (6)$$

Because $\partial^2 \Pi_R(p)/\partial p^2 = -4\beta < 0$, $\Pi_R(p)$ is concave with respect to the retail price $p$; that is to say, there does exist retail price $p'$ maximizing $\Pi_R(p)$. □

In the following, all best response functions are denoted by the superscript "$'$" and let the superscript "$*$" denote the equilibrium decisions.

**Proposition 2.** *For any given $w$ and $y$, the retailer best response function satisfies the following:*

$$p' = \frac{a_1 + a_2 - \mu_1 y + \mu_2 y + 2\beta w - \alpha\beta y}{4\beta}. \quad (7)$$

*Proof.* It can be easily proved by solving the first condition $\partial \Pi_R(p)/\partial p = 0$. □

Paying attention to the impact of wholesale price and optional service level on the retailer's best response function, we derive the first derivatives of $p'$ with respect to wholesale price and optional service level; that is, $\partial p/\partial w = 0.5 > 0$ and $\partial p/\partial w = (-\mu_1 + \mu_2 - \alpha\beta)/4\beta$. We can easily see that the higher the wholesale price, the higher the retail price. However, the effect of optional service level on the retail price is not very evident.

Obviously, if the manufacturer increases the wholesale price $w$ of the product, the retailer tends to increase the sales price $p$ of the product, which is very consistent with our intuition. And for the optional service level $y$, because the optional service price is linear with optional service level $y$, when the correlation coefficient $\alpha$ is relatively large, the retailer, tending to decrease retail price, increases sales to increase his profit.

**Lemma 3.** *The inequality $8\beta(\eta - 2\alpha\mu_2) > (\alpha\beta - \mu_1 + \mu_2)^2$ holds for any practical problem.*

*Proof.* As the retailer's operation problem, its profit must have a maximum value in any practical case.

Bringing the formulation of $p'$ into the manufacturer profit function $\Pi_M(w, y)$, we can get

$$\Pi_M = (w - c)\left(a_1 + a_2 - 2\beta\left(\frac{a_1 + a_2 - \mu_1 y + \mu_2 y + 2\beta w - \alpha\beta y}{4\beta}\right) - \mu_1 y + \mu_2 y\right) - \frac{\eta y^2}{2}. \quad (8)$$

Consider the following derivatives:

$$\frac{\partial \Pi_M}{\partial w} = \frac{a_1 + a_2 - \mu_1 y + \mu_2 y - 4\beta w + \alpha\beta y}{2} + \beta c,$$
$$\frac{\partial \Pi_M}{\partial y} = (w - c)\frac{-\mu_1 + \mu_2 + \alpha\beta}{2} - \eta y,$$
$$\begin{pmatrix} \frac{\partial^2 \Pi_M}{\partial w^2} & \frac{\partial^2 \Pi_M}{\partial w \partial y} \\ \frac{\partial^2 \Pi_M}{\partial y \partial w} & \frac{\partial^2 \Pi_M}{\partial y^2} \end{pmatrix} = \begin{pmatrix} -2\beta & \frac{\alpha\beta - \mu_1 + \mu_2}{2} \\ \frac{\alpha\beta - \mu_1 + \mu_2}{2} & -\eta \end{pmatrix}. \quad (9)$$

From the existence condition of extreme function for multivariable function, there must be

$$\Delta_2 = 2\beta\eta - \left(\frac{\alpha\beta - \mu_1 + \mu_2}{2}\right)^2 > 0. \quad (10)$$

Therefore, the lemma holds. □

**Proposition 4.** *The manufacturer equilibrium decisions satisfy*

$$w^* = \frac{2(a_1 + a_2)\eta + 4\beta\eta c - c(-\mu_1 + \mu_2 + \alpha\beta)^2}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2},$$
$$y^* = \frac{(a_1 + a_2 - 2\beta c)(-\mu_1 + \mu_2 + \alpha\beta)}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2}. \quad (11)$$

*Proof.* With the best response function for retail price $p'$ in (7), the manufacturer's best response function for the product can be achieved by solving $w^*, y^* \in \arg\max \Pi_M(w, y)$.

Let $\partial\Pi_M(x, y)/\partial y = (w - c)((-\mu_1 + \mu_2 + \alpha\beta)/2) - \eta y = 0$; we can get

$$y = \frac{(w - c)(-\mu_1 + \mu_2 + \alpha\beta)}{2\eta}. \quad (12)$$

Bring it into $\partial\Pi_M/\partial w = (a_1 + a_2 - \mu_1 y + \mu_2 y - 4\beta w + \alpha\beta y)/2 + \beta c$. Making it zero, we can get the manufacturer equilibrium decisions:

$$w^* = \frac{2(a_1 + a_2)\eta + 4\beta\eta c - c(-\mu_1 + \mu_2 + \alpha\beta)^2}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2}, \quad (13)$$

$$y^* = \frac{(a_1 + a_2 - 2\beta c)(-\mu_1 + \mu_2 + \alpha\beta)}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2}. \quad (14)$$

Bringing $w^*$, $y^*$ into $p' = (a_1 + a_2 - \mu_1 y + \mu_2 y + 2\beta w - \alpha\beta y)/4\beta$, we can get the retailer equilibrium decisions:

$$p^* = \frac{6\eta(a_1 + a_2) - 2c(-\mu_1 + \mu_2 + \alpha\beta)(-\mu_1 + \mu_2) - \alpha(a_1 + a_2)(-\mu_1 + \mu_2 + \alpha\beta) + 4\beta\eta c}{2\left(8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2\right)}. \quad (15)$$

□

**Proposition 5.** *The equilibrium wholesale price $w^*$ is increasing in $a_1$, $a_2$, $\mu_2$, and $\alpha$ but decreasing in $\mu_1$.*

*Proof.* From (13), we get the following derivatives:

$$\frac{\partial w^*}{\partial a_1} = \frac{2\eta}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2} > 0,$$
$$\frac{\partial w^*}{\partial a_2} = \frac{2\eta}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2} > 0. \quad (16)$$

From (16), we can see that the equilibrium wholesale price $w^*$ is increasing in $a_1$, $a_2$.

From (13), it is easily got that

$$w^* = \frac{2(a_1 + a_2)\eta - 4\beta\eta c}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2} + c. \quad (17)$$

From (17), we can see that when $\mu_1$ is increasing, the equilibrium wholesale price $w^*$ is decreasing in $\mu_1$. And the equilibrium wholesale price $w^*$ is increasing in $\mu_2$ and $\alpha$. □

**Proposition 6.** *If $c > 4(a_1 + a_2)$, the equilibrium wholesale price $w^*$ is increasing in $\eta$; otherwise, the equilibrium wholesale price $w^*$ is decreasing in $\eta$.*

*Proof.* From (13), we get the following derivative:

$$\frac{\partial w^*}{\partial \eta} = \frac{(c - 4(a_1 + a_2))(-\mu_1 + \mu_2 + \alpha\beta)^2}{2\left(8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2\right)^2}. \quad (18)$$

It is easy to know that if $c > 4(a_1 + a_2)$, the equilibrium wholesale price $w^*$ is increasing in $\eta$; otherwise, the equilibrium wholesale price $w^*$ is decreasing in $\eta$. □

**Proposition 7.** *If $4\beta\eta > (-\mu_1 + \mu_2 + \alpha\beta)^2$, the equilibrium wholesale price $w^*$ is increasing in $c$; otherwise, the equilibrium wholesale price $w^*$ is decreasing in $c$.*

*Proof.* From (13), we get the following derivative:

$$\frac{\partial w^*}{\partial c} = \frac{4\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2}{8\beta\eta - (-\mu_1 + \mu_2 + \alpha\beta)^2}. \quad (19)$$

It is easy to know that if $4\beta\eta > (-\mu_1 + \mu_2 + \alpha\beta)^2$, the equilibrium wholesale price $w^*$ is increasing in $c$; otherwise, the equilibrium wholesale price $w^*$ is decreasing in $c$. □

**Proposition 8.** *The equilibrium optional service level $y^*$ is increasing in $a_1$, $a_2$, $\mu_2$, and $\alpha$ but decreasing in $\mu_1$, $\eta$, and $c$.*

*Proof.* From (12), it is easily got that

$$y^* = \frac{(w^* - c)(-\mu_1 + \mu_2 + \alpha\beta)}{2\eta}. \quad (20)$$

From (20), we can see that when $\mu_1$ is increasing, $-\mu_1 + \mu_2 + \alpha\beta$ is decreasing and is greater than zero. And because $w^*$ is decreasing in $\mu_1$, when $\mu_1$ is increasing, $w^* - c$ is decreasing and is greater than zero. Hence, the equilibrium optional service level $y^*$ is decreasing in $\mu_1$. And the equilibrium wholesale price $w^*$ is increasing in $\mu_2$ and $\alpha$.

From (20), we can see that when $a_1$ is increasing, $w^*$ is increasing, so $w^* - c$ is increasing and is greater than zero. Hence, the equilibrium optional service level $y^*$ is decreasing in $a_1$. And the equilibrium wholesale price $w^*$ is increasing in $a_2$.

From (14), we can see that when $\eta$ is increasing, the equilibrium optional service level $y^*$ is decreasing in $\eta$. □

**Proposition 9.** *(i) If $6\eta > \alpha(-\mu_1 + \mu_2 + \alpha\beta)$, the equilibrium retail price $p^*$ is increasing in $a_1$ and $a_2$; otherwise, the equilibrium retail price $p^*$ is decreasing in $a_1$ and $a_2$. (ii) If $2\beta\eta > (-\mu_1 + \mu_2)(-\mu_1 + \mu_2 + \alpha\beta)$, the equilibrium retail price $p^*$ is increasing in $c$; otherwise, the equilibrium retail price $p^*$ is decreasing in $c$.*

*Proof.* From (15), it is easily got that

$$\frac{\partial p^*}{\partial a_1} = \frac{6\eta - \alpha\left(-\mu_1 + \mu_2 + \alpha\beta\right)}{2\left(8\beta\eta - \left(-\mu_1 + \mu_2 + \alpha\beta\right)^2\right)},$$
$$\frac{\partial p^*}{\partial a_2} = \frac{6\eta - \alpha\left(-\mu_1 + \mu_2 + \alpha\beta\right)}{2\left(8\beta\eta - \left(-\mu_1 + \mu_2 + \alpha\beta\right)^2\right)}. \tag{21}$$

It is easy to know that if $6\eta > \alpha(-\mu_1 + \mu_2 + \alpha\beta)$, the equilibrium retail price $p^*$ is increasing in $a_1$ and $a_2$; otherwise, the equilibrium retail price $p^*$ is decreasing in $a_1$ and $a_2$.

From (15), it is easily got that

$$\frac{\partial p^*}{\partial c} = \frac{2\beta\eta - \left(-\mu_1 + \mu_2\right)\left(-\mu_1 + \mu_2 + \alpha\beta\right)}{8\beta\eta - \left(-\mu_1 + \mu_2 + \alpha\beta\right)^2}. \tag{22}$$

It is easy to know that if $2\beta\eta > (-\mu_1 + \mu_2)(-\mu_1 + \mu_2 + \alpha\beta)$, the equilibrium retail price $p^*$ is increasing in $c$; otherwise, the equilibrium retail price $p^*$ is decreasing in $c$. □

## 5. The Optional Service Level Decided by the Retailer

In this part, the optional service level $y$ is determined by the retailer, the retailer also determines the retail price $p$, and the wholesale price $w$ is determined by the manufacturer.

Hence, the manufacturer profit function is

$$\Pi_M = (w - c)\left(d_1(p, y) + d_2(p, y)\right). \tag{23}$$

Then, the profit function of retailer can be formulated by the following:

$$\Pi_R = (p - w)\, d_1(p, y) + (p + \alpha y - w)\, d_2(p, y) - \frac{\eta y^2}{2}. \tag{24}$$

In the above equation, the first item is sales revenue, and the second is the cost of the optional service, as Tsay and Agrawal [10] does.

Based on (1), (23), and (24), the general optimization model can be described as follows:

$$\max \Pi_M = (w - c)\left(a_1 + a_2 - 2\beta p - \mu_1 y + \mu_2 y\right) \tag{25}$$

subject to

$$\max \Pi_R = (p - w)\left(a_1 - \beta p - \mu_1 y\right) + (p + \alpha y - w)\left(a_2 - \beta p + \mu_2 y\right) - \frac{\eta y^2}{2}. \tag{26}$$

We can see that, based on the analysis stated above, in this model the manufacturer determines the wholesale price $w$, and the retailer determines the retail price $p$ and the optional service level $y$, respectively.

## 6. Results and Discussions for the Retailer Case

We now derive chain members' optimal strategies by maximizing their profits in the above general optimization model.

Firstly, the following parameters hold.

**Lemma 10.** *The inequalities $4\beta(\eta - 2\alpha\mu_2) > (-\alpha\beta - \mu_1 + \mu_2)^2$ and $2\alpha\mu_2 < \eta$ hold for any practical problem.*

*Proof.* As the retailer's operation problem, its profit must have a maximum value in any practical case. Consider the following derivatives:

$$\frac{\partial \Pi_R}{\partial p} = a_1 + a_2 - 4\beta p + 2\beta w - \alpha\beta y - \mu_1 y + \mu_2 y,$$
$$\frac{\partial \Pi_R}{\partial y} = \alpha a_2 - \alpha\beta p + \mu_1 w - \mu_2 w - \mu_1 p + \mu_2 p + 2\alpha\mu_2 y - \eta y,$$
$$\begin{pmatrix} \frac{\partial^2 \Pi_R}{\partial p^2} & \frac{\partial^2 \Pi_R}{\partial p \partial y} \\ \frac{\partial^2 \Pi_R}{\partial y \partial p} & \frac{\partial^2 \Pi_R}{\partial y^2} \end{pmatrix} = \begin{pmatrix} -4\beta & -\alpha\beta - \mu_1 + \mu_2 \\ -\alpha\beta - \mu_1 + \mu_2 & 2\alpha\mu_2 - \eta \end{pmatrix}. \tag{27}$$

From the existence condition of extreme function for multivariable function, there must be

$$\Delta_3 = 2\alpha\mu_2 - \eta < 0,$$
$$\Delta_4 = 4\beta\left(\eta - 2\alpha\mu_2\right) - \left(-\alpha\beta - \mu_1 + \mu_2\right)^2 > 0. \tag{28}$$

Therefore, the lemma holds. □

**Proposition 11.** *For any given $w$, the retailer best response functions satisfy*

$$p' = \frac{\zeta\left(A + 2\beta w\right) + \lambda\left(\alpha a_2 + \mu_1 w - \mu_2 w\right)}{\Delta_4},$$
$$y' = \frac{\lambda\zeta\left(A + 2\beta w\right) + 4\beta\zeta\left(\alpha a_2 + \mu_1 w - \mu_2 w\right)}{\zeta\Delta_4}. \tag{29}$$

*Proof.* Let $\partial\Pi_R/\partial y = \alpha a_2 - \alpha\beta p + \mu_1 w - \mu_2 w - \mu_1 p + \mu_2 p + 2\alpha\mu_2 y - \eta y = 0$; we can get

$$y = \frac{\alpha a_2 - \alpha\beta p + \mu_1 w - \mu_2 w - \mu_1 p + \mu_2 p}{\eta - 2\alpha\mu_2}. \tag{30}$$

Bring it into $\partial\Pi_R/\partial p = a_1 + a_2 - 4\beta p + 2\beta w - \alpha\beta y - \mu_1 y + \mu_2 y$. Making it zero, we can get the retailer's best response decisions:

$$p = \frac{\zeta\left(A + 2\beta w\right) + \lambda\left(\alpha a_2 + \mu_1 w - \mu_2 w\right)}{\Delta_4}. \tag{31}$$

It is easy to get that

$$y = \frac{\lambda\zeta\left(A + 2\beta w\right) + 4\beta\zeta\left(\alpha a_2 + \mu_1 w - \mu_2 w\right)}{\zeta\Delta_4}. \tag{32}$$

Among them, $A = a_1 + a_2$, $\sigma = \mu_1 - \mu_2 - \alpha\beta$, $\xi = -\mu_1 + \mu_2$, and $\Delta_4 = 4\beta\zeta - \lambda^2$; that is, $\zeta = \eta - 2\alpha\mu_2$, $\lambda = -\alpha\beta - \mu_1 + \mu_2$. □

Paying attention to the impact of wholesale price on the retailer's best response function, we derive the first derivatives of $p'$ with respect to wholesale price and the first derivatives of $y'$ with respect to wholesale price; that is, $\partial p/\partial w = (2\beta\zeta + \lambda(\mu_1 - \mu_2))/\Delta_4$ and $\partial y/\partial w = 2\beta\zeta(\mu_1 - \mu_2 - \alpha\beta)/\zeta\Delta_4 < 0$. We can easily see that the higher the wholesale price, the higher the optional service level. However, the effect of the wholesale price on the retail price is not very evident.

**Theorem 12.** *The profit function $\Pi_M(w)$ is concave with respect to the wholesale price $w$.*

*Proof.* Bringing the retailer best response functions into the manufacturer profit function $\Pi_M(w)$, we can get

$$\Pi_M = (w - c)\frac{A\zeta\Delta_4 + \left(\lambda\zeta\left(-\mu_1 + \mu_2\right) - 2\beta\zeta^2\right)\left(A + 2\beta w\right) + \left(4\beta\zeta\left(-\mu_1 + \mu_2\right) - 2\beta\zeta\lambda\right)\left(\alpha a_2 + \mu_1 w - \mu_2 w\right)}{\zeta\Delta_4}. \tag{33}$$

Consider the following derivatives:

$$\frac{\partial\Pi_M}{\partial w} = \frac{A\zeta\Delta_4 + \left(\lambda\zeta\left(-\mu_1 + \mu_2\right) - 2\beta\zeta^2\right)\left(A + 2\beta w\right) + \left(4\beta\zeta\left(-\mu_1 + \mu_2\right) - 2\beta\zeta\lambda\right)\left(\alpha a_2 + \mu_1 w - \mu_2 w\right)}{\zeta\Delta_4}$$
$$+ (w - c)\frac{2\beta\left(\lambda\zeta\left(-\mu_1 + \mu_2\right) - 2\beta\zeta^2\right) + \left(\mu_1 - \mu_2\right)\left(4\beta\zeta\left(-\mu_1 + \mu_2\right) - 2\beta\zeta\lambda\right)}{\zeta\Delta_4}, \tag{34}$$

$$\frac{\partial^2\Pi_M}{\partial w^2} = \frac{-8\beta^2\zeta^2 - 8\alpha\zeta\beta^2\left(-\mu_1 + \mu_2\right)}{\zeta\Delta_4} < 0. \tag{35}$$

Because $\partial^2\Pi_M/\partial w^2 = (-8\beta^2\zeta^2 - 8\alpha\zeta\beta^2(-\mu_1 + \mu_2))/\zeta\Delta_4 < 0$, $\Pi_M(w)$ is concave with respect to the wholesale price $w$; that is to say, there does exist wholesale price $w^*$ maximizing $\Pi_M(w)$. □

**Proposition 13.** *The manufacturer equilibrium decisions satisfy*

$$w^* = \frac{A\zeta\Delta_4 + \left(\lambda\zeta\xi - 2\beta\zeta^2\right)\left(A - 2\beta c\right) + 2\beta\zeta\sigma\left(\alpha a_2 + \xi c\right)}{8\beta^2\zeta^2 + 8\alpha\zeta\beta^2\xi}. \tag{36}$$

*Proof.* Let (34) be equal to zero; we can get

$$w^* = \frac{A\zeta\Delta_4 + \left(\lambda\zeta\xi - 2\beta\zeta^2\right)\left(A - 2\beta c\right) + 2\beta\zeta\sigma\left(\alpha a_2 + \xi c\right)}{8\beta^2\zeta^2 + 8\alpha\zeta\beta^2\xi}. \tag{37}$$

□

Bringing $w^*$ into (31) and (32), we can get the retailer equilibrium decisions:

$$p^* = \frac{8\beta^2\zeta\left(\zeta + \alpha\xi\right)\left(\zeta A + \lambda\alpha a_2\right) + \left(2\beta\zeta - \lambda\xi\right)\left(A\Delta_4\zeta + \left(\lambda\zeta\xi - 2\beta\zeta^2\right)\left(A - 2\beta c\right) + 2\beta\zeta\sigma\left(\alpha a_2 + \xi c\right)\right)}{8\Delta_4\left(\beta^2\zeta^2 + \alpha\zeta\beta^2\xi\right)},$$
$$y^* = \frac{8\beta^2\zeta^2\left(\zeta + \alpha\xi\right)\left(\lambda A + 4\alpha\beta a_2\right) + 2\beta A\Delta_4\zeta^2\sigma + 2\beta\zeta\sigma\left(\lambda\zeta\xi - 2\beta\zeta^2\right)\left(A - 2\beta c\right) + 4\beta^2\zeta^2\xi\sigma^2\left(\alpha a_2 + \xi c\right)}{8\zeta\Delta_4\left(\beta^2\zeta^2 + \alpha\zeta\beta^2\xi\right)}. \tag{38}$$

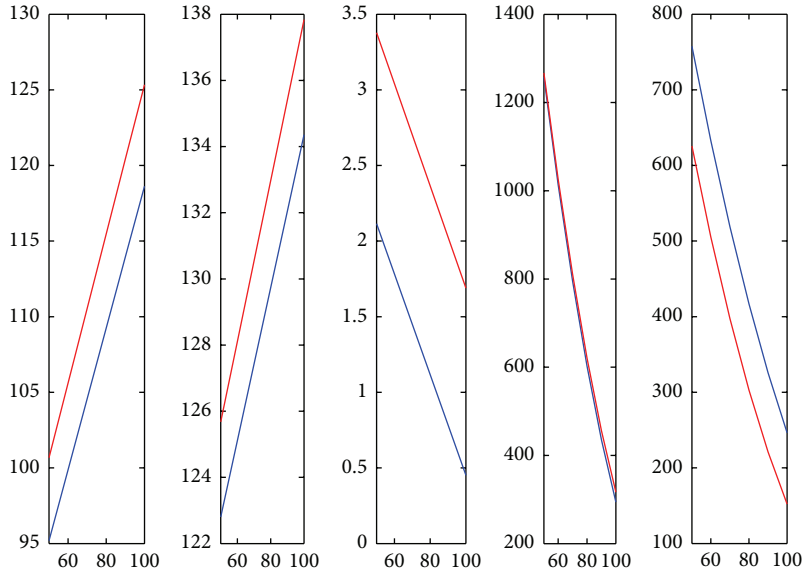



Figure 1: Effects of unit production cost.

## 7. Numerical Simulation

In this section, we will analyze the changes of different parameters of supply chain members by numerical simulation. The objective of numerical analyses is to examine how the equilibrium profit functions change when the values of parameters change. And we compare the equilibrium decisions and the equilibrium profits under two models. The baseline parameters are set as follows: $c = 80$, $a_1 = 100$, $a_2 = 50$, $\alpha = 0.2$, $\mu_1 = 0.1$, $\mu_2 = 0.4$, $\eta = 3$, and $\beta = 0.5$. In order to compare directly, we will use a diagram showing the curves of two models. The red curve represents the equilibrium decisions and the equilibrium profits where the optional service level is determined by the manufacturer and the blue curve represents the equilibrium decisions and the equilibrium profits where the optional service level is determined by the retailer.

The ordinate axes in Figures 1–8, from left to right, respectively, represent the equilibrium wholesale price, the equilibrium retail price, the equilibrium optional service level, the manufacturer's profit, and the retailer's profit.

Figure 1 indicates how the equilibrium decisions and the profits of all chain members change with respect to $c$. It is easy to see that as $c$ increases, the retail price and the wholesale price are linearly increasing, but the optimal optional service level and the profits of all chain members are linearly decreasing. In Figure 1, $c$ is from 50 to 100. We can see that the equilibrium decisions and the manufacturer profit where the optional service level is determined by the manufacturer are greater than the equilibrium decisions and the manufacturer profit where the optional service level is determined by the retailer. But the retailer profit where the optional service level is determined by the manufacturer is less than the retailer profit where the optional service level is determined by the retailer.

Figure 2 indicates how the equilibrium decisions and the profits of all chain members change with respect to $\eta$. It is easy to see that as $\eta$ increases, in the model where the optional service level is determined by manufacturer, the retail price and the wholesale price and the optional service level are decreasing, but in the model where the optional service level is determined by the retailer, the retail price and the wholesale price are increasing and the optional service level is decreasing. In Figure 2, $\eta$ is from 2 to 6. We can see that the equilibrium decisions and the manufacturer profit where the optional service level is determined by the manufacturer are greater than the equilibrium decisions and the manufacturer profit where the optional service level is determined by the retailer. But the retailer profit where the optional service level is determined by the manufacturer is less than the retailer profit where the optional service level is determined by the retailer. For $\mu_1$, there are similar results. It is represented in Figure 3. The only difference is that the retailer's profit in the model where the optional service level is determined by the retailer firstly decreases and then increases.

Figure 4 indicates how the equilibrium decisions and the profits of all chain members change with respect to $a_1$. It is easy to see that as $a_1$ increases, all of the equilibrium decisions and the profits of all chain members are increasing. In Figure 4, $a_1$ is from 50 to 200. We can see that the optimal retail price and the optimal wholesale price and the manufacturer profit in the model where the optional service level is determined by manufacturer are greater than the optimal retail price and the optimal wholesale price and the manufacturer profit in the model where the optional service level is determined by the retailer. But the retailer's profit in the model where the optional service level is determined by the manufacturer is less than the retailer's profit in the model where the optional service level is determined by the

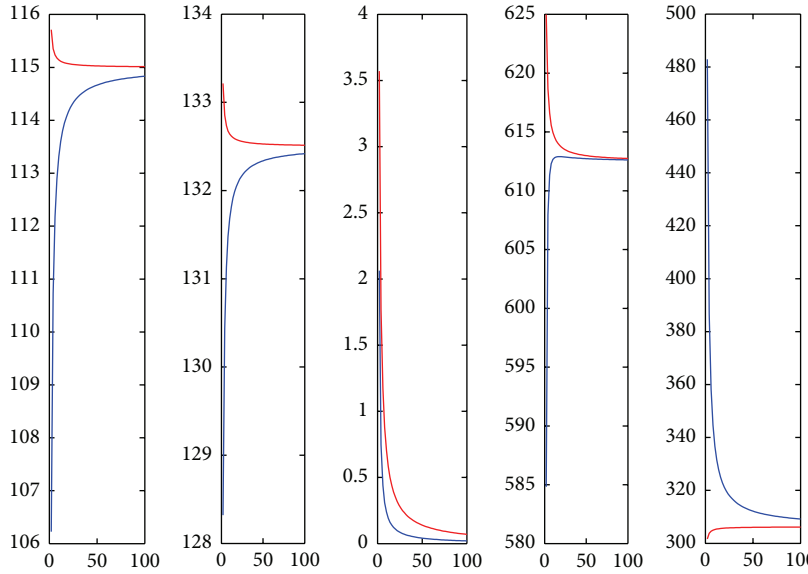



Figure 2: Effects of optional service cost coefficient.

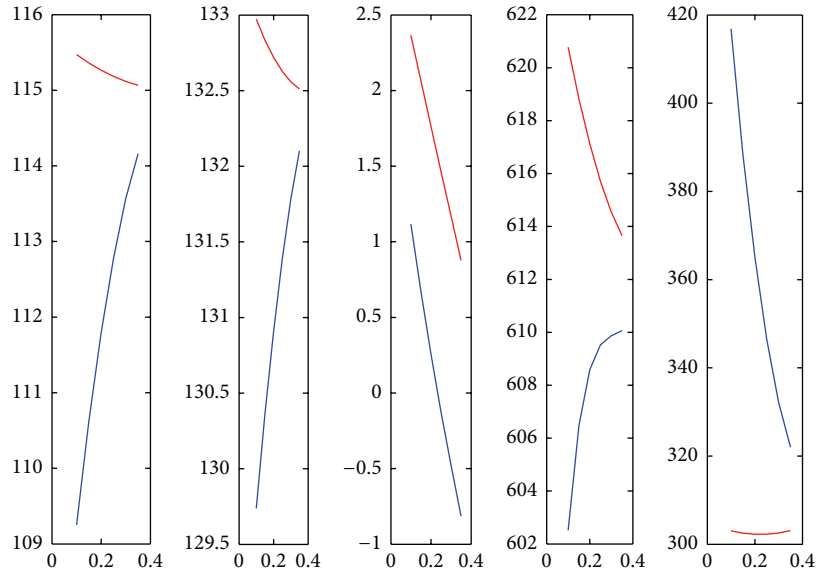



Figure 3: Effects of service elastic coefficient of the first category of customers.

retailer. Moreover, the optimal optional service level in the model where the optional service level is determined by the manufacturer is firstly less and then greater than the optimal optional service level in the model where the optional service level is determined by the retailer. For $a_2$, there are similar results. It is represented in Figure 5. The only difference is that the optimal optional service level in the model where the optional service level is determined by the manufacturer is firstly greater and then less than the optimal optional service level in the model where the optional service level is determined by the retailer.

Figure 6 indicates how the equilibrium decisions and the profits of all chain members change with respect to $\beta$. It is easy to see that as $\beta$ increases, all of the equilibrium decisions and the profits of all chain members are decreasing. In Figure 6, $\beta$ is from 0.1 to 0.5. We can see that the optimal retail price and the optimal wholesale price and the manufacturer profit in the model where the optional service level is determined by the manufacturer are greater than the optimal retail price and the optimal wholesale price and the manufacturer profit in the model where the optional service level is determined by the retailer. But the retailer's profit in the model where the

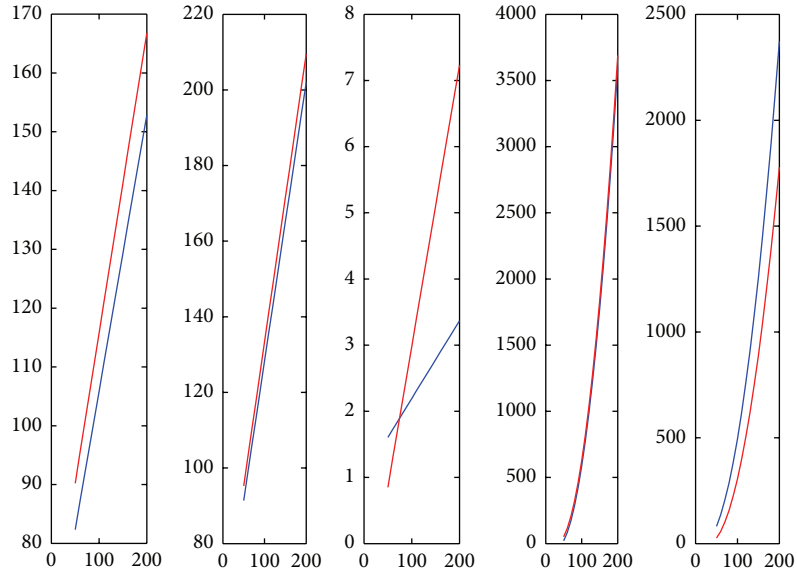



Figure 4: Effects of market base of the first category of customers.

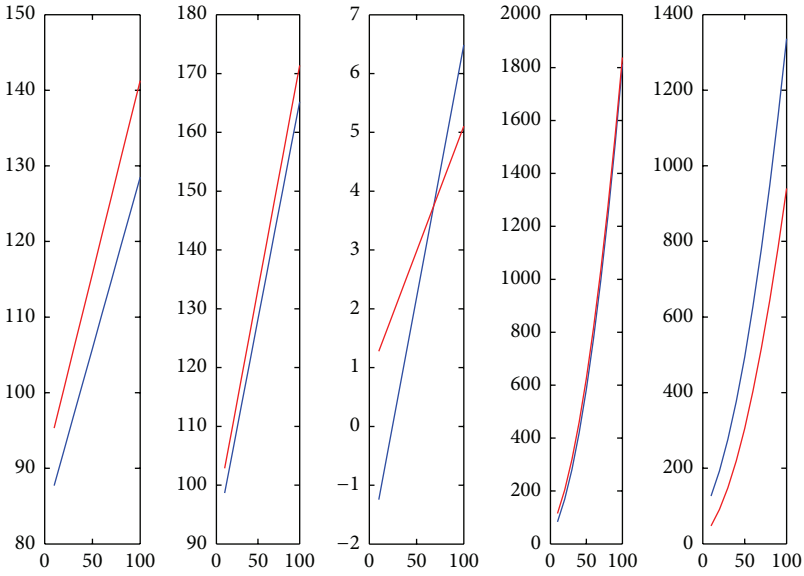



Figure 5: Effects of market base of the second category of customers.

optional service level is determined by the manufacturer is less than the retailer's profit in the model where the optional service level is determined by the retailer. Moreover, the optimal optional service level in the model where the optional service level is determined by the manufacturer is firstly greater and then less than the optimal optional service level in the model where the optional service level is determined by the retailer.

Figure 7 indicates how the equilibrium decisions and the profits of all chain members change with respect to $\mu_2$. It is easy to see that as $\mu_2$ increases, all of the equilibrium decisions and the profits of all chain members in the model where the optional service level is determined by the manufacturer are increasing, and the optimal optional service level and the retailer's profit in the model where the optional service level is determined by the retailer are increasing. In Figure 7, $\mu_2$ is from 0.3 to 0.8. We can see that the optimal retail price and the optimal wholesale price and the manufacturer's profit in the model where the optional service level is determined by the manufacturer are greater than the optimal retail price

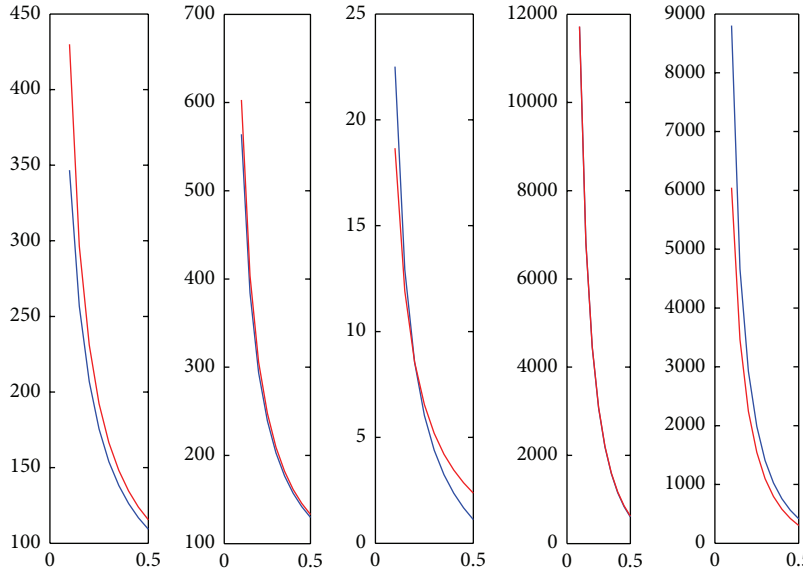



Figure 6: Effects of price sensitive coefficient.

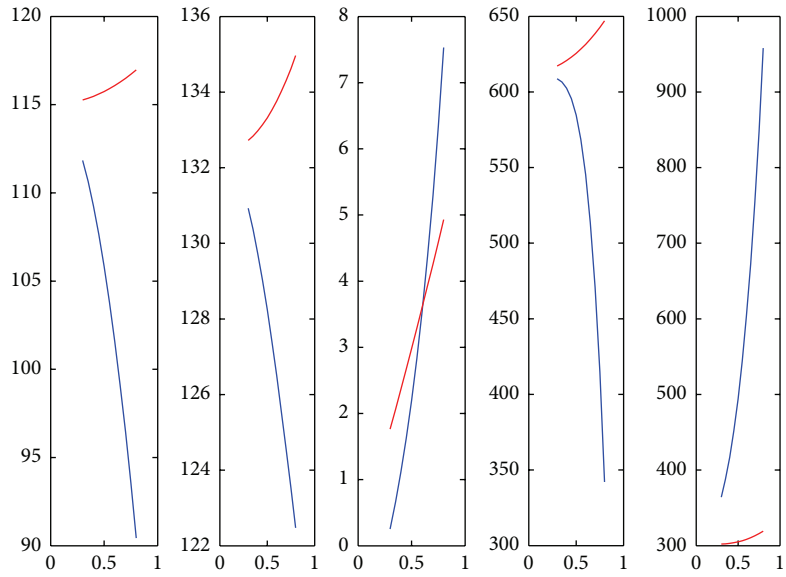



Figure 7: Effects of service elastic coefficient of the second category of customers.

and the optimal wholesale price and the manufacturer profit in the model where the optional service level is determined by the retailer. But the retailer's profit in the model where the optional service level is determined by the manufacturer is less than the retailer's profit in the model where the optional service level is determined by the retailer. Moreover, the optimal optional service level in the model where the optional service level is determined by the manufacturer is firstly greater and then less than the optimal optional service level in the model where the optional service level is determined by the retailer. For $\alpha$, there are similar results. It is represented in Figure 8. The difference is that the optimal optional service level in the model where the optional service level is determined by the retailer and the retailer's profit in the model where the optional service level is determined by the manufacturer are decreasing and the optimal optional service level in the model where the optional service level is determined by the manufacturer is firstly greater than the optimal optional service level in the model where the optional service level is determined by the retailer.

## 8. Conclusions and the Future Research Directions

Considering the existence of price and service competition, the supply chain equilibrium decision problem has become an important problem in the field of management science.

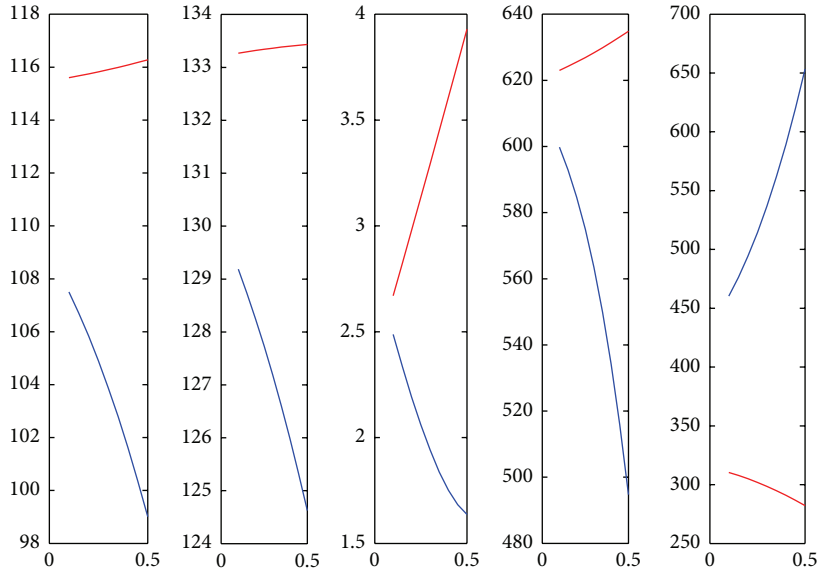



Figure 8: Effects of service elastic coefficient.

Optional service is one aspect of the service and also becomes more and more common in real life. In this paper, we assume that the basic service level is known and we use a general demand function to describe the impact of product price and optional service level on market demand in such a competitive environment. Moreover, a Stackelberg model structure is proposed between the retailer and the manufacturer in a two-echelon supply chain. We obtain the optimal wholesale price, the optimal retail price, and the optimal optional service level when the optional service level is, respectively, determined by the manufacturer and the retailer. Finally, in order to compare the two models in the optimal equilibrium decision, the profit of the members, and the whole supply chain, we carried out numerical simulation, which has certain guiding significance to the enterprise management.

However, our research leaves several unanswered questions for future research. In this paper, we assume that the demand function is determined, while, in real life, the market demand is often random. Assuming that the demand function is random is a good extension of this paper. We also add some constraints to some known parameters. The other limitation of the model is that we cannot compare the equilibrium decisions of the two models in theory because of the complexity of the equilibrium decision. Hence, how to simplify the model is another important research direction.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

This research is supported by the National Natural Science Foundation of China (NSFC), Research Fund no. 71302112.

## References

[1] S. Kameshwaran, N. Viswanadham, and V. Desai, "Bundling and pricing of product with after-sale services," *International Journal of Operational Research*, vol. 6, no. 1, pp. 92–109, 2009.

[2] J.-C. Lu, Y.-C. Tsao, and C. Charoensiriwath, "Competition under manufacturer service and retail price," *Economic Modelling*, vol. 28, no. 3, pp. 1256–1264, 2011.

[3] T. Xiao and D. Yang, "Price and service competition of supply chains with risk-averse retailers under demand uncertainty," *International Journal of Production Economics*, vol. 114, no. 1, pp. 187–200, 2008.

[4] M. A. Cohen and S. Whang, "Competing in product and service: a product life-cycle model," *Management Science*, vol. 43, no. 4, pp. 535–545, 1997.

[5] G. Li, F. F. Huang, T. C. E. Cheng, Q. Zheng, and P. Ji, "Make-or-buy service capacity decision in a supply chain providing after-sales service," *European Journal of Operational Research*, vol. 239, no. 2, pp. 377–388, 2014.

[6] H. Kurata and S.-H. Nam, "After-sales service competition in a supply chain: optimization of customer satisfaction level or profit or both?" *International Journal of Production Economics*, vol. 127, no. 1, pp. 136–146, 2010.

[7] H. Kurata and S.-H. Nam, "After-sales service competition in a supply chain: does uncertainty affect the conflict between profit maximization and customer satisfaction?" *International Journal of Production Economics*, vol. 144, no. 1, pp. 268–280, 2013.

[8] M. Xu, G. Yu, and H. Zhang, "Game analysis in a supply chain with service provision," *Journal of Management Sciences in China*, vol. 2, no. 4, pp. 18–27, 2006.

[9] G. Iyer, "Coordinating channels under price and nonprice competition," *Marketing Science*, vol. 17, no. 4, pp. 338–355, 1998.

[10] A. A. Tsay and N. Agrawal, "Channel dynamics under price and service competition," *Manufacturing and Service Operations Management*, vol. 2, no. 4, pp. 372–391, 2000.

[11] Y. Xia and S. M. Gilbert, "Strategic interactions between channel structure and demand enhancing services," *European Journal of Operational Research*, vol. 181, no. 1, pp. 252–265, 2007.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 3040343, 13 pages
http://dx.doi.org/10.1155/2016/3040343

*Research Article*

# Sourcing for Quality: Cooperating with a Single Supplier or Developing Two Competing Suppliers?

**Jingxian Chen**[1,2]

[1]*School of Business, Nantong University, 9 Seyuan Road, Nantong 226019, China*
[2]*School of Management, University of Science & Technology of China, 96 Jinzhai Road, Hefei 230026, China*

Correspondence should be addressed to Jingxian Chen; jxchenms@vip.163.com

Received 25 January 2016; Accepted 5 April 2016

Academic Editor: Alireza Amirteimoori

Copyright © 2016 Jingxian Chen. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplier efforts regarding product quality are an important issue in outsourcing and play a critical role in a manufacturer's choice of sourcing strategy. Consider a manufacturer that wants to outsource the manufacturing of two substitute products to external suppliers. This paper studies the strategic interactions under two sourcing strategies: single and dual sourcing. A four-stage noncooperative game model is established to describe each member's decisions. We further propose four decision scenarios: single sourcing with and without manufacturer quality investment sharing and dual sourcing when suppliers cooperate or do not cooperate on quality decisions. By the backward induction approach, we obtain analytical equilibrium solutions for each decision scenario. By comparing each pair of equilibrium profiles, we find that an appropriate proportion of quality investment sharing by the manufacturer can enable a cooperating strategy with a single supplier to be the dominant strategy. When the manufacturer does not want to share or does not want to share a relatively large portion of its supplier's quality investment, it will always prefer to develop two competing suppliers when the cost of dual sourcing is sufficiently low. However, dual sourcing can be extremely risky for the manufacturer because the suppliers could provide a relatively low product quality level by cooperating on the quality decision to extract the manufacturer's profit.

## 1. Introduction

Due to many benefits, such as reduced costs, decreased investments, increased productivity, and increased core competency, outsourcing has become a pervasive supply chain feature and has contributed significantly to the growth of the global economy [1–3]. A recent report estimated that, globally, organisations will outsource US $507 billion worth of business and information technology (IT) services in 2014 alone [4]. However, some argue that outsourcing has also created a new set of risks and challenges for organisations, as it has become an increasingly common business practice. Reduced control of supplier quality, a well-known downside of outsourcing, will drive poor product (service) quality, thus representing a risk to the buyer. For example, a 2012 Global Outsourcing and Insourcing Survey found that 48% of companies had previously terminated an outsourcing contract, primarily due to service quality concerns [5]. Therefore, how to encourage supplier quality improvement plays an important role in strategic outsourcing. The objective of this paper is to link the sourcing strategy with the supplier quality effort when the buyer wishes to outsource the manufacturing of multiple products to external suppliers.

With increasing overemphasis on cost reduction, lead time compression, and capacity expansion by the buyer, outsourcing enables suppliers to have no incentives on quality improvement, particularly when the quality cannot be regulated by a contract [6, 7]. As a result, firms typically enforce supply quality control by resorting to quality inspection policies. A more rigorous quality standard would be beneficial to the buyer for enhancing quality in the end market. However, such a standard could negatively affect the supplier and disrupt the supply chain due to the pressure of a standard implementation. For example, one factory of the giant contract manufacturer, Hon Hai Limited (also known as Foxconn Technology Group), in mainland China underwent a large-scale strike triggered by instructions to strengthen quality inspections for the iPhone 5; these instructions were given

by Hon Hai's customer, Apple Inc. [8]. Hence, enabling the supplier to participate in quality improvement could be critical to the buyer.

The majority of opinions regarding the handling of supplier quality incentives could be divided into two main streams. The first stream advocates that the buyer cooperate directly with its supplier, such as by providing technology assistance through sending technical engineers to the supplier site and offering quality management education and training to the supplier's employees [7]. For example, Boeing personnel are embedded into supplier factories at the global scale to monitor quality, work with suppliers on process improvements, and ensure adherence to Boeing standards and schedules [9]. These policies could be categorised as investments in supplier quality improvement or payment for supplier quality efforts, that is, cooperative quality investment [10].

There is another avenue for the buyer to pursue regarding supplier quality improvement—developing multiple competing suppliers. Because competition could be used as a mechanism for the buyer to elicit quality from suppliers, in practice, many giant manufacturers take pleasure in introducing competition among suppliers to mitigate the supplier quality risk in outsourcing [11]. For example, Apple Inc. continually seeks alternative suppliers and prefers to choose whoever has the best quality and best production rate [12]. Toyota invested in its auto seat supplier, Trim Masters, to introduce competition for its other supplier, Johnson Controls ([13] and reference therein).

Although vertical quality cooperation and horizontal quality competition are commonly used as supplier quality improvement mechanisms, which strategy will be more suitable for quality aspects and/or manufacturer profit remains unclear. Specifically, most of the extant literature established a quality enhancement mechanism for single product sourcing [1, 7, 10, 11, 13–15]. In many industries, however, manufacturers outsource the provisioning of a series of substitute products to external suppliers. For example, Apple provides iPads and iPad minis with different prices and features to the panel computer market. Nearly all of these iPads and iPad minis are assembled by the contract manufacturers Foxconn and Pegatron, respectively [16].

Product quality has a positive effect on market demand [10, 17, 18]. When a manufacturer outsources two substitute products and each product is manufactured by a special supplier, then the suppliers could have more incentives to improve product quality for competing for a larger demand share. Thus, each supplier's product decision could be different from the case of single sourcing. This paper aims to investigate the effects of the sourcing strategies on suppliers' quality decisions.

Consider a manufacturer that wants to outsource the manufacturing of two substitute products to external suppliers. This paper studies the strategic interactions for two sourcing strategies: single and dual sourcing. In addition, we propose four decision scenarios: single sourcing with and without manufacturer quality investment sharing and dual sourcing when suppliers cooperate or do not cooperate on quality decisions. This paper addresses the following research questions using a four-stage noncooperative model.

(i) Which strategy works better for improving supplier quality? If the manufacturer is willing to share a sufficiently large portion of quality investment for both products, cooperating with a single supplier can be a dominant strategy for the quality level; otherwise, the strategy of developing two competing suppliers can provide the highest product quality.

(ii) Which strategy works better for manufacturer profit? If the sourcing cost is sufficiently low, the manufacturer prefers to use the strategy of developing two competing suppliers. Otherwise, if the portion of the quality investment defrayed by the manufacturer is sufficiently small, cooperating with a single supplier benefits the manufacturer by increasing profit. Moreover, when the manufacturer's share is sufficiently large, using a single supplier without investing in its quality is better for the manufacturer.

(iii) Which strategy works better for product quality and manufacturer profit? An appropriate proportion of quality investment sharing by the manufacturer can enable dominance of the cooperating with a single supplier strategy; when the manufacturer does not want to share or does not want to share a relatively large portion of its quality investment funds, it will prefer to develop two competing suppliers when the cost of dual sourcing is sufficiently low.

In addition, dual sourcing can be extremely risky for the manufacturer because the suppliers could establish a relatively low product quality level by cooperating on the quality decision and extract the manufacturer profit.

The remainder of this paper is organised as follows: Section 2 briefly reviews the related literature, and Section 3 details our model's general framework. In Sections 4 and 5, we investigate four decision models of the quality choice: single sourcing with and without a manufacturer quality investment sharing and dual sourcing when the suppliers cooperate or do not cooperate on quality decisions. Section 6 studies which strategy is more effective for product quality and/or manufacturer profit. Concluding remarks and future directions are presented in Section 7. All proofs are deferred to the Appendix in Supplementary Material available online at http://dx.doi.org/10.1155/2016/3040343 for clarity of exposition.

## 2. Literature Review

There are three streams of literature related to our paper: strategic quality choice in a vertical channel, mechanisms of product (service) quality improvement, and sourcing strategies considering product quality.

Quality, as an important competitive feature in the majority of industries, has received considerable attention. Pioneering studies focused on the strategic quality decision that maximised social value or a monopoly firm's profit [19, 20]. Thereafter, Banker et al. [18] and Chambers et al. [17], among others, began to analyse the equilibrium quality level in a competitive environment. These studies provided fundamental insights into the issue of the effort tied to firm's quality choice. However, they did not investigate the strategic interaction among

the players in a supply chain channel environment. We refer to Chen et al. [10] for a comprehensive survey on this topic.

As a seminal work of strategic quality choice in a vertical channel, Reyniers and Tapiero [21] highlighted the importance of strategic and contractual issues in supply chain quality management. Tsay and Agrawal [22] investigated a distribution system in which a manufacturer supplies a common product to two independent retailers, who in turn use both service and retail price to directly compete for end customers. The authors derived the closed-loop equilibrium solution of retailers' price and service and then characterised the structure of wholesale pricing mechanisms to coordinate the system. Xiao and Yang [23] extended this model to two competing supply chains and considered the effect of uncertain demand and the risk attitude of the players. Xu [24] investigated a joint pricing and product quality decision in a manufacturer-retailer channel, in which the manufacturer sells a product through the retailer. Other related literature on this topic includes Balasubramanian and Bhardwaj [25], Lu et al. [26], Xie et al. [27], and Xie et al. [28]. In these studies, the information related to quality level is assumed to be symmetric among supply chain members. A number of researchers recognised that asymmetric quality-related information, such as quality level and quality cost, could be vital to quality choice for supply chain players and developed many contractual agreements for mitigating the impact of asymmetric quality information [6, 29]. However, these studies did not investigate the effects of vertical quality cooperation or horizontal quality competition on product quality.

The second stream related to our paper is that of quality improvement mechanisms and associated contract issues. Benjaafar et al. [11] developed two approaches (supplier allocation and supplier selection) to allocate fixed demand to elicit service quality from the suppliers. They concluded that a buyer could indeed orchestrate competition among potential suppliers to promote service quality. Jin and Ryan [30] extended the model into the framework of supplier's price and service competition; however, the authors only considered supplier allocation. Chen and Deng [31] introduced a certification mechanism to improve supplier quality considering asymmetric quality information. They showed that a deterministic certification may lead to underinvestment in quality improvement technology for efficient suppliers, whereas a noisy certification may alleviate this underinvestment problem. Yan [32] studied contract efficiency for a decentralised supply chain in the presence of quality improvements. Using a numerical experiment, the author also showed that a quality effort-sharing contract could be effective for improving product quality. Other examples include Ching et al. [33] and Elahi [14]. Although these studies address the competition and cooperation mechanism for improving supplier quality, they did not demonstrate which strategy would be more effective. Zhu et al. [34] considered cases in which both a buyer and a supplier invest in quality improvement. They showed that the buyer's involvement in quality investments can have a significant impact; in addition, they investigated the effect of supplier competition on quality improvement. However, they did not provide an answer for the question of which strategy is better for quality improvement when focused on quality investment decisions as well as its effects on the buyer's order quantity and the supplier's production lot size.

The third stream related to our paper is that of supplier competition under the framework of outsourcing mode. This literature focused either on price competition ([1, 35, 36], etc.) or on product competition [11]. None of them consider the case of suppliers' competitions both on price and on product quality. Jiang [37] considered a setting in which a manufacturer sequentially sources two components and uses reverse auction to select a supplier with the lowest bidding price for each component. However, the author did not investigate the effects of sourcing strategies on suppliers' decisions and the product quality decisions are not considered. Nagarajan and Bassok [38] examine a decentralised supply chain in which a single assembler buys complementary components from $n$ suppliers and assembles the final product in anticipation of demand. The authors employed the Nash bargain technique to investigate the effects of suppliers' cooperative coalition on the manufacturer's performance. The paper also neglects the impacts of different sourcing strategies.

Recently, many scholars investigated firm product quality-related sourcing decisions. Hsieh and Kuo [15] modelled a service provision game between two vendors under symmetric and asymmetric cost structures with the objective of winning the larger share of the buyer's fixed reward. Lee and Li [7] studied three different strategies that a buyer could use to manage supplier quality: cooperation, incentivisation, and inspection. Xiao et al. [3] employed Hotelling's model to characterise the strategic outsourcing decisions of two competing manufacturers whose key components have quality improvement opportunities. Jin et al. [39] integrated supplier qualification into the manufacturer's sourcing decision and developed a dual-sourcing process model. Yim [40] derived a closed-form characterisation of the optimal quota allocation for the latent defect-undependable product-external failure in single and multiple sourcing. These articles mainly integrated product quality into sourcing decisions; they also analyse the optimal (equilibrium) decision when a buyer wishes to outsource a single product but does not consider multiproduct sourcing.

Our paper differs from the above studies in that (1) we consider a manufacturer that outsources two substitute products and develop a game model to describe strategic interactions among supply chain players; (2) we integrate quality choice into sourcing decisions; and (3) we compare single and dual sourcing in terms of quality level and profits. Moreover, we also investigate the effects of vertical cooperation and horizontal competition on product quality and profits.

## 3. Model Development

We consider a sourcing system in which a monopolistic manufacturer seeks to outsource the provisioning of two substitute products to external suppliers. In purchasing a high-quality product, the manufacturer can choose a strategy among three alternatives. We label these strategies separately as the traditional strategy (T-strategy), cooperation strategy (C-strategy), and competition strategy (D-strategy). With the T-strategy, the manufacturer sources two products from

a single supplier; however, with the cooperating strategy, the manufacturer partially shares the single supplier's quality investment expenses for improving quality. With the D-strategy, the manufacturer purchases substitute products from two individual suppliers. The first two strategies are single-sourcing strategies, whereas the third strategy is a dual-sourcing strategy. In addition, we consider that suppliers can cooperate or not cooperate on quality decisions when the manufacturer uses a dual-sourcing strategy, as discussed in Section 5. This paper will model the strategic interactions of supply chain players for each decision scenario. For a better understanding of these models, we first summarise the basic notations of this paper in the Notations section.

For the Notations section, $P_i$, $W_i$, and $X_i$ are decision variables. Other related parameters are exogenous, which means that they are known to all supply chain members.

*3.1. Demand Structure.* In this paper, we characterise each product's demand by employing the Bertrand-style model, which has precedent in previous studies, such as Ingene and Parry [41], Liu et al. [42], and Zhang et al. [43]. We propose the demand function of product $i$ in the following equation:

$$Q_i\left(P_i, P_j, X_i, X_j\right) = \frac{A_i + kX_i - \theta\left(A_j + kX_j\right) - P_i + \theta P_j}{1-\theta^2}. \tag{1}$$

The demand function shown in the above equation captures both product substitutability and the impact of quality effort. Equation (1) indicates that the manufacturer has an incentive to improve product quality because it could obtain more customer demand through investing in quality enhancement. However, improving product quality is not without cost. Hence, given the manufacturer's sourcing strategy, each supplier should trade off the cost and benefits of quality improvement and then optimally select the quality effort level. To facilitate our analyses, we further assume that product $i$'s quality investment cost function is $\Theta_i X_i^2$ with $\Theta_i > 0$, that is, improving the quality level has an increasing incremental cost at high levels, which yields a diminishing return on the quality expenditure. This relationship is commonly observed in the extant literature [10, 18, 22, 26, 32].

*3.2. Sequence of Events.* We model the strategic interactions of vertical members in the outsourced supply chain as a four-stage noncooperative game, as shown in Figure 1.

Given the sourcing strategy, Figure 1 illustrates a standard supplier-Stackelberg game process, that is, the supplier first sets the quality level and wholesale price for each product, and the manufacturer responds and establishes each product's sales price after observing the supplier's choice. If the dual-sourcing mode is selected at stage 1, there are two independent suppliers that simultaneously set the product quality and wholesale price. This decision structure is commonly used in the literature of channel prices and quality decisions [28, 42, 44, 45].

In an outsourcing mode, one can expect that the OEM always has full power on the choice of the product quality, whereas we consider that the CM is in charge of setting the quality for the manufactured products. We suggest this assumption given the fact that contract manufacturers (CMs) have grown rapidly in many industries and often achieved some degree of power over OEMs [46]. For example, Kaya and Ozer [6] noted that several CMs in the pharmaceuticals industry have been undertaking research and development, creating a $30 billion drug-development and manufacturing market. These facts provide some rational explanations for the assumption that the CM optimally chooses the product quality, which has also been studied in the literature [6, 11, 30].

## 4. Single-Sourcing Strategy

We first investigate the case in which the manufacturer selects the single-sourcing strategy at stage 1, that is, the C- and T-strategies will be analysed in this section. The C-strategy implies that the manufacturer proportionally shares the supplier quality investment under the single-sourcing configuration. We assume that the manufacturer's share of quality investment of product $i$ is $\tau_i$, where $0 \le \tau_i < 1$. The case of $\tau_i = 0$ means that all quality investment expenses are defrayed by the supplier, that is, the T-strategy. Hence, the equilibrium outcomes of the T-strategy can be derived by setting $\tau_i = 0$ in the equilibrium of the C-strategy.

According to (1) and the quality investment function, we derive the supply chain member's profit function of the single-sourcing case as

$$\Pi_m = \sum_{i=1, j\neq i}^{2} \left[ \left(P_i - W_i\right) \cdot \left( \frac{A_i + kX_i - \theta\left(A_j + kX_j\right) - P_i + \theta P_j}{1-\theta^2} \right) - \tau_i \Theta_i X_i^2 \right] - f^S, \tag{2}$$

$$\Pi_s = \sum_{i=1, j\neq i}^{2} \left[ W_i \left( \frac{A_i + kX_i - \theta\left(A_j + kX_j\right) - P_i + \theta P_j}{1-\theta^2} \right) - \left(1-\tau_i\right)\Theta_i X_i^2 \right]. \tag{3}$$

From (3), this paper assumes that the unit production cost of each product is zero, which is commonly used in the related literature [25, 47]. Relaxing this assumption will not change our main results or the models' managerial insights.

Based on the players' payoff functions shown in (2) and (3), we solve this dynamic game model by using the backward induction approach and derive the following proposition.

**Proposition 1.** *For the C-strategy case, if* $8(1-\tau_i)\Theta_i(1-\theta) \ge k^2$*, the equilibrium quality level of product i is*

Table 1: Equilibrium outcomes of the single-sourcing strategy.

| Item | Equilibrium outcome |
|---|---|
| Quality level | $X_i^{C*} = \dfrac{k\left\{A_i\left[8\left(1-\tau_j\right)\Theta_j - k^2\right] - 8\theta A_j\left(1-\tau_j\right)\Theta_j\right\}}{k^4 - 8k^2\left[\left(1-\tau_i\right)\Theta_i + \left(1-\tau_j\right)\Theta_j\right] + 64\left(1-\theta^2\right)\left(1-\tau_i\right)\Theta_i\left(1-\tau_j\right)\Theta_j}$ |
| Wholesale price | $W_i^{C*} = \dfrac{64\left(1-\theta^2\right)A_i\left(1-\tau_i\right)\Theta_i\left(1-\tau_j\right)\Theta_j - 8k^2\left[A_i\left(1-\tau_i\right)\Theta_i + \theta A_j\left(1-\tau_j\right)\Theta_j\right]}{2\left\{k^4 - 8k^2\left[\left(1-\tau_i\right)\Theta_i + \left(1-\tau_j\right)\Theta_j\right] + 64\left(1-\theta^2\right)\left(1-\tau_i\right)\Theta_i\left(1-\tau_j\right)\Theta_j\right\}}$ |
| Sales price | $P_i^{C*} = \dfrac{6\left\{8\left(1-\theta^2\right)A_i\left(1-\tau_i\right)\Theta_i\left(1-\tau_j\right)\Theta_j - k^2\left[A_i\left(1-\tau_i\right)\Theta_i + \theta A_j\left(1-\tau_j\right)\Theta_j\right]\right\}}{k^4 - 8k^2\left[\left(1-\tau_i\right)\Theta_i + \left(1-\tau_j\right)\Theta_j\right] + 64\left(1-\theta^2\right)\left(1-\tau_i\right)\Theta_i\left(1-\tau_j\right)\Theta_j}$ |
| Demand quantity | $Q_i^{C*} = \dfrac{2\left(1-\tau_i\right)\Theta_i\left[8\left(A_i - \theta A_j\right)\left(1-\tau_j\right)\Theta_j - A_i k^2\right]}{k^4 - 8k^2\left[\left(1-\tau_i\right)\Theta_i + \left(1-\tau_j\right)\Theta_j\right] + 64\left(1-\theta^2\right)\left(1-\tau_i\right)\Theta_i\left(1-\tau_j\right)\Theta_j}$ |
| Manufacturer's profit | $\Pi_m^{C*} = \sum_{i=1}^{2}\left[\left(P_i^{C*} - W_i^{C*}\right)Q_i^{C*} - \tau_i\Theta_i\left(X_i^{C*}\right)^2\right] - f^s$ |
| Supplier's profit | $\Pi_s^{C*} = \sum_{i=1}^{2}\left[W_i^{C*}Q_i^{C*} - \left(1-\tau_i\right)\Theta_i\left(X_i^{C*}\right)^2\right]$ |

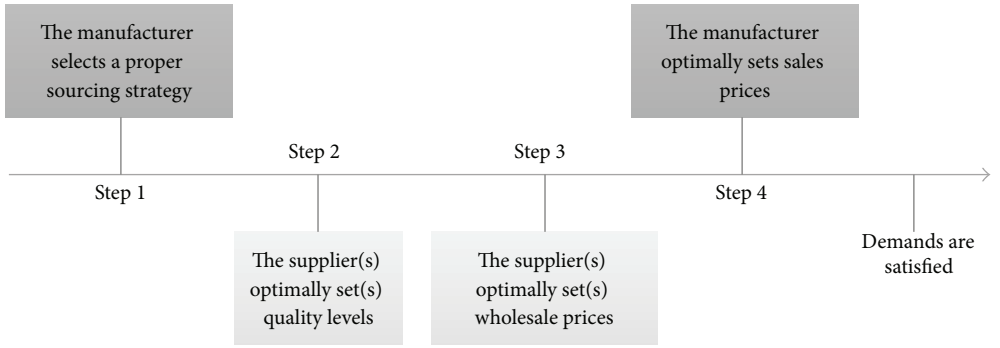



Figure 1: Sequence of decision events.

$$X_i^{C*} = \frac{k\left\{A_i\left[8\left(1-\tau_j\right)\Theta_j - k^2\right] - 8\theta A_j\left(1-\tau_j\right)\Theta_j\right\}}{k^4 - 8k^2\left[\left(1-\tau_i\right)\Theta_i + \left(1-\tau_j\right)\Theta_j\right] + 64\left(1-\theta^2\right)\left(1-\tau_i\right)\left(1-\tau_j\right)\Theta_i\Theta_j}. \tag{4}$$

In (4), the superscript C∗ denotes the optimal value for the C-strategy case. We present all proofs in the Appendix. Substituting the equilibrium product quality level into the other decision variables and the payoffs, we derive all equilibrium outcomes of this strategy, as shown in Table 1. Moreover, the equilibrium outcomes of the T-strategy can be derived by setting the proportion coefficient $\tau_i$ to zero and substituting it into the second column of Table 1. The conditions that guarantee the equilibrium can be derived by the same approach, that is, substituting $\tau_i = 0$ into the condition $8(1-\tau_i)\Theta_i(1-\theta^2) \geq k^2$.

The outcomes shown in Table 1 are not sufficiently simple to establish the effects of key parameters. Fortunately, we can investigate the effect of each parameter by imposing simplifying conditions on the remaining parameters, as is the extant literature [22, 26]. Corollaries 2 and 3 summarise the impacts of the market base, unit quality investment, and coefficient of the proportion of the quality investment share on the equilibrium outcomes under the single-sourcing setting.

**Corollary 2.** *For the C-strategy case, the comparison results of the equilibrium outcomes with regard to the asymmetric parameters are shown in Table 2.*

**Corollary 3.** *For the C-strategy case, the effects of market base, unit quality investment, and coefficient of the proportion of the quality investment share on the equilibrium outcomes are shown in Table 3.*

## 5. Dual-Sourcing Strategy

Supplier competition could be used as a mechanism for eliciting quality improvement in a decentralised supply chain [11, 14, 48]. A manufacturer may enable product quality improvements by choosing the dual-sourcing strategy to outsource

Table 2: Comparison results for the single-sourcing model when the parameters are asymmetric.

| Case | Quality level | Wholesale price | Sales price | Demand quantity |
|---|---|---|---|---|
| $\Theta_i = \Theta_j, \tau_i = \tau_j, A_i > A_j$ | $X_i^{C*} > X_j^{C*}$ | $W_i^{C*} > W_j^{C*}$ | $P_i^{C*} > P_j^{C*}$ | $Q_i^{C*} > Q_j^{C*}$ |
| $A_i = A_j, \tau_i = \tau_j, \Theta_i > \Theta_j$ | $X_i^{C*} < X_j^{C*}$ | $W_i^{C*} < W_j^{C*}$ | $P_i^{C*} < P_j^{C*}$ | $Q_i^{C*} < Q_j^{C*}$ |
| $A_i = A_j, \Theta_i = \Theta_j, \tau_i > \tau_j$ | $X_i^{C*} > X_j^{C*}$ | $W_i^{C*} > W_j^{C*}$ | $P_i^{C*} > P_j^{C*}$ | $Q_i^{C*} > Q_j^{C*}$ |

Table 3: Effect of key parameters on the equilibrium outcomes in the single-sourcing case.

| | $X_i^{C*}$ | $W_i^{C*}$ | $P_i^{C*}$ | $Q_i^{C*}$ | $\Pi_m^{C*}$ | $\Pi_s^{C*}$ |
|---|---|---|---|---|---|---|
| Case 1 | $\Theta_i = \Theta_j = \Theta, \tau_i = \tau_j = \tau$ | | | | | |
| $A_i$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ for $M_i^1 \geq 0$<br>$\leq 0$ for $M_i^1 \leq 0$ | $\geq 0$ |
| $A_j$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\geq 0$ for $M_j^1 \geq 0$<br>$\leq 0$ for $M_j^1 \leq 0$ | $\geq 0$ for $S_j^1 \geq 0$<br>$\leq 0$ for $S_j^1 \leq 0$ |
| Case 2 | $A_i = A_j = A, \tau_i = \tau_j = \tau$ | | | | | |
| $\Theta_i$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\geq 0$ for $M_i^2 \geq 0$<br>$\leq 0$ for $M_i^2 \leq 0$ | $\leq 0$ |
| $\Theta_j$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ for $M_j^2 \geq 0$<br>$\leq 0$ for $M_j^2 \leq 0$ | $\leq 0$ |
| Case 3 | $A_i = A_j = A, \Theta_i = \Theta_j = \Theta$ | | | | | |
| $\tau_i$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ for $M_i^3 \geq 0$<br>$\leq 0$ for $M_i^3 \leq 0$ | $\geq 0$ |
| $\tau_j$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\geq 0$ for $M_j^3 \geq 0$<br>$\leq 0$ for $M_j^3 \leq 0$ | $\geq 0$ |

Note: definitions of notations $M_i^1$, $M_i^2$, $M_i^3$, and $S_j^1$ are presented in online supplements.

the provisioning of two substitute products to two external suppliers, that is, the D-strategy is selected at stage 1 of the four-stage noncooperative game process (see Figure 1). The endogenous demand function of (1) means that each product's demand is not solely dependent on price but is also relative to product quality; this creates challenges for verifying the dual-sourcing strategy's quality improvement advantage and causes our studies to differ from the extant literature, such as Benjaafar et al. [11], Jin and Ryan [30], and Elahi [14]. Those authors considered customer demand as fixed and not affected by product price and quality.

In practice, different suppliers are often employed to manufacture substitute products in business practices. For example, in mainland China, most iPads are assembled at Foxconn's factory, whereas the more inexpensive iPad mini product orders are allocated to Pegatron, Inc. [16]. To investigate whether the dual-sourcing strategy benefits quality improvement, we consider two types of decision structures for the quality decision stage: centralised and decentralised quality decisions. For centralised decisions, two suppliers jointly decide their product's quality level; however, for decentralised decisions, they simultaneously choose the quality level to maximise their own profits. For expositional simplicity, we label the former as the DC-strategy and the latter as the DD-strategy.

*5.1. Centralised Quality Decision.* This case considers the situation in which two suppliers cooperatively select each product's quality level after they observe the upstream firm's dual-sourcing strategy. The decision structure implies that $X_i$ and $X_j$ are jointly set to maximise $\Pi_i + \Pi_j$. We first characterise each supply chain entity's payoff function as follows:

$$\Pi_m = \sum_{i=1, j\neq i}^{2} \left[ (P_i - W_i) \cdot \left( \frac{A_i + kX_i - \theta\left(A_j + kX_j\right) - P_i + \theta P_j}{1-\theta^2} \right) \right] - f^D,$$
$$\Pi_i = W_i \left[ \frac{A_i + kX_i - \theta\left(A_j + kX_j\right) - P_i + \theta P_j}{1-\theta^2} \right] - \Theta_i X_i^2. \tag{5}$$

By using the backward induction approach, we derive the following.

**Proposition 4.** *For the DC-strategy case, if*

$$k^2\left(4 - 3\theta^2 + \theta^4\right) \leq 2\left(1-\theta^2\right)\left(4-\theta^2\right)^2 \Theta_i,$$
$$2k^2\left(4 - 3\theta^2 + \theta^4\right)\left(\Theta_i + \Theta_j\right) \leq k^4\left(1-\theta^2\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^2 \Theta_i\Theta_j, \tag{6}$$

*then the equilibrium quality level of product i is*

$$X_i^{DC*} = \frac{2k\left(4-3\theta^2+\theta^4\right)A_i\Theta_j - 4k\theta\left(2-\theta^2\right)A_j\Theta_j - k^3\left(1-\theta^2\right)A_i}{k^4\left(1-\theta^2\right) - 2k^2\left(4-3\theta^2+\theta^4\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^2\Theta_i\Theta_j}. \tag{7}$$

In (7), the superscript DC$*$ denotes the optimal value of the DC-strategy case. The other equilibrium outcomes of this strategy are provided in Table 4.

*5.2. Decentralised Quality Decision.* In this case, two suppliers noncooperatively set the product quality level at the second stage, that is, $X_i$ is set to maximise $\Pi_i$. The following proposition is obtained from the payoff functions in (5).

**Proposition 5.** *For the DD-strategy case, if*

$$\begin{aligned} &k^2\left(2-\theta^2\right)^2 \le 2\left(1-\theta^2\right)\left(4-\theta^2\right)^2\Theta_i, \\ &2k^2\left(2-\theta^2\right)^2\left(4-\theta^2\right)\left(\Theta_i+\Theta_j\right) \\ &\quad \le k^4\left(2-\theta^2\right)^2 + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^3\Theta_i\Theta_j, \end{aligned} \tag{8}$$

*then the equilibrium quality level of product i is*

$$X_i^{DD*} = \frac{k\left(2-\theta^2\right)\left[2\left(2-\theta^2\right)\left(4-\theta^2\right)A_i\Theta_j - k^2\left(2-\theta^2\right)A_i - 2\theta\left(4-\theta^2\right)A_j\Theta_j\right]}{k^4\left(2-\theta^2\right)^2 - 2k^2\left(2-\theta^2\right)^2\left(4-\theta^2\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^3\Theta_i\Theta_j}. \tag{9}$$

In (9), the superscript DD$*$ denotes the optimal value of the DD-strategy case. The other equilibrium outcomes of this strategy are provided in Table 4.

*5.3. Effects of the Key Parameters.* To examine the effects of the various parameters, such as market base and unit quality investment, we first compare the equilibrium outcomes with the asymmetric setting on the related parameters. We derive the following corollary based on the analytical solutions in Table 4.

**Corollary 6.** *For the DC- and DD-strategy cases, the comparison results of the equilibrium outcomes with regard to the asymmetric parameters are as shown in Table 5.*

**Corollary 7.** *For the DC- and DD-strategy cases, the effects of the market base and unit quality investment on the equilibrium outcomes are shown in Table 6.*

## 6. Which Sourcing Strategy Is More Effective for the Manufacturer?

We now begin the analysis of the choice of strategy in the first stage, that is, which sourcing strategy will be selected by the manufacturer. There are two quality decision channel structures for the dual sourcing. Therefore, we must compare each pair of equilibrium outcomes for the above four decision cases. To facilitate our analysis, we assume that two substitute products are identical in the market base, unit quality investment, and manufacturer's investment sharing and then define $A_i = A_j = A$, $\Theta_i = \Theta_j = \Theta$, and $\tau_i = \tau_j = \tau$. This assumption could allow this paper to focus on the comparison results of the sourcing strategy choice. In the interim, this assumption will facilitate our computation. Furthermore, we perform the comparison based on two performance aspects: product quality level and manufacturer's profit.

*6.1. A Manufacturer Who Pursues Quality Level.* This case primarily investigates which strategy is more effective in terms of quality improvement. To answer this question, we must compare each equilibrium quality level, which was established in the aforementioned discussion. Formally, Proposition 8 provides the preferable sourcing strategy with respect to quality level.

**Proposition 8.** *Assume that $A_i = A_j = A$, $\Theta_i = \Theta_j = \Theta$, and $\tau_i = \tau_j = \tau$; there exists a unique threshold value $\tau_0$ for the manufacturer's sharing of its supplier's quality investment such that when $0 \le \tau < \tau_0$, $X_i^{DD*} > X_i^{C*} \ge X_i^{T*} > X_i^{DC*}$ (the equality is satisfied if and only if $\tau = 0$) and when $\tau_0 \le \tau < 1$, $X_i^{C*} \ge X_i^{DD*} > X_i^{T*} > X_i^{DC*}$ (the equality is satisfied if and only if $\tau = \tau_0$). The threshold value is defined as*

$$\tau_0 = \frac{\theta\left[4-\theta\left(2+\theta\right)\right]}{4\left(2-\theta^2\right)}. \tag{10}$$

The above proposition implies that if the manufacturer is willing to share a sufficiently large quality investment, it will prefer to cooperate with a single supplier. However, if the manufacturer's quality investment share is smaller than the threshold value $\tau_0$, it will prefer to develop two competing suppliers. We illustrate this result in Figure 2. Based on the parameters in this illustration, if $0 \le \tau < \tau_0 = 0.1964$, $X_i^{\mathrm{DD}*} > X_i^{\mathrm{C}*}$, whereas if $0.1964 = \tau_0 \le \tau < 1$, then $X_i^{\mathrm{C}*} \ge X_i^{\mathrm{DD}*}$

Table 4: Equilibrium outcomes in the dual-sourcing case.

| | Centralised quality decision (DC-strategy) | Decentralised quality decision (DD-strategy) |
|---|---|---|
| Quality level | $X_i^{DC*} = \frac{2k\left(4-3\theta^2+\theta^4\right)A_i\Theta_j - 4k\theta\left(2-\theta^2\right)A_j\Theta_j - k^3\left(1-\theta^2\right)A_i}{k^4\left(1-\theta^2\right) - 2k^2\left(4-3\theta^2+\theta^4\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^2\Theta_i\Theta_j}$ | $X_i^{DD*} = \frac{k\left(2-\theta^2\right)\left[2\left(2-\theta^2\right)\left(4-\theta^2\right)A_i\Theta_j - k^2\left(2-\theta^2\right)A_i - 2\theta\left(4-\theta^2\right)A_j\Theta_j\right]}{k^4\left(2-\theta^2\right)^2 - 2k^2\left(2-\theta^2\right)^2\left(4-\theta^2\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^3\Theta_i\Theta_j}$ |
| Wholesale price | $W_i^{DC*} = \frac{2\left(1-\theta^2\right)\left[2\left(2-\theta^2\right)\left(4-\theta^2\right)A_i\Theta_i\Theta_j - k^2\left(2-\theta^2\right)A_i\Theta_i - 2\theta\left(4-\theta^2\right)A_j\Theta_i\Theta_j - k^2\theta A_j\Theta_j\right]}{k^4\left(1-\theta^2\right) - 2k^2\left(4-3\theta^2+\theta^4\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^2\Theta_i\Theta_j}$ | $W_i^{DD*} = \frac{2\left(1-\theta^2\right)\left(4-\theta^2\right)\Theta_i\left[2\left(2-\theta^2\right)\left(4-\theta^2\right)A_i\Theta_j - 2\theta\left(4-\theta^2\right)A_j\Theta_j - k^2\left(2-\theta^2\right)A_i\right]}{k^4\left(2-\theta^2\right)^2 - 2k^2\left(2-\theta^2\right)^2\left(4-\theta^2\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^3\Theta_i\Theta_j}$ |
| Sales price | $P_i^{DC*} = \frac{2\left(4-5\theta^2+\theta^4\right)\left[2\left(3-\theta^2\right)A_i - \theta A_j\right]\Theta_i\Theta_j - 2k^2\left(3-3\theta^2+\theta^4\right)A_i\Theta_i - k^2\theta\left(5-3\theta^2\right)A_j\Theta_j}{k^4\left(1-\theta^2\right) - 2k^2\left(4-3\theta^2+\theta^4\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^2\Theta_i\Theta_j}$ | $P_i^{DD*} = \frac{\left(4-\theta^2\right)\left\{2\left(4-5\theta^2+\theta^4\right)\left[2\left(3-\theta^2\right)A_i - \theta A_j\right]\Theta_i\Theta_j - k^2\left[\left(6-7\theta^2+2\theta^4\right)A_i\Theta_i + \theta\left(2-\theta^2\right)\right]A_j\Theta_j\right\}}{k^4\left(2-\theta^2\right)^2 - 2k^2\left(2-\theta^2\right)^2\left(4-\theta^2\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^3\Theta_i\Theta_j}$ |
| Demand quantity | $Q_i^{DC*} = \frac{2\left(2-\theta^2\right)\left(4-\theta^2\right)A_i\Theta_i\Theta_j - k^2\left[\left(2-\theta^2\right)A_i\Theta_i + \theta A_j\Theta_j\right] - 2\theta\left(4-\theta^2\right)A_j\Theta_i\Theta_j}{k^4\left(1-\theta^2\right) - 2k^2\left(4-3\theta^2+\theta^4\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^2\Theta_i\Theta_j}$ | $Q_i^{DD*} = \frac{\left(4-\theta^2\right)\Theta_i\left[2\left(2-\theta^2\right)\left(4-\theta^2\right)A_i\Theta_j - k^2\left(2-\theta^2\right)A_i - 2\theta\left(4-\theta^2\right)A_j\Theta_j\right]}{k^4\left(2-\theta^2\right)^2 - 2k^2\left(2-\theta^2\right)^2\left(4-\theta^2\right)\left(\Theta_i+\Theta_j\right) + 4\left(1-\theta^2\right)\left(4-\theta^2\right)^3\Theta_i\Theta_j}$ |
| Manufacturer's profit | $\Pi_m^{DC*} = \sum_{i=1}^{2}\left[\left(P_i^{DC*} - W_i^{DC*}\right)Q_i^{DC*}\right] - f^D$ | $\Pi_m^{DD*} = \sum_{i=1}^{2}\left[\left(P_i^{DD*} - W_i^{DD*}\right)Q_i^{DD*}\right] - f^D$ |
| Supplier's profit | $\Pi_i^{DC*} = W_i^{DC*}Q_i^{DC*} - \Theta_i\left(X_i^{DC*}\right)^2$ | $\Pi_i^{DD*} = W_i^{DD*}Q_i^{DD*} - \Theta_i\left(X_i^{DD*}\right)^2$ |

Table 5: Comparison results for the dual-sourcing model when the parameters are asymmetric.

| Strategy | Case | Quality level | Wholesale price | Sales price | Demand quantity | Supplier's profit |
|---|---|---|---|---|---|---|
| DC-strategy | $\Theta_i = \Theta_j, A_i > A_j$ | $X_i^{DC*} > X_j^{DC*}$ | $W_i^{DC*} > W_j^{DC*}$ | $P_i^{DC*} > P_j^{DC*}$ | $Q_i^{DC*} > Q_j^{DC*}$ | $\Pi_i^{DC*} > \Pi_j^{DC*}$ |
| | $A_i = A_j, \Theta_i > \Theta_j$ | $X_i^{DC*} < X_j^{DC*}$ | $W_i^{DC*} < W_j^{DC*}$ | $P_i^{DC*} < P_j^{DC*}$ | $Q_i^{DC*} < Q_j^{DC*}$ | $\Pi_i^{DC*} < \Pi_j^{DC*}$ |
| DD-strategy | $\Theta_i = \Theta_j, A_i > A_j$ | $X_i^{DD*} > X_j^{DD*}$ | $W_i^{DD*} > W_j^{DD*}$ | $P_i^{DD*} > P_j^{DD*}$ | $Q_i^{DD*} > Q_j^{DD*}$ | $\Pi_i^{DD*} > \Pi_j^{DD*}$ |
| | $A_i = A_j, \Theta_i > \Theta_j$ | $X_i^{DD*} < X_j^{DD*}$ | $W_i^{DD*} < W_j^{DD*}$ | $P_i^{DD*} < P_j^{DD*}$ | $Q_i^{DD*} < Q_j^{DD*}$ | $\Pi_i^{DD*} < \Pi_j^{DD*}$ |

Table 6: Effect of key parameters on the equilibrium outcomes of the dual-sourcing case.

| Strategy | | $X_i^{DC(D)*}$ | $W_i^{DC(D)*}$ | $P_i^{DC(D)*}$ | $Q_i^{DC(D)*}$ | $\Pi_m^{DC(D)*}$ | $\Pi_i^{DC(D)*}$ |
|---|---|---|---|---|---|---|---|
| DC-strategy | Case 1 | $\Theta_i = \Theta_j = \Theta$ | | | | | |
| | $A_i$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ for $M_i^4 \geq 0$<br>$\leq 0$ for $M_i^4 \leq 0$ | $\geq 0$ for $S_i^2 \geq 0$<br>$\leq 0$ for $S_i^2 \leq 0$ |
| | $A_j$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\geq 0$ for $M_j^4 \geq 0$<br>$\leq 0$ for $M_j^4 \leq 0$ | $\geq 0$ for $S_j^3 \geq 0$<br>$\leq 0$ for $S_j^3 \leq 0$ |
| | Case 2 | $A_i = A_j = A$ | | | | | |
| | $\Theta_i$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\geq 0$ for $M_i^5 \geq 0$<br>$\leq 0$ for $M_i^5 \leq 0$ | $\geq 0$ for $S_i^4 \geq 0$<br>$\leq 0$ for $S_i^4 \leq 0$ |
| | $\Theta_j$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ for $M_j^5 \geq 0$<br>$\leq 0$ for $M_j^5 \leq 0$ | $\geq 0$ for $S_j^5 \geq 0$<br>$\leq 0$ for $S_j^5 \leq 0$ |
| DD-strategy | Case 1 | $\Theta_i = \Theta_j = \Theta$ | | | | | |
| | $A_i$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ for $M_i^6 \geq 0$<br>$\leq 0$ for $M_i^6 \leq 0$ | $\geq 0$ for $S_i^6 \geq 0$<br>$\leq 0$ for $S_i^6 \leq 0$ |
| | $A_j$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\geq 0$ for $M_j^6 \geq 0$<br>$\leq 0$ for $M_j^6 \leq 0$ | $\geq 0$ for $S_i^7 \geq 0$<br>$\leq 0$ for $S_i^7 \leq 0$ |
| | Case 2 | $A_i = A_j = A$ | | | | | |
| | $\Theta_i$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\leq 0$ | $\geq 0$ for $M_i^7 \geq 0$<br>$\leq 0$ for $M_i^7 \leq 0$ | $\geq 0$ for $S_i^8 \geq 0$<br>$\leq 0$ for $S_i^8 \leq 0$ |
| | $\Theta_j$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ | $\geq 0$ for $M_j^7 \geq 0$<br>$\leq 0$ for $M_j^7 \leq 0$ | $\geq 0$ |

Note: definitions of notations $M_i^4$, $M_i^5$, $M_i^6$, $M_i^7$, $S_i^2$, and $S_i^4$ are presented in online supplements.

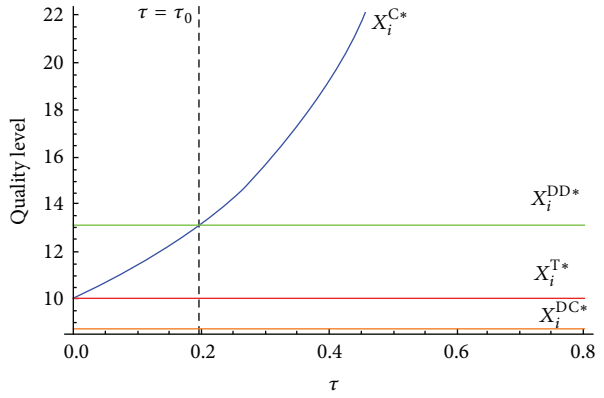



Figure 2: Comparison results of the equilibrium quality level. Note: $A = 100$, $\Theta = 2$, $k = 2$, and $\theta = 0.5$.

is observed. Moreover, the difference $X_i^{C*} - X_i^{DD*}$ increases with increases in the manufacturer's investment share when $\tau \geq \tau_0$.

The second insight observed in Proposition 8 is that the traditional strategy is dominated by the quality cooperation strategy. An interesting observation found in the above proposition is that the DC-strategy is dominated by the other three cases. Thus, the manufacturer should consider supplier quality collusion when it decides to outsource substitute products to two external suppliers. The manufacturer may obtain products with poorer quality levels if the two suppliers cooperate on quality decisions.

*6.2. A Manufacturer Who Pursues Profit.* We now consider which strategy can best service the manufacturer's profit pursuit. By analysing the difference between each pair of the manufacturer's equilibrium payoffs, the following proposition summarises the results of which strategy is more effective for increasing profit. For a better understanding of the effect of strategy choice on the profit, we assume that the fixed cost of single sourcing is zero (i.e., $f^S = 0$), but the dual-sourcing strategy's cost is $F > 0$ (i.e., $f^D = F > 0$). Relaxing this assumption cannot change our primary results and

managerial insights but does increase the computational complexity. Furthermore, the fixed cost $F$ cannot exceed the manufacturer's reservation profit $F^{\max}$ (i.e., $F < F^{\max}$), which is defined in (11). Without this constraint, the manufacturer cannot derive a positive profit in the dual-sourcing case. Consider

$$\begin{aligned} F &\le F^{\max} \\ &= \min\left\{\frac{2A^2(1+\theta)\left(4-\theta^2\right)^2\Theta^2}{\left[2(2-\theta)^2(1+\theta)(2+\theta)\Theta-k^2\left(2-\theta^2\right)\right]^2},\right. \\ &\quad \left.\frac{2A^2(1+\theta)(2-\theta)^2\Theta^2}{\left[2(2-\theta)^2(1+\theta)\Theta-k^2(1-\theta)\right]^2}\right\} \\ &= \frac{2A^2(1+\theta)(2-\theta)^2\Theta^2}{\left[2(2-\theta)^2(1+\theta)\Theta-k^2(1-\theta)\right]^2}. \end{aligned} \tag{11}$$

**Proposition 9.** *Assume that $A_i = A_j = A$, $\Theta_i = \Theta_j = \Theta$, and $\tau_i = \tau_j = \tau$; when $0 \le \tau < \tau_1$ and $0 \le F < F_1$, the DD-strategy is the dominant strategy; however, when $0 \le \tau < \tau_1$ and $F_1 \le F < F^{\max}$, the C-strategy is the dominant strategy. When $\tau_1 \le \tau < 1$ and $0 \le F < F_2$, the DD-strategy is the dominant strategy; however, when $\tau_1 \le \tau < 1$ and $F_2 \le F < F^{\max}$, the T-strategy is the dominant strategy. The threshold values $\tau_1$, $F_1$ and $F_2$ are defined as*

$$\begin{aligned} \tau_1 &= \frac{k^2\left[8(1+\theta)\Theta-k^2\right]}{4(1+\theta)\Theta\left[16(1+\theta)\Theta-k^2\right]}, \\ F_1 &= \max\left\{0, -2A^2\Theta\left\{\frac{4(1+\theta)(1-\tau)^2\Theta-k^2\tau}{\left[8(1+\theta)(1-\tau)\Theta-k^2\right]^2}\right.\right. \\ &\quad \left.\left.-\frac{(1+\theta)\left(4-\theta^2\right)^2\Theta}{\left[2(1+\theta)(2+\theta)(2-\theta)^2\Theta-k^2\left(2-\theta^2\right)\right]^2}\right\}\right\}, \\ F_2 &= \max\left\{0, -2A^2(1+\theta)\Theta^2\left\{\frac{4}{\left[8(1+\theta)\Theta-k^2\right]^2}\right.\right. \\ &\quad \left.\left.-\frac{\left(4-\theta^2\right)^2}{\left[2(1+\theta)(2+\theta)(2-\theta)^2\Theta-k^2\left(2-\theta^2\right)\right]^2}\right\}\right\}. \end{aligned} \tag{12}$$

The above proposition illustrates that when the fixed cost of dual sourcing is sufficiently low, the manufacturer will prefer to develop two competing suppliers regardless of the investment share. However, when the dual-sourcing cost is sufficiently high, the manufacturer will prefer to cooperate with the single supplier when sharing a sufficiently small portion of the quality investment. In contrast, the manufacturer will prefer to use the traditional strategy when the quality investment share is sufficiently high. Figure 3 illustrates these results through a numerical example.

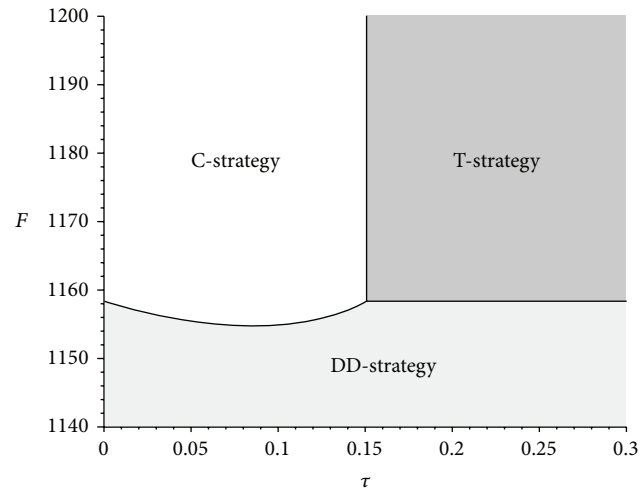



Figure 3: Dominant strategy of the manufacturer who pursues profit. Note: $A = 100$, $\Theta = 2$, $k = 2$, and $\theta = 0.5$.

In addition, the profit-pursuing manufacturer will never choose the dual-sourcing strategy when both suppliers cooperatively set product quality. In conjunction with Proposition 8, the DC-strategy cannot be a dominant strategy in both aspects of quality level and profit. Therefore, the manufacturer should prevent supplier quality collusion when it decides to use the dual-sourcing strategy. However, the suppliers have incentives for quality cooperation because they can extract more profits by producing a low-quality product (the relation $\Pi_i^{DC*} > \Pi_i^{DD*}$ can be derived by computing the difference of $\Pi_i^{DC*}$ and $\Pi_i^{DD*}$). Therefore, in this situation, it is important to design an efficient contract to encourage improvements in supplier quality. A quality contest among suppliers could be used to enable quality improvement among suppliers; this is also established in the fixed demand model [11, 30]. However, with the quality-dependent demand model, contesting for demand may not be applicable because the equilibrium solution cannot be derived or, at a minimum, cannot prove the existence and uniqueness; thus, the effect of this mechanism on quality level will be unclear. The manufacturer can design a quality contest among suppliers to strive for a fixed quality investment subsidy. However, this problem is not the main focus of this study; thus, we would encourage researchers to devote additional attention to this issue in the future.

*6.3. Summary.* Propositions 8 and 9 summarise the results of which strategy is more effective for product quality and manufacturer profit, respectively. We now consider whether there is a sourcing strategy that benefits quality level in addition to manufacturer's profit. First, the T- and DC-strategies are excluded because the former is strictly dominated by the C-strategy in the aspect of quality and the latter is the worst case of the four strategies in both aspects. Therefore, we consider the C- and DD-strategy cases. The following observation provides useful materials for answering this question.

*Observation 1.* Assume that $A_i = A_j = A$, $\Theta_i = \Theta_j = \Theta$, and $\tau_i = \tau_j = \tau$; then, we obtain the following conclusions: (i) when the unit quality investment is neither sufficiently large nor sufficiently small, that is, $\Theta^{\min} \le \Theta \le \Theta^{\max}$, the C-strategy could be beneficial to both supplier quality improvement and manufacturer's profit enhancement if and only if $\tau_0 \le \tau \le \tau_1$ and $F \ge F_1$; and (ii) when the fixed cost of dual sourcing is sufficiently small, that is, $F \le \min(F_1, F_2)$, the DD-strategy could be beneficial to both supplier quality improvement and manufacturer profit enhancement if and only if $0 \le \tau < \tau_0$.

The Appendix provides theoretical support for this observation. Parameters $\Theta^{\min}$ and $\Theta^{\max}$ can also be found in the Appendix. The following insights can be obtained from the above observation:

(i) Quality cost information plays an important role in implementing the strategy of cooperating with a single supplier. Because the supplier has incentives to hide the private quality cost information [6], the manufacturer should design a mechanism to reveal the type of supplier, with regard to quality cost, when the manufacturer wants to employ the C-strategy to improve quality and profits. Moreover, ensuring that the quality investment share is in the appropriate range is also important to successfully implement the C-strategy. In the interim, a relatively high sourcing cost for dual sourcing could ensure that the manufacturer has an incentive to implement C-strategy.

(ii) When the manufacturer does not want to share or does not want to share a large portion of the supplier quality investment, developing two competing suppliers could be the optimal strategy for improving quality level and profit if the sourcing cost is sufficiently low. Moreover, if the supplier demands a large share of the quality investment, the manufacturer will switch to developing two competing suppliers. In this scenario, preventing collusion on the two suppliers' quality becomes a critical issue for the manufacturer. Otherwise, the manufacturer could migrate into the worst case if the two suppliers cooperatively set the product quality level.

## 7. Conclusions

In this paper, we examine the equilibrium quality levels of two substitute products in four decision scenarios belonging to two sourcing strategies: single sourcing and dual sourcing. There are two different decision scenarios for the single-sourcing strategy: the supplier optimally sets each product's quality level with and without the manufacturer's investment sharing. Two decision scenarios also exist for the dual-sourcing strategy: two suppliers cooperatively or noncooperatively set product quality levels. This study employs a four-stage noncooperative game process to model the strategic interactions among supply chain players. Given each decision scenario, we derive the closed-loop equilibrium outcomes of each game model and propose the conditions that guarantee that the Nash equilibrium of each game uniquely exists. Moreover, this paper examines the effects of key parameters, such as the market base, quality cost margin, and investment share, on players' decisions and the equilibrium payoffs.

Finally, we compare the equilibrium quality levels and manufacturer's profits of each decision scenario. The question of which strategy is more effective for product quality or manufacturer's profit is resolved. By combining these results, we further demonstrate that cooperating with a single supplier and developing two competing suppliers could be the dominant strategy in terms of both quality level and profit. The conditions for guaranteeing the optimal strategy are also established. These results will help to effectively implement the sourcing strategy in practice. This paper also provides valuable insights into the quality management strategy that leads to quality improvement and the selection of a supply chain channel structure for supply chain outsourcing.

There are several directions for future research. First, we derive the effects of the key parameters on the equilibrium outcomes by assuming partially or completely symmetrical parameters settings. A similar assumption could be found in the comparison results between each pair of decision scenarios. One possible extension is to compute these results with complete asymmetrical parameters. Second, we prove that suppliers' quality cooperation in the dual-sourcing strategy is dominated by the other three strategies. Hence, integrating well-known mechanisms, such as a quality contest into the vertical channel to break through suppliers' quality collusion, could be worth investigating. Finally, we assume that the quality investment margin of each product is common knowledge in this study. However, quality effort costs can be private information. Moreover, the manufacturer's strategy choice could be dependent on quality investment unit. Thus, it is interesting but challenging to discuss sourcing strategy when suppliers keep quality cost information private.

## Notations

| | |
|---|---|
| $i = 1, 2$; $j = 1, 2$, and $j \ne i$: | Index to denote each product |
| $Q_i$: | Gross demand of product $i$ |
| $A_i$: | Market base of product $i$ |
| $\theta$ ($0 \le \theta < 1$): | Coefficient of capturing product substitutability |
| $k$ ($k > 0$): | Coefficient of capturing the impact of quality level on market base |
| $P_i$: | Product $i$'s sales price |
| $X_i$: | Product $i$'s quality level |
| $W_i$: | Product $i$'s wholesale price |
| $\Theta_i$: | Unit quality investment for the $i$th product |
| $f^S$: | Fixed configuration cost of single-sourcing mode |
| $f^D$: | Fixed configuration cost of dual-sourcing mode |
| $\Pi_m$: | Manufacturer's profit |

$\Pi_s$: Supplier's profit in single-sourcing mode
$\Pi_i$: Supplier $i$'s profit in dual-sourcing mode.

## Competing Interests

The author declares no competing interests.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 71401082) and Research Fund for Humanities and Social Sciences of Chinese Ministry of Education (no. 14YJC630009).

## References

[1] Q. Feng and L. X. Lu, "The strategic perils of low cost outsourcing," *Management Science*, vol. 58, no. 6, pp. 1196–1210, 2012.

[2] K. M. Gilley and A. Rasheed, "Making more by doing less: an analysis of outsourcing and its effects on firm performance," *Journal of Management*, vol. 26, no. 4, pp. 763–790, 2000.

[3] T. Xiao, Y. Xia, and G. P. Zhang, "Strategic outsourcing decisions for manufacturers competing on product quality," *IIE Transactions*, vol. 46, no. 4, pp. 313–329, 2014.

[4] L. Plunkett Research, "Introduction to the outsourcing & offshoring industry," 2014, http://www.plunkettresearch.com/outsourcing-offshoring-bpo-market-research/industry-and-business-data.

[5] Deloitte, "Managing outsourcing risks at the early stages," *The Wall Street Journal*, 2014, http://deloitte.wsj.com/riskandcompliance/2014/03/03/managing-early-stage-outsourcing-risks/.

[6] M. Kaya and O. Ozer, "Quality risk in outsourcing: noncontractible product quality and private quality cost information," *Naval Research Logistics*, vol. 56, no. 7, pp. 669–685, 2009.

[7] H.-H. Lee and C. Li, "Sourcing for quality: incentives, cooperation, and inspection," Working Paper, University of Hongkong, 2012.

[8] P. Elmer-Dewitt, What's really going on at Apple's iPhone 5 factory in Zhengzhou, China. Fortune, 2012, http://fortune.com/2012/10/07/whats-really-going-on-at-apples-iphone-5-factory-in-zhengzhou-china/.

[9] Boeing, "World Class Supplier Quality," 2014, http://787updates.newairplane.com/787-Suppliers/World-Class-Supplier-Quality.

[10] J. Chen, L. Liang, and F. Yang, "Cooperative quality investment in outsourcing," *International Journal of Production Economics*, vol. 162, pp. 174–191, 2015.

[11] S. Benjaafar, E. Elahi, and K. L. Donohue, "Outsourcing via service competition," *Management Science*, vol. 53, no. 2, pp. 241–259, 2007.

[12] Knowledge@Wharton, *The Foxconn-Sharp Alliance: Which Company Will Gain the Most?*, 2012, http://knowledge.wharton.upenn.edu/article/the-foxconn-sharp-alliance-which-company-will-gain-the-most/.

[13] C. Li, "Sourcing for supplier effort and competition: design of the supply base and pricing mechanism," *Management Science*, vol. 59, no. 6, pp. 1389–1406, 2013.

[14] E. Elahi, "Outsourcing through competition: what is the best competition parameter?" *International Journal of Production Economics*, vol. 144, no. 1, pp. 370–382, 2013.

[15] C.-C. Hsieh and P.-L. Kuo, "Analysis of a dual-sourcing service provision game with symmetric and asymmetric vendors," *Applied Mathematical Modelling*, vol. 35, no. 8, pp. 3998–4007, 2011.

[16] L. Luk, "Hon Hai profit growth slows," *The Wall Street Journal*, 2014, http://online.wsj.com/articles/hon-hai-second-quarter-profit-growth-slows-1407927508.

[17] C. Chambers, P. Kouvelis, and J. Semple, "Quality-based competition, profitability, and variable costs," *Management Science*, vol. 52, no. 12, pp. 1884–1895, 2006.

[18] R. D. Banker, I. Khosla, and K. K. Sinha, "Quality and competition," *Management Science*, vol. 44, no. 9, pp. 1179–1192, 1998.

[19] A. Dixit, "Quality and quantity competition," *The Review of Economic Studies*, vol. 46, no. 4, pp. 587–599, 1979.

[20] A. M. Spence, "Monopoly, quality, and regulation," *The Bell Journal of Economics*, vol. 6, no. 2, pp. 417–429, 1975.

[21] D. J. Reyniers and C. S. Tapiero, "The delivery and control of quality in supplier-producer contracts," *Management Science*, vol. 41, no. 10, pp. 1581–1589, 1995.

[22] A. A. Tsay and N. Agrawal, "Channel dynamics under price and service competition," *Manufacturing & Service Operations Management*, vol. 2, no. 4, pp. 372–391, 2000.

[23] T. Xiao and D. Yang, "Price and service competition of supply chains with risk-averse retailers under demand uncertainty," *International Journal of Production Economics*, vol. 114, no. 1, pp. 187–200, 2008.

[24] X. Xu, "Optimal price and product quality decisions in a distribution channel," *Management Science*, vol. 55, no. 8, pp. 1347–1352, 2009.

[25] S. Balasubramanian and P. Bhardwaj, "When not all conflict is bad: manufacturing-marketing conflict and strategic incentive design," *Management Science*, vol. 50, no. 4, pp. 489–502, 2004.

[26] J.-C. Lu, Y.-C. Tsao, and C. Charoensiriwath, "Competition under manufacturer service and retail price," *Economic Modelling*, vol. 28, no. 3, pp. 1256–1264, 2011.

[27] G. Xie, S. Wang, and K. K. Lai, "Quality improvement in competing supply chains," *International Journal of Production Economics*, vol. 134, no. 1, pp. 262–270, 2011.

[28] G. Xie, W. Yue, S. Wang, and K. K. Lai, "Quality investment and price decision in a risk-averse supply chain," *European Journal of Operational Research*, vol. 214, no. 2, pp. 403–410, 2011.

[29] S. Baiman, P. E. Fischer, and M. V. Rajan, "Information, contracting, and quality costs," *Management Science*, vol. 46, no. 6, pp. 776–789, 2000.

[30] Y. Jin and J. K. Ryan, "Price and service competition in an outsourced supply chain," *Production and Operations Management*, vol. 21, no. 2, pp. 331–344, 2012.

[31] Y.-J. Chen and M. Deng, "Supplier certification and quality investment in supply chains," *Naval Research Logistics*, vol. 60, no. 3, pp. 175–189, 2013.

[32] X. Yan, "Contract efficiency for a decentralized supply chain in the presence of quality improvement," *International Transactions in Operational Research*, vol. 22, no. 4, pp. 713–734, 2015.

[33] W.-K. Ching, S.-M. Choi, and X. Huang, "Inducing high service capacities in outsourcing via penalty and competition," *International Journal of Production Research*, vol. 49, no. 17, pp. 5169–5182, 2011.

[34] K. Zhu, R. Q. Zhang, and F. Tsung, "Pushing quality improvement along supply chains," *Management Science*, vol. 53, no. 3, pp. 421–436, 2007.

[35] G. P. Cachon and P. T. Harker, "Competition and outsourcing with scale economies," *Management Science*, vol. 48, no. 10, pp. 1314–1333, 2002.

[36] A. Arya, B. Mittendorf, and D. E. M. Sappington, "The make-or-buy decision in the presence of a rival: strategic outsourcing to a common supplier," *Management Science*, vol. 54, no. 10, pp. 1747–1758, 2008.

[37] L. Jiang, "Managing supplier competition and sourcing sequence for component manufacturing," *Production and Operations Management*, vol. 24, no. 2, pp. 287–310, 2015.

[38] M. Nagarajan and Y. Bassok, "A bargaining framework in supply chains: the assembly problem," *Management Science*, vol. 54, no. 8, pp. 1482–1496, 2008.

[39] Y. Jin, J. K. Ryan, and W. Yund, "Sourcing decisions with competitive suppliers and imperfect information," *Decision Sciences*, vol. 45, no. 2, pp. 229–254, 2014.

[40] A. Yim, "Failure risk and quality cost management in single versus multiple sourcing decision," *Decision Sciences*, vol. 45, no. 2, pp. 341–354, 2014.

[41] C. A. Ingene and M. E. Parry, *Mathematical Models of Distribution Channels*, Kluwer Academic, Boston, Mass, USA, 2004.

[42] B. Liu, G. Cai, and A. A. Tsay, "Advertising in asymmetric competing supply chains," *Production and Operations Management*, vol. 23, no. 11, pp. 1845–1858, 2014.

[43] R. Zhang, B. Liu, and W. Wang, "Pricing decisions in a dual channels system with different power structures," *Economic Modelling*, vol. 29, no. 2, pp. 523–533, 2012.

[44] S. C. Choi, "Price competition in a channel structure with a common retailer," *Marketing Science*, vol. 10, no. 4, pp. 271–296, 1991.

[45] S. C. Choi, "Price competition in a duopoly common retailer channel," *Journal of Retailing*, vol. 72, no. 2, pp. 117–134, 1996.

[46] J. V. Gray, B. Tomlin, and A. V. Roth, "Outsourcing to a powerful contract manufacturer: the effect of learning-by-doing," *Production and Operations Management*, vol. 18, no. 5, pp. 487–505, 2009.

[47] G. Iyer, "Coordinating channels under price and nonprice competition," *Marketing Science*, vol. 17, no. 4, pp. 338–355, 1998.

[48] N. Gans, "Customer loyalty and supplier quality competition," *Management Science*, vol. 48, no. 2, pp. 207–221, 2002.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 6153898, 7 pages
http://dx.doi.org/10.1155/2016/6153898

*Research Article*

# Research on Medicine Distribution Route Optimization for Community Health Service Institutions

**Zhao Li, Zhiqiang Ma, Wei Shi, and Xianggen Qian**

*School of Management, Jiangsu University, No. 301 Xuefu Road, Zhenjiang, Jiangsu 212013, China*

Correspondence should be addressed to Wei Shi; sw586022@163.com

Received 25 January 2016; Accepted 26 April 2016

Academic Editor: Zhimin Huang

Copyright © 2016 Zhao Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

At present, the optimization of medicine distribution route has become an urgent issue that needs to be solved in the unified medicine distribution for community health service institutions. Considering the characteristics of medicine distribution for community health service institutions, to minimize the overall cost (including refrigeration storage cost, vehicle fixed cost, and transportation cost) of medicine distribution in a certain region, a transport-distance-constrained local community medicine distribution route optimizing model is established. Then, a tabu-search-based algorithm originated from client direct arrangement ideology is put forward and MATLAB language is used to simulate the medicine distribution procedure. The simulated results show that the proposed algorithm is capable of obtaining an optimum distribution scheme cost with minimum transportation cost.

## 1. Introduction

Developing community health service is an effective method to optimize the allocation of medical health resource and to make medical services in China more affordable and accessible for ordinary people. Recent researches show that the exceedingly high logistics and distribution cost is one of the major causes contributing to soaring medical fees [1]. The distribution route is designed directly to affect the distribution cost and delivery speed of medicine and the quality of medical service. As medicare reform deepens continuously the community health service in China will step into a stage of large-scale development, which calls for a unified distribution for community medicine. Therefore, it provides with great theoretical and practical significance to design an optimal medicine distribution route for community health service institutions.

The community health service institutions, which undertake part of medicare function, are mainly oriented towards the disease prevention, health care, and health education. Therefore, the medical distribution for these institutions has its distinguishing characteristics. Compared to pharmacies and hospitals, they are prone to such logistics problems as small medicine demand, frequent delivery, widely and randomly distributed branches, and high logistics cost. On the other hand, they require more on temperature-sensitive medicines, like vaccines, insulin injections, and tincture. High temperature during delivery may lead to the deterioration of these medicines and make them become toxic. According to Zhao's research, community health service institution has developed its own new pharmaceutical supply chain, which is different from that of hospitals and pharmacies [2]. Wang established a medicine distribution model for community health service institution based on the system of logistic chain management in his doctoral thesis, but he failed to present detailed model algorithm due to its theoretical complexity [3]. J. Zhang and K. Zhang optimized the medicine distribution routing using ant colony algorithm [4]. Zheng et al. set up a model optimizing the dynamic medicine distribution route in urban areas by use of the Genetic Algorithm [5].

Through an overall review of current researches, it is found that while numerous researches have been done on medicine distribution for hospitals and pharmacies, few researches have been made specially on medicine distribution for community health service institutions and its distinguishing characteristics have failed to be taken into account. Therefore, compared with current researches, this

paper has three different points. Firstly, this paper reflects the characteristics of community medicine distribution such as the precise time, the special vehicles, and the fixed branches. So the overall distribution cost includes three key elements in the objective function, which are fixed cost, transport cost, and refrigeration storage cost. Secondly, this paper constructs a model more flexible and adaptable for community medicine distribution. It can complete the optimal distribution route design according to different conditions on its own. Thirdly, this paper improves the tabu-search algorithm based on the idea of client-direct-arrangement and applies it into the model solution to seek the optimal distribution plan meeting all constraint conditions. The result should be contributed to provide some new insights into the design of optimal medicine distribution for community health service institutions and thus reduce the regional medicine distribution cost.

## 2. Problems Description in Community Medical Distribution Route

The characteristics of medicine distribution for community health service institutions can be described as follows: (1) strict requirements for the accuracy and time of delivery. For example, in Zhenjiang, Jiangsu Province, there are a total of 53 community health service centers and stations widely distributed within a region of 130 square kilometers. The long distance between scattered branches complicates the distribution. (2) The second characteristic is the high requirements for the control of temperature, humidity, exposure to light, and sanitary conditions during transit. Therefore, special vehicles are required for the task of delivery. (3) The third one is the basically fixed locations and clients. Community health service institutions mainly serve for middle-aged and elder citizens suffering from chronic diseases, children who need to be vaccinated, and patients with common diseases, which keeps their demand for medicine stable in quantity and variety [6] and makes it possible to conduct unified medicine distribution.

In view of the characteristics described above, the optimization of medicine distribution route can be simplified into a mathematic subject: given a fixed distribution center and $k$ number of medicine delivering vehicles with refrigerating devices. Let the load capacity of each vehicle be $Q$, and let the maximum traveling distance of each vehicle once a time be $L$. Suppose that there are $n$ numbers of branches around that need medicine (with exactly known location), and the requested amount of medicine for the $i$th institution is $q_i$. Suppose that the starting point and destination for each vehicle are the distribution center, the fixed operating cost for each vehicle is $c_0$, transportation cost per kilometer is $c_1$, and the cost for cooling per hour is $c_2$. With all these parameters known, an optimized distribution scheme needs to be given to reduce the cost for a unified distribution.

## 3. Model Assumptions and Constraints

### 3.1. Assumptions

(1) The amount and location of community health service institutions are constant.

(2) There is a distribution center with fixed location and adequate distributing capacity.

(3) The demand for medicine in each community health service branch can be satisfied readily.

(4) The usage of vehicles is fixed; after delivery is done, vehicle should return to the distribution center.

(5) In order to avoid overload issues, the amount of each delivery should not exceed the overall requested amount each time.

(6) One-to-one rule is applied, which means that one vehicle only serves one certain route.

### 3.2. Constraints

(1) Load on vehicles: the overall amount of medicine requested by branches along certain route should not exceed the load capacity of the vehicle.

(2) Maximum distribution distance: the overall distance of a certain route should not exceed the maximum distribution distance of the vehicle.

(3) One-way rule: a vehicle only follows one direction along a certain route, which means that the vehicle cannot turn back and each branch only can be served by one vehicle once.

## 4. Construction of Community Medicine Distribution Route Optimization Model

Based on the assumptions and constraints described above, the model is established as follows:

$$\min \quad Z = m \cdot c_0 + c_1 \cdot \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=1}^{m} L_{ij} \cdot x_{ijk} + c_2 \cdot \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=1}^{m} t_{ij} \cdot x_{ijk}, \tag{1}$$

$$\text{s.t.} \quad \sum_{i=0}^{n} q_i \cdot y_{ik} \le Q \quad k = 1, 2, \ldots, m; \tag{2}$$

$$\sum_{i=0}^{n} \sum_{j=0}^{n} L_{ij} \cdot x_{ijk} \le L \quad k = 1, 2, \ldots, m; \tag{3}$$

$$\sum_{k=1}^{m} y_{ik} = 1 \quad i = 0, 1, \ldots, n; \tag{4}$$

$$\sum_{i=0}^{m} x_{ijk} = y_{jk} \quad j = 0, 1, \ldots, n;\ k = 1, 2, \ldots, m; \tag{5}$$

$$\sum_{j=0}^{n} x_{ijk} = y_{ik} \quad i = 0, 1, \ldots, n;\ k = 1, 2, \ldots, m; \tag{6}$$

$$y_{ik} = \begin{cases} 1, & \text{the } i\text{th client is served by vehicle } k, \\ 0, & \text{others;} \end{cases} \tag{7}$$

$$x_{ijk} = \begin{cases} 1, & \text{vehicle } k \text{ drives from client } i \text{ to client } j, \\ 0, & \text{others} \end{cases} \tag{8}$$

in which $L_{ij}$ is the distance between client $i$ and client $j$, $L$ is the maximum distance traveled by a vehicle each time, $t_{ij}$ is the time a vehicle takes to drive from client $i$ to client $j$, $i = 0$ means distribution center, and the amount of medicine requested is 0 ($q_0 = 0$). Other parameters are the following: $m$ is the number of delivery vehicles, $c_0$ is the fixed operating cost of vehicles, $c_1$ is delivery cost per distance unit of vehicle, and $c_2$ is cooling cost per unit time. The meanings for each expression shown above are as follows.

Expression (1): overall cost of community medicine distribution cost includes fixed operating cost of vehicles, delivery cost per distance unit of vehicle, and cooling cost; Expression (2): vehicles are not overloaded; Expression (3): the overall distance of one certain route does not exceed the maximum delivery distance of vehicle; Expression (4): each vehicle only serves one client once; Expressions (5)-(6): the route for each vehicle is closed: logistic center—one or more community health service institutions—logistic center; Expressions (7)-(8): they reflect integer constraints.

Time window is not taken into account in this model, considering the fact that the delivery time requested by each branch is almost the same, generally before 8 a.m. Therefore, as long as the shortest route is achieved, the least time for distribution can be assured. In addition, for the sake of saving energy, after distribution is done, cooling devices on vehicles are suggested to be turned off on the way back to the distribution center.

## 5. Algorithm Design

Tabu search, TS for short, is a metaheuristic algorithm proposed by Glover in 1986. It is actually extension of local neighborhood search [7]. Through simulating the memory mechanism of human intelligence, TS algorithm introduces an agile storage structure and corresponding tabu criteria to avoid circuitous search. It can achieve a global optimum by remitting some tabooed good status according to remitting criteria [8–10]. Being different from genetic algorithm and simulated annealing algorithm, tabu search algorithm is an intelligent random algorithm, which can overcome the defection of precocious convergence during the search process and finally achieve a global optimum [11]. Now, it has been applied to solve classic issues like traveling businessmen issue, siting issue, and route optimization problem [12–14]. In order to find out an optimized route for community medicine distribution, this paper improves the traditional tabu search algorithm mainly based on the idea of customer straight arrangement [10].

### 5.1. Key Elements Analysis

*5.1.1. Initial Solution.* The choice of initial solution affects the search efficiency and convergence speed of TS algorithm. Generally, the traditional tabu search algorithm generates initial solution randomly. But it has large probability of randomly generated infeasible initial solution in the community medicine distribution route optimization, because of too many conditions of satisfaction and infeasible solutions. So, we construct the initial solution based on the idea of customer straight arrangement and divide the initial solution into feasible solutions which satisfy the constraint conditions by the method of greed segmentation. We prevent the problem that random initial solution is infeasible as far as possible.

*5.1.2. Evaluation Function.* The meaning of evaluation function is a method of determining the solution evaluation. It determines complexity and performance of the algorithm implementation. The method of edge-directed solution arrangement is much used for evaluation function construction in the existing literature, but it has its deficiency such as being time-consuming, search-inefficient, and convergence-slow. In this paper, we design the evaluation function of TS algorithm also based on the idea of customer straight arrangement. It reduces the number of infeasible solutions, narrows the search scope, and improves the search efficiency. Therefore, it provides the possibility for the improvement of the whole performance of TS algorithm.

The evaluation function for tabu algorithm is designed following the ideology: take the route optimization of one

distribution center and 6 communities in one certain region, for example. Set logistic center as zero, current solution $S$ equals 342156, and constraints are the maximum load capacity and maximum traveling distance of vehicle for each single delivery. Firstly, add client 3 as the 1st client into route 1; then judge whether it can satisfy the constraints, namely, to judge whether the amount of medicine requested by client 3 and the whole distance of route 0-3-0 can satisfy the constraints. If it does, add the client 4 as the 2nd one into this route, and go on with next calculation. If not, end calculation for route 0-3-0, add client 4 into route 2, and go on with the calculation. Assume that 3 routes are derived, 0-3-0, 0-4-2-1-0, and 0-5-6-0. $L_0$ is the sum of the distance of these 3 routes. The evaluation function of current solution can be calculated by $f(x) = 3c_0 + L_0 * c_1$.

*5.1.3. Neighborhood and Candidate.* Neighborhood structure design is very important to ensure the efficiency of TS algorithm and the solution of the overall situation. Combined with the characteristics and requirement of community medicine distribution route optimization problem, we adopt the SWAP operation to build the neighborhood structure, namely, to exchange the positions of two random elements in solution. The number of neighborhood solutions is $C_n^2$ produced by SWAP operation. According to the characteristics of the problem, the candidate solutions are preferential selected in the field of current state.

*5.1.4. Tabu Table and Termination Criterion.* This paper uses the queue storage structure (FIFO: first in and first out) to update tabu table: when a new feasible solution is generated, its reverse move will be inserted to the bottom of the tabu list, and the first tabu element will be shifted out from the top; then the other tabu elements are moving up one unit. For the algorithm performance, too overloaded value of length of tabu list could increase the search cost, and too short value of length of tabu list could let the algorithm fell into local circulation. Generally, two equally effective ways could be used to set length of tabu list, which are static setting and dynamic setting. This paper adopts static setting to determine the length of tabu list in order to be convenient for calculation and solving. The static length of tabu list should be less than TabuL = round($\sqrt{n(n-1)/2}$) according to literature research, and $n$ is the number of search points (namely, the number of community health service institutions in Section 6).

In addition, this paper uses the specified maximum iteration step length method as the termination criterion of TS algorithm. Maxnum represents the number of feasible solutions, also used as the largest iterative steps. The search algorithm is ended when the total number of the detecting feasible solutions reaches to Maxnum. Then, we save and output the current best solution as the optimal solution.

*5.2. Steps of Tabu Search Algorithm.* According to the basic theory of tabu search algorithm, steps for searching an optimized solution are as follows.

*Step 1.* Use a computer to generate an element collection randomly, and use greedy segmentation method to decompose it into a number of subsets which satisfy the constraint conditions, set it as the initial solution $S_0$, and make the tabu table $H = \Phi$; at the same time make the algorithm iterate counter Num = 0.

*Step 2.* Use customer straight alignment method and evaluation function to construct the evaluation value $f(x)$ of the algorithm, and make the current optimal solution as BSF = $S_0$.

*Step 3.* Use pairwise interchange to conduct local neighborhood operation: exchange the positions of two elements in current solution randomly, and save it as a new current solution in alternative solution set Candidate (the number of elements in this solution set does not exceed $n * (n-1)/2$).

*Step 4.* Select one solution $S_1$ from the set Candidate as current solution, and make a comparison between its evaluation value $f(S_1)$ and the evaluation value of current optimum solution $f$(BSF); if $f(S_1) < f$(BSF), then let BSF = $S_1$; otherwise, go to Step 3, continue to select a new solution from Candidate set for comparison, and then update the tabu table $H$.

*Step 5.* Determine whether Num $\leq$ Maxnum is valid; if it is valid, return to Step 3 to continue the search; otherwise, terminate the algorithm and output current optimal solutions BSF.

## 6. Application

We choose Jingkou District of Zhenjiang City in Jiangsu Province as the application area. The community medicine distribution is unified and entrusted to a specific company by the medicine distribution center. However, the decision-making of this company about the required number of vehicles and the actual distribution path depends on too much experience. Moreover, this company is short of initiative for optimization of the community medicine distribution. Therefore, this paper tests and verifies the effectiveness of the optimization model by means of logical data processing. At the same time, we hope that the conclusion could contribute to the improvement of the actual operation efficiency of community medicine distribution.

There are one medicine distribution center, two community health service centers, and sixteen community health service stations in Jingkou District. So the total number of community health service institutions is 18; namely, $n = 18$. And the medicine distribution center has several vehicles with 8-ton load capacity. We set the coordinate of the distribution center as $(0, 0)$ and get the corresponding coordinate of each community health service institution referring to $(0, 0)$. The coordinates and the amount of medicine requested by each community health service institution are shown in Table 1.

In MATLAB 6.5, the parameters of tabu search algorithm are set as follows.

Table 1: Coordinates (km) and the amount of medicine requested by each community health service institution.

| Delivery destination | Service center 1 | Service center 2 | Service station 1 | Service station 2 | Service station 3 | Service station 4 |
|---|---|---|---|---|---|---|
| Coordinates | (13, 15) | (−40, −60) | (60, 48) | (49, 23) | (−72, 80) | (−51, −39) |
| The amount of medicine requested | 3 | 2 | 0.1 | 0.3 | 0.2 | 0.1 |
| Delivery destination | Service station 5 | Service station 6 | Service station 7 | Service station 8 | Service station 9 | Service station 10 |
| Coordinates | (38, −58) | (60, −45) | (73, 69) | (−65, 70) | (20, 25) | (30, −20) |
| The amount of medicine requested | 0.4 | 0.5 | 1 | 0.6 | 0.3 | 0.8 |
| Delivery destination | Service station 11 | Service station 12 | Service station 13 | Service station 14 | Service station 15 | Service station 16 |
| Coordinates | (40, −18) | (50, 35) | (−25, 60) | (−15, −10) | (92, 42) | (−20, −35) |
| The amount of medicine requested | 0.2 | 1 | 1.2 | 0.2 | 0.5 | 0.6 |

Table 2: Results of tabu algorithm.

| Iteration | Distribution route | Time for each iteration (s) | Numbers of vehicle/m | Overall cost/c |
|---|---|---|---|---|
| 1 | 0 → 16 → 18 → 6 → 2 → 8 → 17 → 9 → 3 → 14 → 4 → 13 → 12 → 0<br>0 → 1 → 11 → 7 → 15 → 5 → 10 → 0 | 37.14 | 2 | 8790.6 |
| 2 | 0 → 1 → 11 → 16 → 18 → 6 → 2 → 8 → 12 → 0<br>0 → 10 → 5 → 15 → 7 → 9 → 17 → 3 → 14 → 4 → 13 → 0 | 28.46 | 2 | 8998.5 |
| 3 | 0 → 18 → 2 → 6 → 10 → 5 → 15 → 1 → 16 → 0<br>0 → 11 → 7 → 9 → 17 → 3 → 14 → 4 → 8 → 13 → 12 → 0 | 38.5 | 2 | 8553.4 |
| 4 | 0 → 1 → 16 → 18 → 2 → 6 → 10 → 5 → 15 → 0<br>0 → 11 → 7 → 9 → 17 → 3 → 14 → 4 → 8 → 13 → 12 → 0 | 35.02 | 2 | 8466.6 |
| 5 | 0 → 11 → 4 → 14 → 3 → 17 → 9 → 7 → 15 → 5 → 10 → 0<br>0 → 1 → 12 → 13 → 8 → 2 → 6 → 18 → 16 → 0 | 34.38 | 2 | 8358.7 |
| 6 | 0 → 18 → 2 → 6 → 16 → 7 → 9 → 17 → 3 → 14 → 4 → 12 → 13 → 8 → 0<br>0 → 1 → 11 → 15 → 5 → 10 → 0 | 39.34 | 2 | 8929.4 |
| 7 | 0 → 1 → 15 → 5 → 10 → 6 → 2 → 18 → 16 → 0<br>0 → 11 → 7 → 9 → 17 → 3 → 14 → 4 → 8 → 13 → 12 → 0 | 33.25 | 2 | 8217.6 |
| 8 | 0 → 6 → 2 → 8 → 13 → 12 → 4 → 14 → 3 → 17 → 9 → 7 → 11 → 0<br>0 → 1 → 15 → 5 → 10 → 18 → 16 → 0 | 29.92 | 2 | 9042.7 |
| 9 | 0 → 1 → 11 → 16 → 18 → 6 → 2 → 8 → 12 → 0<br>0 → 10 → 5 → 15 → 7 → 9 → 17 → 3 → 14 → 4 → 13 → 0 | 27.41 | 2 | 8998.5 |
| 10 | 0 → 18 → 2 → 6 → 16 → 11 → 7 → 15 → 5 → 10 → 0<br>0 → 1 → 14 → 3 → 9 → 17 → 4 → 8 → 13 → 12 → 0 | 36.95 | 2 | 9046.7 |

The length of tabu table is TabuL = round($\sqrt{n(n-1)/2}$) = 12, the number of candidate solutions is 144, the largest iterative step is Maxnum = 500, fixed operating cost of vehicles $c_0$ = 500 RMB, transportation cost $c_1$ = 10 RMB/km, cooling cost for medicine $c_2$ = 20 RMB/hour, average velocity of vehicles $v$ = 60 km/hour, maximum distribution distance in one single time $L$ = 500 km, and maximum load capacity $T$ = 8 t. Run the calculation 10 times randomly; we can get result as shown in Table 2.

According to Table 2, it can be seen that the average distribution cost is 8740.3 RMB, numbers of vehicles are 2 sets, average iteration time is 34 s, minimum distribution cost is 8217.6 RMB, and optimum route is the 1st route, 0 → 1 → 15 → 5 → 10 → 6 → 2 → 18 → 16 → 0, and the 2nd route: 0 → 11 → 7 → 9 → 17 → 3 → 14 → 4 → 8 → 13 → 12 → 0. The search procedure and optimum route are shown in Figures 1 and 2.

From Figure 1, it can be seen that the improved tabu search algorithm converged rapidly and could locate optimum solution after 20 iterations, which means that this algorithm is quite effective in seeking an optimum route for community health service medicine distribution. In addition, according to the simulation results, after the calculation is converged, the search procedure remained stable, which means that this algorithm is reliable and stable.

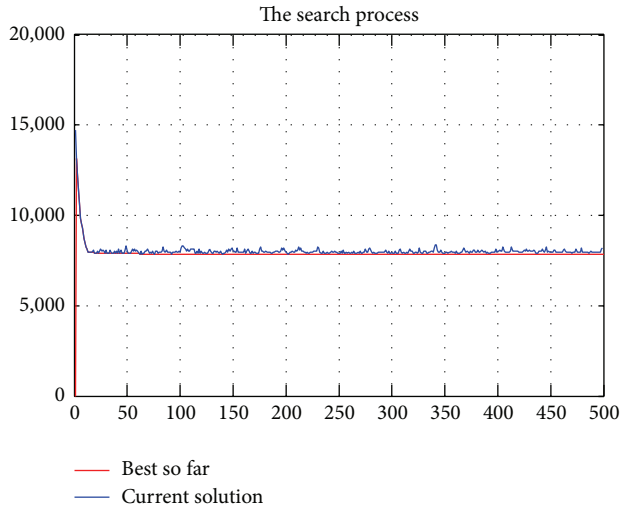



Figure 1: Procedure of tabu algorithm.

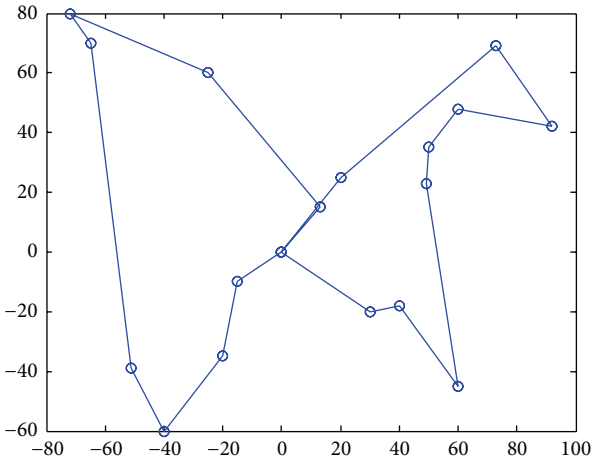



Figure 2: Optimum route.

From Figure 2, it can be found that in order to finish this distribution we need 2 vehicles in one time or use 1 vehicle to deliver two times. Also, there are no intersections in these two routes, which means that these two routes are intuitively preferable.

## 7. Conclusion

Unified and professional medicine distribution for communities can effectively integrate and allocate the already scarce medical resources and thus is of great significance to provide more affordable and easily accessible medical and health services to citizens. Considering the special characteristics of community medicine distribution, the paper establishes a model to optimize the community medicine distribution route, with the objective of minimizing the overall distribution cost (including storage cost, vehicle consuming cost, and transportation cost) and with the load capacity of vehicles and the maximum transport distance each time as constraint conditions. Then, a tabu-search-based algorithm originated from costumer direct arrangement ideology is put forward and MATLAB language is used to simulate the medicine distribution procedure. The simulated results show that the proposed algorithm can effectively solve the medicine distribution route optimization issues with uncertain number of vehicles and provide a scheme with minimum distribution cost for community health service institution. Compared to conventional distribution scheme which depends on experience and intuition, the proposed algorithm can reduce the overall distribution distance and delivery time, thus lowering the overall cost and improving the quality of medicare service. In addition, the algorithm can adjust the route readily to the changes in initial values, constraint conditions, and the amount of medicine requested by community health service institutions, so it has good adaptability in practice. It is believed that the advantage of the optimized distribution route will stand out with the increasing expansion of the community health service institutions. However, the road conditions, traffic jams, and weather are not considered in this study, which will become the focus of future research.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

The financial support from National Natural Science Foundation of China (Grant no. 71573109 for Professor Zhiqiang Ma), Social Sciences Foundation of Jiangsu Province in China (Grant no. 13GLB009 for Professor Zhiqiang Ma), the Doctoral Program of Higher Education Research Foundation from Education Ministry of China (Grant no. 20123227110023 for Professor Zhiqiang Ma), and the Innovation Research Foundation for Graduate Students of Jiangsu Province in China (Grant no. KYZZ_0282 for Dr. Zhao Li) is acknowledged.

## References

[1] Z. Wang and L. Yao, "Discussion on the medicine distribution of community health service institution," *China Logistics and Procurement*, vol. 30, no. 13, pp. 33–36, 2009.

[2] Y. Zhao, "Research on the construction and management of community pharmaceutical supply chain," *China Circulation Economy*, vol. 23, no. 6, pp. 33–36, 2009.

[3] Z. Wang, *Research on the supply mode of community fundamental medicine based on the management of supply chain [Ph.D. thesis]*, Huazhong University of Science and Technology, Wuhan, China, 2010.

[4] J. Zhang and K. Zhang, "Application of ant colony algorithm in medical logistics distribution routing optimization," *Journal of Taiyuan University of Technology*, vol. 40, no. 6, pp. 600–603, 2009.

[5] G. Zheng, X. Zhou, and L. Zhang, "Model and algorithm for dynamic optimization of city pharmaceutical distribution route with time window," *Journal of Railway Science and Engineering*, vol. 8, no. 4, pp. 80–85, 2011.

[6] Y. Liu and P. Zhang, "Medicines direct supply distribution solutions," *China Logistics and Procurement*, vol. 32, no. 13, pp. 44–45, 2011.

[7] F. Glover, "Future paths for integer programming and links to artificial intelligence," *Computers & Operations Research*, vol. 13, no. 5, pp. 533–549, 1986.

[8] F. Glover, "Tabu search—part I," *ORSA Journal on Computing*, vol. 1, no. 3, pp. 190–206, 1989.

[9] F. Glover, "Tabu search: part II," *ORSA Journal on Computing*, vol. 2, no. 1, pp. 4–32, 1990.

[10] M. Lang and X. Hu, "Research on the tabu search algorithm of vehicle routing problem," *Journal of Management Engineering*, vol. 18, no. 1, pp. 81–84, 2004.

[11] L. Wang, *Intelligent Optimization Algorithm and Its Application*, Tsinghua University Press, Beijing, China, 2001.

[12] R. Liu, X. Xie, V. Augusto, and C. Rodriguez, "Heuristic algorithms for a vehicle routing problem with simultaneous delivery and pickup and time windows in home health care," *European Journal of Operational Research*, vol. 230, no. 3, pp. 475–486, 2013.

[13] A. D. Legües, J. A. Ferland, C. C. Ribeiro, J. R. Vera, and A. Weintraub, "A tabu search approach for solving a difficult forest harvesting machine location problem," *European Journal of Operational Research*, vol. 179, no. 3, pp. 788–805, 2007.

[14] D. Taş, N. Dellaert, T. van Woensel, and T. de Kok, "Vehicle routing problem with stochastic travel times including soft time windows and service costs," *Computers and Operations Research*, vol. 40, no. 1, pp. 214–224, 2013.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 4313247, 10 pages
http://dx.doi.org/10.1155/2016/4313247

*Research Article*

# Multiple Attribute Decision Making Based on Cross-Evaluation with Uncertain Decision Parameters

**Tao Ding,[1] Liang Liang,[2] Min Yang,[1] and Huaqing Wu[3]**

[1]*School of Management, Hefei University of Technology, 193 Tunxi Road, Hefei, Anhui 230009, China*
[2]*Hefei University of Technology, 193 Tunxi Road, Hefei, Anhui 230009, China*
[3]*School of Economics, Hefei University of Technology, 193 Tunxi Road, Hefei, Anhui 230009, China*

Correspondence should be addressed to Min Yang; yangmin@hfut.edu.cn

Received 27 December 2015; Accepted 31 March 2016

Academic Editor: Victor Shi

Copyright © 2016 Tao Ding et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Multiple attribute decision making (MADM) problem is one of the most common and popular research fields in the theory of decision science. A variety of methods have been proposed to deal with such problems. Nevertheless, many of them assumed that attribute weights are determined by different types of additional preference information which will result in subjective decision making. In order to solve such problems, in this paper, we propose a novel MADM approach based on cross-evaluation with uncertain parameters. Specifically, the proposed approach assumes that all attribute weights are uncertain. It can overcome the drawback in prior research that the alternatives' ranking may be determined by a single attribute with an overestimated weight. In addition, the proposed method can also balance the mean and deviation of each alternative's cross-evaluation score to guarantee the stability of evaluation. Then, this method is extended to a more generalized situation where the attribute values are also uncertain. Finally, we illustrate the applicability of the proposed method by revisiting two reported studies and by a case study on the selection of community service companies in the city of Hefei in China.

## 1. Introduction

Multiple attribute decision making (MADM) is an important part of modern decision science [1–4]. It assumes that there exists a set of alternatives with multiple attributes which a decision maker (DM) should evaluate and analyze. The aim of MADM is to find the most desirable alternative or rank the feasible alternatives for supporting decision makings. As an active research area, MADM problems have been tried to be solved by some classical methods such as the simple additive weighting method (SAW), the analytic hierarchy process (AHP), and the technique for order preference by similarity to ideal solution (TOPSIS) [5–7]. It also has been extensively applied to many aspects such as economics, management, engineering, and technology [8–15].

In MADM problems, the needed decision making information includes attribute values and attribute weights. The attribute values describe characteristics, qualities, and performances of alternatives. The attribute weights are used to measure the importance of attributes. Due to the complexity of the real world and limited knowledge and perception capability of human beings, both attribute values and attribute weights are uncertain [16, 17]. The decision makers are not able to express their preferences or assessments explicitly [18]. In order to deal with such problems, many researchers try to describe attribute values and attribute weights accurately in uncertain environments. Generally, the researches in this area can be classified into three categories as follows.

Firstly, fuzzy set theory that appeared in 1965 has gradually become the mainstream in the field of representing and handling uncertain attribute values [19]. Since the inception of the fuzzy set theory, various extensions have been proposed, such as interval-valued fuzzy set [20], fuzzy multiset [21], intuitionistic fuzzy set (IFS) [22], linguistic fuzzy set [23], and hesitant fuzzy set (HFS) [24].

Secondly, several approaches have been proposed to determine attribute weights. Generally, they fall into three categories: subjective, objective, and hybrid methods [25]. The subjective methods utilize the preferences of a decision

maker to select attribute weights [26–28]. The objective methods determine attribute weights based on the objective information [29, 30]. Hybrid methods [31–34] combine both the preferences of a decision maker and objective information to determine attribute weights.

Thirdly, as a nonparametric programming efficiency-rating technique for evaluating a set of decision making units (DMUs) with multiple inputs and outputs, date envelopment analysis (DEA) requires relatively less additional information [35]. DEA has been considered as an effective tool to solve MADM problems without information on weights. For instance, Doyle [36] presents an approach based on DEA method, which derives a set of attribute weights from the data itself. Bernroider and Stix [37] introduce an approach based on DEA with restricted multipliers for accountable and understandable MADM. Wu and Liang [38] develop a new multiple attribute ranking method based on DEA game cross-efficiency model, in which each alternative is viewed as a player that seeks to maximize its own performance.

Although existing methods handle uncertainty to some extent, some drawbacks are also exposed which are listed as follows. First, both different defuzzy approaches and weight determination methods may generate different attribute weights and produce different optimal solutions [39, 40]. Therefore, it is difficult to identify which optimal solution is the best one. Second, behavior factors are not considered in existing MADM methods which may lead to deviation results in decision making. Specially, they do not avoid the bad situation that the ranking results may be determined by a single attribute with an exorbitant weight. In other words, the alternative rank order is consistent with the rank order of overestimated attribute which means that other attributes are totally useless in evaluation. Obviously, such result is unfair and unreasonable because the optimal alternative should be determined by multiple attributes instead of an individual attribute. Third, although DEA models can evaluate MADM problems without weight information, it also suffers from the following three drawbacks. One is that the optimal weights derived by DEA may not be consistent with the opinions of the decision maker [41]. Another shortage is that DEA identifies too many efficient alternatives and cannot well distinguish them [42, 43]. Furthermore, the nonuniqueness solution phenomenon is also common in DEA based MADM methods [44].

Considering the above analysis, we attempt to develop a cross-evaluation based approach to solve MADM problems with uncertain decision information. To be specific, we firstly define and analyze the attribute weight space which contains all feasible weights. Specially to avoid the single dominating attribute situation, we calculate the upper bound weight of each attribute. Then, considering a condition where the attribute values are certain and the attribute weights are uncertain, the return and risk of closeness indexes are calculated based on cross-evaluation in which each alternative seeks to maximize its closeness index by selecting a set of optimal weights. And the coefficient of variation is utilized to integrate the results and produce an alternative rank order. In a more generalized condition where the attribute values are denoted by interval values or ordinal values, the cross-evaluation models considering DM's risk preference are developed.

The main contributions of the current paper can be summarized as follows: (1) we firstly defined the attribute weight space and determined the upper bound weight of each attribute to conduct the uncertain information; (2) both mean and deviation measures of closeness index under cross-evaluation are considered to integrate a unique solution to the MADM problem; (3) the proposed method is extended to more generalized situations where the attribute values are described by interval values and ordinal values, respectively.

The remainder of this paper is organized as follows. In Section 2, we briefly review the relative works. In Section 3, we propose a cross-evaluation based method to solve MADM problems with uncertain weight information. In Section 4, the proposed method is extended to more generalized forms where the attribute information is expressed by interval values and ordinal values. Section 5 illustrates the methods with reported MADM studies and a case study of the selection of community service companies. Conclusions and future research will be displayed in Section 6.

## 2. Preliminaries

*2.1. Formulation of MADM Problems.* Suppose $X = \{X_1, X_2, \ldots, X_m\}$ is a discrete set of alternatives and $Y = \{Y_1, Y_2, \ldots, Y_n\}$ is a set of $n$ attributes with the weight vector $w = (w_1, w_2, \ldots, w_n)^T$, where $w_j > 0$, $\sum_{j=1}^{n} w_j = 1$. Let $A = (a_{ij})_{m\times n}$ be a decision matrix, where $a_{ij}$ denotes the attribute value that the $i$th alternative $X_i \in X$ takes with respect to the $j$th attribute $Y_j \in Y$.

In general, there are both benefit type attributes (i.e., the larger the better) and cost type attributes (i.e., the smaller the better) in a MADA problem. We transform the original decision matrix $A = (a_{ij})_{m\times n}$ into a normalized matrix $R = (r_{ij})_{m\times n}$ by using the method from Hwang and Yoon [7]:

$$r_{ij} = \begin{cases} \dfrac{a_{ij}}{\left\{\max_i a_{ij}\right\}}, & \text{for benefit attribute } Y_j \\ \dfrac{\left\{\min_i a_{ij}\right\}}{a_{ij}}, & \text{for cost attribute } Y_j, \end{cases} \quad (1)$$

$$i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n.$$

*2.2. Characteristics of the TOPSIS Method.* The basic principle of TOPSIS method is that the optimal alternative should have both the shortest distance from the positive ideal solution and the furthest distance from the negative ideal solution [14].

The TOPSIS procedure is performed in six stages as follows:

(a) Normalize the decision matrix

$$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{j=1}^{n} a_{ij}^2}}, \quad i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n, \quad (2)$$

where $a_{ij}$ is the $j$th attribute value of alternative $i$.

(b) Calculate the weighted normalized decision matrix

$$v_{ij} = w_j r_{ij}, \quad i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n, \tag{3}$$

where $w_j$ is the weight of the $j$th attribute, and $\sum_{j=1}^{n} w_j = 1$.

(c) Determine the positive ideal and negative ideal solution

$$\text{Positive ideal solution } A^+ = \{v_1^+, v_2^+, \ldots, v_n^+\} = \left\{\left(\max_i v_{ij} \mid j \in I'\right), \left(\min_i v_{ij} \mid j \in I''\right)\right\},$$
$$\text{Negative ideal solution } A^- = \{v_1^-, v_2^-, \ldots, v_n^-\} = \left\{\left(\min_i v_{ij} \mid j \in I'\right), \left(\max_i v_{ij} \mid j \in I''\right)\right\}, \tag{4}$$

where $I'$ indicate the index set of the benefit type attributes and $I''$ the index set of the cost type attributes.

(d) Measure the distances from the positive ideal solution and negative ideal solution, respectively:

$$D_i^+ = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^+\right)^2}, \quad i = 1, 2, \ldots, m,$$
$$D_i^- = \sqrt{\sum_{j=1}^{n} \left(v_{ij} - v_j^-\right)^2}, \quad i = 1, 2, \ldots, m. \tag{5}$$

(e) Calculate the relative closeness to the ideal solutions

$$C_i^* = \frac{D_i^-}{\left(D_i^+ + D_i^-\right)}, \quad i = 1, 2, \ldots, m. \tag{6}$$

(f) Rank the alternatives.

The larger the closeness index of an alternative, the higher the rank of the alternative.

*2.3. DEA Cross-Efficiency Evaluation.* We assume that there are $n$ DMUs to be evaluated in terms of $m$ inputs and $n$ outputs. The $i$th input and $r$th output of $\text{DMU}_j$ are denoted by $x_{ij}$ $(i = 1, \ldots, m)$ and $y_{rj}$, respectively. The efficiency of a $\text{DMU}_d$ can be calculated by the following CCR model in linear programming formulation [35]:

$$\begin{aligned}
\max \quad & \sum_{r=1}^{s} u_r y_{rd} = \theta_d \\
\text{s.t.} \quad & \sum_{i=1}^{m} w_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0, \quad j = 1, 2, \ldots, n, \\
& \sum_{i=1}^{m} w_i x_{id} = 1, \\
& w_i \geq 0, \quad i = 1, 2, \ldots, m, \\
& u_r \geq 0, \quad r = 1, 2, \ldots, s.
\end{aligned} \tag{7}$$

When the above model is solved, a set of optimal weights $w_{1d}^*, \ldots, w_{md}^*, u_{1d}^*, \ldots, u_{sd}^*$ is obtained. Then the $d$ cross-efficiency of any $\text{DMU}_j$ is computed as [38]

$$E_{dj} = \frac{\sum_{r=1}^{s} u_{rd}^* y_{rj}}{\sum_{i=1}^{m} w_{id}^* x_{ij}}, \quad d, j = 1, \ldots, n. \tag{8}$$

For $\text{DMU}_j$ $(j = 1, \ldots, n)$, the average of all $E_{dj}$ $(d = 1, \ldots, n)$, denoted by $\overline{E}_j = (1/n) \sum_{d=1}^{n} E_{dj}$, is considered as the cross-efficiency score of $\text{DMU}_j$.

## 3. Proposed Method

In this section, we will introduce the so-called single dominating attribute issue and propose a model to support decisions by incorporating considerations on this issue in Section 3.1; furthermore, in Section 3.2, we will introduce the cross-evaluation framework and the accompanying risk consideration; finally, in Section 3.3, we propose a procedure summarizing all the aforementioned considerations for MADM.

*3.1. Determination of the Weight Space.* We observe that, in MADM problems, an individual attribute can determine the final ranks of all alternatives when the corresponding weight is greater than some threshold value. In other words, the rank order of the alternatives is completely in accordance with the rank order of the dominating attribute and the rest of the attributes are totally ignored in the evaluation. It should be noted that the ignored attributes are taken into the evaluation only because the decision maker regards them as import factors or they could have excluded them from the decision process in the first place. Therefore, we argue this is unacceptable. In order to address this issue, it is necessary to define and compute the threshold upper bound of each attribute weight. The upper bound can aid the decision process by suggesting meaningful weights.

For the simplicity of exposition, if the attribute values $r_{ik}$ $(i = 1, \ldots, m)$ in the $k$th column of the decision matrix $R$ have the relation $r_{i_1 k} \geq r_{i_2 k} \geq \cdots \geq r_{i_m k}$ $(i_s \in \{1, m\}, \ s = 1, \ldots, m)$, then we denote the ranks of the alternatives according to the values of the $k$th attribute $Y_k$ as $X_{i_1} \succ_k X_{i_2} \succ_k \cdots \succ_k X_{i_m}$. With this notation in position, the single dominating attribute happens if the final rank order of all alternates is the same as some $X_{i_1} \succ_k X_{i_2} \succ_k \cdots \succ_k X_{i_m}$ no matter how the decision maker distributes the relative weights of the other attributes. We formulize this notion as Definition 1 as below.

*Definition 1.* Let $k$ be the index of some attribute. $w_k^*$ is the threshold upper bound of $w_k$ if and only if the following conditions are satisfied:

(a) If $w_k \geq w_k^*$, then $\forall 1 \leq i, \ t \leq m, \ r_{ik} \geq r_{tk}$ if and only if $\sum_{j=1}^{n} w_j r_{ij} \geq \sum_{j=1}^{n} w_j r_{tj}$.

(b) If $w_k < w_k^*$, then $\exists 1 \leq i, \ t \leq m$, such that $r_{ik} \geq r_{tk}$ and $\sum_{j=1}^{n} w_j r_{ij} \leq \sum_{j=1}^{n} w_j r_{tj}$ hold true simultaneously.

To implement Definition 1, we provide Theorem 2 below to determine the threshold upper bound.

**Theorem 2.** *The threshold upper bound of the kth attribute ($w_k^*$) is given by the minimization program*

$$\begin{aligned}
\min \quad & w_k \\
\text{s.t.} \quad & w_k r_{i_1 1} + \sum_{j=1, j\neq k}^{n} w_j^{(f)} r_{i_1 j} \geq w_k r_{i_2 1} + \sum_{j=1, j\neq k}^{n} w_j^{(f)} r_{i_2 j} \\
& \geq \cdots \\
& \geq w_k r_{i_m 1} + \sum_{j=1, j\neq k}^{n} w_j^{(f)} r_{i_m j}
\end{aligned} \tag{9}$$

$$\begin{aligned}
& \sum_{j=1}^{n} w_j = 1, \\
& w_j \geq 0, \\
& j = 1, 2, \ldots, n, \\
& f = 1, 2, \ldots, n-1,
\end{aligned} \tag{10}$$

*where $w_j^{(f)} = 1 - w_k$ if $j$ is the $f$th element of the set $\{1, \ldots, N\} \setminus \{k\}$ in ascending order; otherwise, $w_j^{(f)} = 0$. Let us take $k = 1$ as an example. We have $(w_2^{(1)}, w_3^{(1)}, \ldots, w_n^{(1)}) = (1 - w_1, 0, \ldots, 0)$, $(w_2^{(2)}, w_3^{(2)}, \ldots, w_n^{(2)}) = (0, 1 - w_1, 0, \ldots, 0)$, $\ldots, (w_2^{(n-1)}, w_3^{(n-1)}, \ldots, w_n^{(n-1)}) = (0, \ldots, 0, 1 - w_1)$. The subscript $i_h$ ($h \in \{1, \ldots, m\}$) of $r$ in the above model indicates the $i_h$ preferred alternative determined by the preference relation $X_{i_1} \succ_k X_{i_2} \succ_k \cdots \succ_k X_{i_m}$.*

*Proof.* Clearly, there are $(n - 1)$ constraints in (9). When $k = 1$, the weight combinations $(w_2^{(1)}, w_3^{(1)}, \ldots, w_n^{(1)}) = (1 - w_1, 0, \ldots, 0)$, $(w_2^{(2)}, w_3^{(2)}, \ldots, w_n^{(2)}) = (0, 1 - w_1, 0, \ldots, 0)$, $\ldots, (w_2^{(n-1)}, w_3^{(n-1)}, \ldots, w_n^{(n-1)}) = (0, \ldots, 0, 1 - w_1)$ are applied to the $n-1$ constraints in order. The first constraint guarantees that the alternatives rank result has nothing to do with attribute $Y_2$. Similarly, the remaining $n - 2$ constraints ensure that the alternatives rank result has nothing to do with the other attributes $Y_3, \ldots, Y_n$. So the optimal minimization $w_1$ satisfies (a) in Definition 1.

Conversely, we suggest that a constraint belonging to the $l_{\text{th}}$ ($l = 1, \ldots, n - 1$) group is removed from the model. Then, attribute $Y_l$ may play a part in the ranking because it is not completely suppressed. In other words, (b) in Definition 1 holds.

Similarly, the upper bound of each attribute weight ($w_k^*$, $k = 2, \ldots, n$) can be calculated by the model in Theorem 2 with $k = 2, \ldots, n$. This proof is completed. □

*Example 3.* Ma et al. [31] consider a case of a robot user expecting to select a robot and there are four alternatives for him to choose. There are four attributes to evaluate them, that is, cost, velocity, repeatability, and load capacity. The normalized decision matrix is as follows:

$$A = \begin{pmatrix} 0 & 1 & \frac{5}{6} & \frac{1}{3} \\ \frac{5}{12} & \frac{3}{5} & 1 & 0 \\ 1 & 0 & 0 & 1 \\ \frac{2}{3} & \frac{2}{5} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}. \tag{11}$$

Suppose that the information on attribute weights is completely unknown. The upper bound weight of the first attribute cost can be calculated by the above model.

By calculating the above model provided in Theorem 2, we obtain the optimal value $w_1^* = 2/3$ denoting the threshold upper bound of the first attribute $w_1$. Analogously, each upper bound of attribute weight can be calculated by similar models, and the values are $w_2^* = 10/13$, $w_3^* = 4/5$, $w_4^* = 9/14$.

*3.2. Mean and Deviation of Closeness Index Based Cross-Evaluation Models.* In this subsection, in the environment of uncertain attribute weights, we develop cross-evaluation models to calculate the mean and deviation of closeness index for each assessed alternative, respectively. Then, the computed indexes are integrated to obtain the evaluation results.

Cross-efficiency evaluation method, since being proposed by Sexton et al. [45], has been used to evaluate and rank DMUs. Its main idea is to use cross-evaluation instead of a self-evaluation. There are at least two advantages for cross-evaluation method. One is that it obtains a unique ranking among the DMUs. The other is that the cross-efficiency means can act effectively to differentiate between good and bad alternatives [46]. Hence, the cross-evaluation method is widely used for ranking performance of alternatives, for example, equipment design [47], supplier selection [48], and stock market [49].

To calculate the self-evaluation closeness index for each alternative, we consider the following mathematical programming problem for a given alternative $X_i$:

$$\begin{aligned}
\max \quad & \frac{D_i^-}{D_i^- + D_i^+} = C_i^{\max} \\
& w \in \Omega(w).
\end{aligned} \tag{12}$$

Here, $D_i^-$ and $D_i^+$ represent the distance from negative ideal point and ideal point (defined in (5)), respectively. For ease of calculation, the ideal point and negative point are defined as $A^+ = \{w_1, w_2, \ldots, w_n\}$ and $A^- = \{0, 0, \ldots, 0\}$. Then $D_i^+$ and $D_i^-$ are denoted by $D_i^+ = \sqrt{\sum_{j=1}^{n} (w_j r_{ij} - w_j)^2}$ and $D_i^- = \sqrt{\sum_{j=1}^{n} (w_j r_{ij})^2}$, respectively. The constraint $w \in \Omega(w)$ depicts the possible set of attribute weights. Obviously, when the weights information of attributes is completely unknown, $\Omega(w) = \{w_j > 0, \ \sum_{j=1}^{n} w_j = 1\}$. Introducing the upper bound weight information calculated in Section 3.1, $\Omega(w) = \{w_j < w_j^*, \ \sum_{j=1}^{n} w_j = 1\}$.

In many practical situations, although the weight values are difficult to obtain, the weight information is partially known and predefined by the weight space. As follows, we introduce two types of partial weight information as references to decision maker.

*Type 1* (pairwise comparison weights information). In MADM, attribute weights $w_j$ are mainly introduced to reflect the importance to attributes $Y_j$ ($j = 1, 2, \ldots, n$). As stated in Section 1, pairwise comparison is a widely accepted subjective method to determine attribute weight.

Attribute $Y_2$ is at least two times and not more than three times as important as attribute $Y_1$.

In notational jargons, this is denoted by

$$2w_1 \le w_2 \le 3w_1. \tag{13}$$

In general terms, we compare $Y_1$ with $Y_j$ ($j = 2, 3, \ldots, n$), respectively. The bound linear constraints are denoted by

$$l_j w_1 \le w_j \le u_j w_1, \tag{14}$$

in which $l_j$ and $u_j$ represent the lower and upper bounds of times $w_j$ compared with $w_1$, respectively. Hence, the attribute weights space is obtained by the above pairwise comparison information, which is expressed as $\Omega(w) = \{w_j > 0,\ l_j w_1 \le w_j \le u_j w_1,\ j = 1, 2, \ldots, n\}$. Clearly, $l_1 = u_1 = 1$ holds.

Compared with conventional subjective weight determination approaches such as the AHP, weights space is more feasible and applicative because the decision makers need not express their preferences explicitly. In many situations, experts are pleased to give such interval information rather than exact value due to the complexity of actual problems.

*Type 2* (ordinal weights information). Outranking relation [50], usually denoted by $S$, was proposed by Roy whose aim was to establish a realistic representation of the three basic preference situations: strict preference, indifference, and incomparability. Motivated by the widely accepted outranking relation, we introduce the concept of rank-order weights information. In many situations, although the weight values are difficult to give by decision makers, the ranking of each attribute is easy to obtain according to importance. Take the Olympic Games as an example, the achievements of athletes and nations are judged by gold, silver, and bronze medals. Apparently, a gold medal is more important than a silver medal, and a silver medal is more important than a bronze medal. Then, the weights space can be denoted by $\Omega(w) = \{w_1 > w_2 > w_3,\ w_1, w_2, w_3 > 0\}$, in which $w_1, w_2, w_3$ represent the weight of gold, silver, and bronze medal, respectively.

The above self-evaluation model (see (12)) shows that all alternatives choose their own attribute weights in order to maximize their own self-evaluation closeness index. Similar to the form of DEA cross-efficiency, we obtain a set of optimal attribute weights $W_i^* = (w_{i1}^*, \ldots, w_{in}^*)$ according to the above model. When these optimal weights are applied to evaluate alternative $X_k$, a cross score denoted by $C_{ik} = D_{ik}^- / (D_{ik}^- + D_{ik}^+)$ is obtained.

For alternative $k$ ($k = 1, \ldots, m$), we use the average of all $C_{ik}$ to measure the expected return of cross-evaluation closeness index, which is written as

$$r(X_k) = \frac{1}{m} \sum_{i=1}^{m} C_{ik}. \tag{15}$$

Furthermore, we use the standard variance

$$\delta(X_k) = \sqrt{\frac{\sum_{i=1}^{m} (C_{ik} - r(X_k))^2}{n}} \tag{16}$$

to measure the risk caused by different sets of optimal weights produced by each alternative.

In general, there are both return of closeness index (defined by the mean of cross-evaluation score) and risk of closeness index (defined by the deviation between return and each cross-evaluation closeness index) in cross-evaluation framework. It is not hard to see that the larger return score the better and the smaller risk score the better. Specifically, given two alternatives $X_s$ and $X_t$, if $R(X_s) > R(X_t)$ and $\delta(X_s) < \delta(X_t)$, then $X_s \succ X_t$ ($X_s$ is better than $X_t$) clearly holds. However, when $R(X_s) > R(X_t)$ and $\delta(X_s) > \delta(X_t)$, it is not easy to compare the two alternatives directly.

Then, we integrate the return and risk measures by the well-known coefficient of variation [51]

$$\text{C.V}(X_k) = \frac{\delta(X_k)}{r(X_k)} \times 100\%. \tag{17}$$

For two given alternatives, one having a smaller coefficient of variation is considered to be better than the other one. Thus, all the alternatives are ranked by the comparison of their corresponding coefficient of variation.

*3.3. Procedure of the Proposed Approach.* Based on the above analysis, the procedure for generating a solution to the MADM problem is elaborated in the following.

*Step 1* (formulation of a MADM problem). The decision maker identifies $n$ attributes to evaluate $m$ alternatives. The original attribute values are standardized by (1).

*Step 2* (obtainment of the attribute weight space). First, the upper bounds of each attribute are calculated by Theorem 2 to guarantee the effectiveness of all the attributes. Then, different types of constraint of attribute weights are determined by the decision maker to reflect the importance of each attribute. Last, the weight space including all the feasible weight combinations is obtained.

*Step 3* (calculation of the mean and deviation measures). Firstly, the maximized closeness index of each alternative is calculated by (12). Secondly, the mean measure of each alternative's cross closeness index is computed by (15). Third, the deviation measure of each alternative's cross closeness index is obtained by (16).

*Step 4* (evaluation of the alternatives via coefficient of variation). Since the mean and deviation measures of each alternative's cross closeness index are obtained in Step 3, the coefficient of variation defined in (17) is utilized to integrate the results.

*Step 5* (generation of rank order of the alternatives). Through comparing the coefficients of variation among the alternatives, the rank order can be directly obtained.

## 4. Extension with Uncertain Attribute Values

*4.1. Model with Interval Attribute Value.* It is widely acknowledged that uncertainty is associated with attribute values due to complexity of real world and limited knowledge of human beings. In many situations, the attribute value $r_{ij}$ is uncertain, but we know that the feasible attribute value $r_{ij} \in [r_{ij}^-, r_{ij}^+]$. Here, we present the following self-evaluation model with interval attribute values [52] and uncertain attribute weights:

$$\begin{aligned}
\max \quad & \alpha \frac{\sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}^+\right)^2}}{\sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}^+\right)^2} + \sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}^+ - w_j\right)^2}} \\
& + \beta \frac{\sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}^-\right)^2}}{\sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}^-\right)^2} + \sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}^- - w_j\right)^2}} \qquad (18) \\
& = C_i^{\max} \\
& w \in \Omega(w), \\
& \alpha + \beta = 1, \quad \alpha \geq 0, \ \beta \geq 0,
\end{aligned}$$

in which $\sqrt{\sum_{j=1}^{n}(w_j r_{ij}^+)^2}$, $\sqrt{\sum_{j=1}^{n}(w_j r_{ij}^+ - w_j)^2}$, $\sqrt{\sum_{j=1}^{n}(w_j r_{ij}^-)^2}$, and $\sqrt{\sum_{j=1}^{n}(w_j r_{ij}^- - w_j)^2}$ represent distance of upper bounds from ideal point, distance of upper bounds from negative ideal point, distance of lower bounds from ideal point, and distance of lower bounds from negative ideal point, respectively. $\alpha$ and $\beta$ reflect the risk coefficients that combine upper closeness index and lower closeness, respectively. In practice, $\alpha$ and $\beta$ are decided by the decision maker according to his or her risk preference. Specifically, the decision maker is neutral if $\alpha = \beta = 1/2$; $\alpha > \beta$ and $\alpha < \beta$ indicate that he is optimistic and pessimistic, respectively.

*4.2. Model with Ordinal Attribute Value.* Another form of uncertain preferential information is expressed by ordinal attribute value (e.g., "more preferable" and "less preferable"). It is argued to be more stable and more reliable than cardinal value in some situations [53]. In this paper, an approach of threshold value is introduced to transform ordinal value into linear formulation. For instance, given four attributes $a$, $b$, $c$, and $d$, the ordinal values are 1, 2, 3, and 4, respectively. By specifying a threshold value 0.1, the attribute values can be converted to a cardinal attribute space $\Omega(a, b, c, d) = \{a = 1, a - b \geq 0.1, b - c \geq 0.1, c - d \geq 0.1, d = 0\}$. In general, let the ordinal attribute values $a_{ij}$ $(i = 1, \ldots, m)$ have relation $a_{i_1 j} > a_{i_2 j} > \cdots > a_{i_m j}$ $(i_k \in [1, m],\ k = 1, \ldots, m)$, and then the cardinal attribute space $\Omega(r_{ij}) = \{a_{i_m j} = 1, a_{i_{m-1} j} - a_{i_m j} \geq \varepsilon, a_{i_{m-2} j} - a_{i_{m-1} j} \geq \varepsilon, \ldots, a_{i_2 j} - a_{i_1 j} \geq \varepsilon, a_{i_1 j} = 0, \varepsilon > 0\}$ is obtained ($\varepsilon$ is the threshold value given by the decision maker). Here, we present the following self-evaluation model with ordinal attribute values and uncertain attribute weights:

$$\begin{aligned}
\max \quad & \frac{\sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}\right)^2}}{\sqrt{\sum_{j=1}^{n} \left(w_j r_{ij}\right)^2} + \sqrt{\sum_{j=1}^{n} \left(w_j r_{ij} - w_j\right)^2}} \\
& = C_i^{\max} \qquad (19) \\
\text{s.t.} \quad & r \in \Omega(r) \\
& w \in \Omega(w), \\
& i = 1, \ldots, m; \qquad (20) \\
& j = 1, \ldots, n. \qquad (21)
\end{aligned}$$

For each alternative $X_k$ $(k = 1, \ldots, m)$ under evaluation, we obtain a group of optimal attribute values and a set of attribute weights. They are denoted by $A_k^* = (r_{ij}^k)_{m \times n}$ and $W_k^* = (w_1^k, w_2^k, \ldots, w_n^k)$, respectively. When the optimal decision information is applied to evaluate alternative $X_l$ $(l = 1, \ldots, m)$, a cross score denoted by $C_{kl} = \sqrt{\sum_{j=1}^{n}(w_j^k r_{lj}^k)^2} / (\sqrt{\sum_{j=1}^{n}(w_j^k r_{lj}^k)^2} + \sqrt{\sum_{j=1}^{n}(w_j^k r_{lj}^k - w_j^k)^2})$ is obtained.

Similar to that defined by (14) and (15), the mean and deviation measures of alternative $X_l$ are denoted by $r(X_l) = \sum_{k=1}^{m} C_{kl}$ and $\delta(X_l) = \sqrt{\sum_{k=1}^{m}(C_{kl} - r(X_l))^2 / n}$, respectively.

Note that (12), (18), and (19) are nonlinear models; they can be transformed into linear forms according to practical cases.

## 5. Illustrative Examples

In order to illustrate the applications of our proposed methods, three practical MADM problems are presented.

*5.1. Application to House Selection with Cardinal Attribute Values.* Xu [54] stated a purchase house problem: there are four possible houses to be selected, and the attributes are as follows: (1) $Y_1$ is the house price (ten thousand yuan); (2) $Y_2$ is the usable area (square meter); (3) $Y_3$ is the distance between house and the center of city (kilometers); (4) $Y_4$ is the nature environment (scores), where $Y_2$ and $Y_4$ are benefit attributes and $Y_1$ and $Y_3$ are cost attributes. The attribute weights are completely unknown. The decision making matrix $A$ and normalized decision making matrix $R$ are listed in Tables 1 and 2, respectively. Step 1 is completed.

By utilizing (9) and (10), the upper bound of each attribute weight is obtained. And as a result, the attribute weight space is $\Omega(w) = \{0 < w_1 < 0.77, 0 < w_2 < 0.78, 0 < w_3 < 0.62, 0 < w_4 < 0.62, w_1 + w_2 + w_3 + w_4 = 1\}$. Step 2 is completed.

Table 1: Decision making matrix $A$ of house evaluation.

| | $Y_1$ | $Y_2$ | $Y_3$ | $Y_4$ |
|---|---|---|---|---|
| $X_1$ | 3.0 | 100 | 10 | 7 |
| $X_2$ | 2.5 | 80 | 8 | 5 |
| $X_3$ | 1.8 | 50 | 20 | 11 |
| $X_4$ | 2.2 | 70 | 12 | 9 |

Table 2: Normalized decision making matrix $R$ of house evaluation.

| | $Y_1$ | $Y_2$ | $Y_3$ | $Y_4$ |
|---|---|---|---|---|
| $X_1$ | 0.6 | 1.0 | 0.8 | 0.636 |
| $X_2$ | 0.72 | 0.8 | 1.0 | 0.455 |
| $X_3$ | 1.0 | 0.5 | 0.4 | 1.0 |
| $X_4$ | 0.818 | 0.7 | 0.667 | 0.818 |

Table 3: The cross-evaluation measurements.

| Alternatives | $C_i^{\max}$ | $C_{1i}$ | $C_{2i}$ | $C_{3i}$ | $C_{4i}$ | $r(X_i)$ | $\delta(X_i)$ |
|---|---|---|---|---|---|---|---|
| $X_1$ | 0.96 | 0.96 | 0.82 | 0.61 | 0.61 | 0.75 | 0.15 |
| $X_2$ | 0.92 | 0.80 | 0.92 | 0.66 | 0.66 | 0.76 | 0.11 |
| $X_3$ | 1 | 0.52 | 0.51 | 1 | 1 | 0.76 | 0.24 |
| $X_4$ | 0.818 | 0.70 | 0.70 | 0.818 | 0.818 | 0.76 | 0.06 |

Table 4: Alternative efficiency matrix.

| Alternatives | $X_1$ | $X_2$ | $X_3$ | $X_4$ |
|---|---|---|---|---|
| CV $(X_i)$ | 0.20 | 0.14 | 0.32 | 0.08 |

In terms of Step 3 in Section 3.3, we obtain the cross-evaluation measurements of each alternative, which are presented in Table 3. Step 3 is completed.

Based on Table 3 and the calculation of coefficient of variation, the results are shown in Table 4.

Therefore, we can rank the houses in accordance with the CV values presented in Table 4 to get a complete rank order: $X_4 \succ X_2 \succ X_1 \succ X_3$. And $X_4$ is the best house under evaluation. Step 4 is completed.

In order to understand our approach more clearly, we compare our approach with other three similar methods [54], namely, average attribute value method, maximizing deviation method, and information entropy method, respectively, with the same example. The corresponding results of such three methods are listed in Table 5.

According to Tables 4 and 5, the result derived from our cross-evaluation approach is quite different from that of the existing three weight determination methods. In fact, Table 5 shows that the three weight determination methods generate three attribute weight vectors and produce three different optimal solutions. However, there is no single method that can guarantee more convincing weight vector than others. Hence, we treat the attribute weight as a weight space rather than a certain value in this paper. The comparison result validates that our proposed approach can overcome the shortage that we cannot obtain the most dominant result caused by the diversity of the way of weight vector determined.

Table 5: The results of three methods.

| Method | Weight vector | Rank order |
|---|---|---|
| Average attribute value | (0.25, 0.25, 0.25, 0.25) | $X_1 \succ X_4 \succ X_2 \succ X_3$ |
| Maximizing deviation | (0.20, 0.24, 0.29, 0.27) | $X_1 \succ X_2 \succ X_4 \succ X_3$ |
| Information entropy | (0.13, 0.22, 0.36, 0.29) | $X_1 \succ X_2 \succ X_4 \succ X_3$ |

Table 6: Attribute weights.

| $Y_1$ | $Y_2$ | $Y_3$ |
|---|---|---|
| $[w_1^L, w_1^U]$ | $[w_2^L, w_2^U]$ | $[w_3^L, w_3^U]$ |
| [0.35, 0.50] | [0.15, 0.30] | [0.30, 0.35] |

Table 7: Attribute values of alternatives.

| | $Y_1$ | $Y_2$ | $Y_3$ |
|---|---|---|---|
| $A$ | [0.65, 0.85] | [0.68, 0.86] | [0.45, 0.78] |
| $B$ | [0.75, 1] | [0.55, 0.78] | [0.64, 0.80] |
| $C$ | [0.54, 0.68] | [0.46, 0.55] | [0.78, 1] |
| $D$ | [0.77, 0.90] | [0.67, 0.78] | [0.56, 0.77] |
| $E$ | [0.25, 0.45] | [0.89, 1] | [0.75, 0.94] |

*5.2. Application to University Faculty with Interval Attribute Value.* Consider the example in Bryson and Mobolurin's [55] study; that is, five faculty candidates ($A$, $B$, $C$, $D$, $E$) are evaluated for tenure and promotion through three attributes: teaching, research, and service. According to the DM's preference, the information for both attribute and alternatives is given by interval numbers, as shown in Tables 6 and 7, respectively (for the purpose of illustration, our data is different from the literature).

Table 8 shows the computed results of the proposed approach with interval values. The self-evaluation score of each alternative is listed in the second column and the cross-evaluation scores are listed in columns 3–7. Columns 8 and 9 show the mean and deviation measures, respectively. Then, the results in column 10 are applied to generate a rank order of the five candidates $A \succ C \succ E \succ B \succ D$, which is a solution to the problem.

*5.3. Application to the Selection of Community Service Companies with Ordinal Attribute Values.* In this subsection, we apply our approach to the evaluation and selection of five community service companies that are located in Hefei, a city in Anhui province of China. These companies provide various kinds of service for the public, such as property

Table 8: Evaluation results for the comparison of faculty candidates.

| Candidates | $C_i^{\max}$ | $C_{1i}$ | $C_{2i}$ | $C_{3i}$ | $C_{4i}$ | $C_{5i}$ | $r_i$ | $\delta_i$ | C.V $(X_i)$ |
|---|---|---|---|---|---|---|---|---|---|
| $A$ | 0.717 | 0.717 | 0.712 | 0.706 | 0.717 | 0.703 | 0.711 | 0.006 | 0.008 |
| $B$ | 0.792 | 0.792 | 0.792 | 0.788 | 0.792 | 0.746 | 0.782 | 0.018 | 0.023 |
| $C$ | 0.672 | 0.651 | 0.661 | 0.672 | 0.651 | 0.660 | 0.659 | 0.008 | 0.012 |
| $D$ | 0.774 | 0.774 | 0.771 | 0.764 | 0.774 | 0.737 | 0.764 | 0.014 | 0.461 |
| $E$ | 0.645 | 0.525 | 0.526 | 0.540 | 0.525 | 0.645 | 0.552 | 0.047 | 0.018 |

Table 9: Ordinal values of alternatives.

| | $Y_1$ | $Y_2$ | $Y_3$ | $Y_4$ |
|---|---|---|---|---|
| $X_1$ | 4 | 2 | 4 | 2 |
| $X_2$ | 1 | 4 | 3 | 3 |
| $X_3$ | 2 | 5 | 1 | 1 |
| $X_4$ | 5 | 3 | 2 | 5 |
| $X_5$ | 3 | 1 | 5 | 4 |

management, catering, health care, and entertainment. In the case study, the decision maker is an official from the development and reform commission of Hefei, with the help of a group of experts from the governments, relevant enterprises, and community organizations. Four attributes are used to evaluate the six companies. They are quality of service ($Y_1$), diversity of service ($Y_2$), profitability ($Y_3$), and development potential ($Y_4$). The original attribute values are given by ordinal values, as shown in Table 9. The decision procedures are as follows.

After considering the four attributes, the decision maker views that quality of service ($Y_1$) is the most important attribute in evaluation. The development potential ($Y_4$) is the second most important one, profitability ($Y_3$) is the third most important one, and diversity of service ($Y_2$) is the least important one, respectively. And the difference between two adjacent attribute weights should be more than 0.1. In addition, each weight value should be no more than 0.5. Then the weight space is denoted by $\Omega(w) = \{w_1 - 0.1 > w_4, w_4 - 0.1 > w_3, w_3 - 0.1 > w_2, w_1 \le 0.5, w_2 \le 0.5, w_3 \le 0.5, w_4 \le 0.5, w_1 + w_2 + w_3 + w_4 = 1\}$.

By specifying a threshold value of 0.1, ordinal attribute values are transformed into four sets of attribute value spaces. They are denoted by

$$\begin{aligned}
\Omega(Y_1) &= \{r_{21} = 1, r_{21} - r_{31} \ge 0.1, r_{31} - r_{51} \ge 0.1, r_{51} \\
&\quad - r_{11} \ge 0.1, r_{11} - r_{41} \ge 0.1, r_{41} = 0\}; \\
\Omega(Y_2) &= \{r_{52} = 1, r_{52} - r_{12} \ge 0.1, r_{12} - r_{42} \ge 0.1, r_{42} \\
&\quad - r_{22} \ge 0.1, r_{22} - r_{32} \ge 0.1, r_{32} = 0\}; \\
\Omega(Y_3) &= \{r_{33} = 1, r_{33} - r_{43} \ge 0.1, r_{43} - r_{23} \ge 0.1, r_{23} \\
&\quad - r_{13} \ge 0.1, r_{13} - r_{53} \ge 0.1, r_{53} = 0\}; \\
\Omega(Y_4) &= \{r_{34} = 1, r_{34} - r_{14} \ge 0.1, r_{14} - r_{24} \ge 0.1, r_{24} \\
&\quad - r_{54} \ge 0.1, r_{54} - r_{44} \ge 0.1, r_{44} = 0\}.
\end{aligned} \tag{22}$$

Table 10 shows the computed results of the proposed approach. The second column shows the self-evaluation score of each alternative computed by (19). And columns 3–7 list the cross-evaluation scores. Then the mean and deviation measures of each alternative are presented in columns 8 and 9, respectively. Consequently, the results in column 10 are applied to generate a rank order of the five community service companies $X_5 \succ X_1 \succ X_2 \succ X_4 \succ X_3$. The decision maker is pleased to select company $X_5$ with the best performance among the five alternatives.

## 6. Conclusions and Future Research

Due to the complexity and uncertainty of MADM problems in real world, the decision maker is not able to provide the exact attribute weights. And different weight determination methods are difficult to generate a most satisfaction solution. The current paper presents a new MADM method based on cross-evaluation, in which the weight information is expressed by the so-called weight space. It attempts to rectify some shortcomings of existing MADM methods by providing the following important features: (1) it is more feasible to provide a weight space rather than specifying a set of weight values; (2) the single dominating attribute drawback can be avoided by the restriction of the upper bound of each attribute; (3) compared with DEA based methods, the variable attribute weights are consistent with the opinion of decision maker; (4) both mean score and the deviation score are considered in our cross-evaluation models; and (5) the proposed method is also successfully extended to more generalized situations in which attribute values are expressed by interval values and ordinal values, respectively.

We have illustrated the applicability of the proposed methods by revisiting two reported MADM studies and by depicting a case study on the selection of community service companies. Although our proposed method makes some improvement in solving MADM problems, it only can deal with the individual decision making case. Future research

Table 10: Evaluation results for the selection of community service companies.

| Candidates | $C_l^{\max}$ | $C_{1l}$ | $C_{2l}$ | $C_{3l}$ | $C_{4l}$ | $C_{5l}$ | $r_i$ | $\delta_i$ | C.V $(X_i)$ |
|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | 0.489 | 0.489 | 0.317 | 0.295 | 0.317 | 0.317 | 0.347 | 0.072 | 0.207 |
| $X_2$ | 0.532 | 0.433 | 0.532 | 0.508 | 0.458 | 0.504 | 0.487 | 0.036 | 0.074 |
| $X_3$ | 0.535 | 0.526 | 0.526 | 0.535 | 0.526 | 0.526 | 0.528 | 0.004 | 0.008 |
| $X_4$ | 0.223 | 0.202 | 0.223 | 0.211 | 0.223 | 0.223 | 0.216 | 0.009 | 0.042 |
| $X_5$ | 0.422 | 0.349 | 0.349 | 0.103 | 0.349 | 0.422 | 0.314 | 0.109 | 0.347 |

may attempt to employ our cross-evaluation methods to handle group decision making problem.

## Competing Interests

The authors declare that they have no competing interests.

## Acknowledgments

The research is supported by National Key Technology Support Program (no. 2014BAH27F00), National Natural Science Funds of China (no. 71471053), and the Fundamental Research Fund for the Central Universities (no. JZ2014-HGBZ0339).

## References

[1] R. H. Green and J. R. Doyle, "On maximizing discrimination in multiple criteria decision making," *Journal of the Operational Research Society*, vol. 46, no. 2, pp. 192–204, 1995.

[2] R. E. Steuer and P. Na, "Multiple criteria decision making combined with finance: a categorized bibliographic study," *European Journal of Operational Research*, vol. 150, no. 3, pp. 496–515, 2003.

[3] J. Wallenius, J. S. Dyer, P. C. Fishburn, R. E. Steuer, S. Zionts, and K. Deb, "Multiple criteria decision making, multi-attribute utility theory: recent accomplishments and what lies ahead," *Management Science*, vol. 54, no. 7, pp. 1336–1349, 2008.

[4] W. Ho, X. Xu, and P. K. Dey, "Multi-criteria decision making approaches for supplier evaluation and selection: a literature review," *European Journal of Operational Research*, vol. 202, no. 1, pp. 16–24, 2010.

[5] C. W. Churchman, R. L. Ackoff, and N. M. Smith, "An approximate measure of value," *Journal of the Operational Research Society of America*, vol. 2, no. 2, pp. 172–187, 1954.

[6] T. L. Saaty, "The analytic hierarchy process," in *Proceedings of the Second International Seminar on Operational Research in the Basque Provinces*, vol. 4, pp. 189–234, 1980.

[7] C. L. Hwang and K. Yoon, *Multiple Attribute Decision Making: Methods and Applications*, Springer, Berlin, Germany, 1981.

[8] H. Deng, C.-H. Yeh, and R. J. Willis, "Inter-company comparison using modified TOPSIS with objective weights," *Computers and Operations Research*, vol. 27, no. 10, pp. 963–973, 2000.

[9] P. K. Dey, "Integrated project evaluation and selection using multiple-attribute decision-making technique," *International Journal of Production Economics*, vol. 103, no. 1, pp. 90–103, 2006.

[10] G.-D. Li, D. Yamaguchi, and M. Nagai, "A grey-based decision-making approach to the supplier selection problem," *Mathematical & Computer Modelling*, vol. 46, no. 3-4, pp. 573–581, 2007.

[11] A. Shanian and O. Savadogo, "TOPSIS multiple-criteria decision support analysis for material selection of metallic bipolar plates for polymer electrolyte fuel cell," *Journal of Power Sources*, vol. 159, no. 2, pp. 1095–1104, 2006.

[12] G.-H. Tzeng, Y.-P. O. Yang, C.-T. Lin, and C.-B. Chen, "Hierarchical MADM with fuzzy integral for evaluating enterprise intranet web sites," *Information Sciences*, vol. 169, no. 3-4, pp. 409–426, 2005.

[13] T. Yang and C.-C. Hung, "Multiple-attribute decision making methods for plant layout design problem," *Robotics and Computer-Integrated Manufacturing*, vol. 23, no. 1, pp. 126–137, 2007.

[14] A. Kelemenis and D. Askounis, "A new TOPSIS-based multi-criteria approach to personnel selection," *Expert Systems with Applications*, vol. 37, no. 7, pp. 4999–5008, 2010.

[15] B. Bairagi, B. Dey, B. Sarkar, and S. K. Sanyal, "A De Novo multi-approaches multi-criteria decision making technique with an application in performance evaluation of material handling device," *Computers & Industrial Engineering*, vol. 87, pp. 267–282, 2015.

[16] J. S. H. Kornbluth, "Dynamic multi-criteria decision making," *Journal of Multi-Criteria Decision Analysis*, vol. 1, no. 2, pp. 81–92, 1992.

[17] Y.-H. Lin, P.-C. Lee, T.-P. Chang, and H.-I. Ting, "Multi-attribute group decision making model under the condition of uncertain information," *Automation in Construction*, vol. 17, no. 6, pp. 792–797, 2008.

[18] R. Lahdelma and P. Salminen, "SMAA-2: stochastic multicriteria acceptability analysis for group decision making," *Operations Research*, vol. 49, no. 3, pp. 444–454, 2001.

[19] L. A. Zadeh, "Fuzzy sets," *Information and Computation*, vol. 8, pp. 338–353, 1965.

[20] I. B. Turksen, "Interval valued fuzzy sets based on normal forms," *Fuzzy Sets and Systems*, vol. 20, no. 2, pp. 191–210, 1986.

[21] S. Miyamoto, "Multisets and fuzzy multisets," in *Soft Computing and Human-Centered Machines*, Z. Q. Liu and S. Miyamoto, Eds., pp. 9–33, Springer, Berlin, Germany, 2000.

[22] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.

[23] Z. S. Xu, "Uncertain linguistic aggregation operators based approach to multiple attribute group decision making under uncertain linguistic environment," *Information Sciences*, vol. 168, no. 1–4, pp. 171–184, 2004.

[24] V. Torra, "Hesitant fuzzy sets," *International Journal of Intelligent Systems*, vol. 25, no. 6, pp. 529–539, 2010.

[25] C. Kao, "Weight determination for consistently ranking alternatives in multiple criteria decision analysis," *Applied Mathematical Modelling. Simulation and Computation for Engineering and Environmental Systems*, vol. 34, no. 7, pp. 1779–1787, 2010.

[26] D. Horsky and M. R. Rao, "Estimation of attribute weights from preference comparisons," *Management Science*, vol. 30, no. 7, pp. 801–822, 1984.

[27] P. A. Bottomley and J. R. Doyle, "A comparison of three weight elicitation methods: good, better, and best," *Omega*, vol. 29, no. 6, pp. 553–560, 2001.

[28] J. Figueira and B. Roy, "Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure," *European Journal of Operational Research*, vol. 139, no. 2, pp. 317–326, 2002.

[29] L. E. Shirland, R. R. Jesse, R. L. Thompson, and C. L. Iacovou, "Determining attribute weights using mathematical programming," *Omega*, vol. 31, no. 6, pp. 423–437, 2003.

[30] Y.-M. Wang and Y. Luo, "Integration of correlations with standard deviations for determining attribute weights in multiple attribute decision making," *Mathematical & Computer Modelling*, vol. 51, no. 1-2, pp. 1–12, 2010.

[31] J. Ma, Z.-P. Fan, and L.-H. Huang, "A subjective and objective integrated approach to determine attribute weights," *European Journal of Operational Research*, vol. 112, no. 2, pp. 397–404, 1999.

[32] X. Xu, "A note on the subjective and objective integrated approach to determine attribute weights," *European Journal of Operational Research*, vol. 156, no. 2, pp. 530–532, 2004.

[33] P. J. H. Schoemaker and C. C. Waid, "An experimental comparison of different approaches to determining weights in additive utility models," *Management Science*, vol. 28, no. 2, pp. 182–196, 1982.

[34] Y.-M. Wang and C. Parkan, "A general multiple attribute decision-making approach for integrating subjective preferences and objective information," *Fuzzy Sets and Systems*, vol. 157, no. 10, pp. 1333–1345, 2006.

[35] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.

[36] J. R. Doyle, "Multiattribute choice for the lazy decision maker: let the alternatives decide," *Organizational Behavior and Human Decision Processes*, vol. 62, no. 1, pp. 87–100, 1995.

[37] E. Bernroider and V. Stix, "A method using weight restrictions in data envelopment analysis for ranking and validity issues in decision making," *Computers & Operations Research*, vol. 34, no. 9, pp. 2637–2647, 2007.

[38] J. Wu and L. Liang, "A multiple criteria ranking method based on game cross-evaluation approach," *Annals of Operations Research*, vol. 197, no. 1, pp. 191–200, 2012.

[39] C. Fu and K.-S. Chin, "Robust evidential reasoning approach with unknown attribute weights," *Knowledge-Based Systems*, vol. 59, no. 2, pp. 9–20, 2014.

[40] B. S. Ahn, "Extreme point-based multi-attribute decision analysis with incomplete information," *European Journal of Operational Research*, vol. 240, no. 3, pp. 748–755, 2015.

[41] W. W. Cooper, J. L. Ruiz, and I. Sirvent, "Choosing weights from alternative optimal solutions of dual multiplier models in DEA," *European Journal of Operational Research*, vol. 180, no. 1, pp. 443–458, 2007.

[42] G. R. Jahanshahloo and P. F. Shahmirzadi, "New methods for ranking decision making units based on the dispersion of weights and Norm 1 in Data Envelopment Analysis," *Computers & Industrial Engineering*, vol. 65, no. 2, pp. 187–193, 2013.

[43] V. E. Krivonozhko, O. B. Utkin, M. M. Safin, and A. V. Lychev, "On some generalization of the DEA models," *Journal of the Operational Research Society*, vol. 60, no. 11, pp. 1518–1527, 2009.

[44] P. Andersen and N. C. Petersen, "A procedure for ranking efficient units in data envelopment analysis," *Management Science*, vol. 39, no. 10, pp. 1261–1264, 1993.

[45] T. R. Sexton, R. H. Silkman, and A. J. Hogan, "Data envelopment analysis: critique and extensions," *New Directions for Program Evaluation*, vol. 1986, no. 32, pp. 73–105, 1986.

[46] A. Boussofiane, R. G. Dyson, and E. Thanassoulis, "Applied data envelopment analysis," *European Journal of Operational Research*, vol. 52, no. 1, pp. 1–15, 1991.

[47] K. Wang and F. Wei, "Game cross efficiency DEA based project review of equipment design," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 35, no. 10, pp. 1278–1282, 2009.

[48] T. Wu and J. Blackhurst, "Supplier evaluation and selection: an augmented DEA approach," *International Journal of Production Research*, vol. 47, no. 16, pp. 4593–4608, 2009.

[49] S. Lim, K. W. Oh, and J. Zhu, "Use of DEA cross-efficiency evaluation in portfolio selection: an application to Korean stock market," *European Journal of Operational Research*, vol. 236, no. 1, pp. 361–368, 2014.

[50] B. Roy, "The outranking approach and the foundations of electre methods," *Theory and Decision*, vol. 31, no. 1, pp. 49–73, 1991.

[51] G. F. Reed, F. Lynn, and B. D. Meade, "Use of coefficient of variation in assessing variability of quantitative assays," *Clinical & Diagnostic Laboratory Immunology*, vol. 9, no. 6, pp. 1235–1239, 2002.

[52] G. R. Jahanshahloo, F. H. Lotfi, and M. Izadikhah, "An algorithmic method to extend TOPSIS for decision-making problems with interval data," *Applied Mathematics and Computation*, vol. 175, no. 2, pp. 1375–1384, 2006.

[53] H. M. Moshkovich, A. I. Mechitov, and D. L. Olson, "Ordinal judgments in multiattribute decision analysis," *European Journal of Operational Research*, vol. 137, no. 3, pp. 625–641, 2002.

[54] Z. S. Xu, *Uncertain Multiple Attribute Decision Making: Methods and Applications*, Tsinghua Press, Beijing, China, 2005.

[55] N. Bryson and A. Mobolurin, "An action learning evaluation procedure for multiple criteria decision making problems," *European Journal of Operational Research*, vol. 96, no. 2, pp. 379–386, 1997.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 1490712, 10 pages
http://dx.doi.org/10.1155/2016/1490712

*Research Article*

# Optimal Replenishment Policy for Weibull-Distributed Deteriorating Items with Trapezoidal Demand Rate and Partial Backlogging

**Lianxia Zhao**

*School of Management, Shanghai University, Shanghai 200444, China*

Correspondence should be addressed to Lianxia Zhao; zhaolx2004@163.com

Received 2 December 2015; Accepted 27 March 2016

Academic Editor: Alireza Amirteimoori

Copyright © 2016 Lianxia Zhao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An inventory model for Weibull-distributed deteriorating items is considered so as to minimize the total cost per unit time in this paper. The model starts with shortage, allowed partial backlogging, and trapezoidal demand rate. By analyzing the model, an efficient solution procedure is proposed to determine the optimal replenishment and the optimal order quantity and the average total costs are also obtained. Finally, numerical examples are provided to illustrate the theoretical results and a sensitivity analysis of the major parameters with respect to the stability of optimal solution is also carried out.

## 1. Introduction

Deteriorating is a general phenomenon for many products, in that fruits or vegetables are spoiled directly while alcohol undergoes physical depletion over time, and electronic products deteriorate through a gradual loss of potential utility as time passed. Based on the above situation, the effect of deteriorating for items can not be disregarded in many inventory systems. An inventory model with a constant decaying rate for items is proposed by Ghare and Schrader in 1963 [1], while Covert and Philip [2] extended Ghare and Schrader's model to an EOQ model with a two-parameter Weibull-distributed deteriorating rate. Since then, there are many researchers who focused on such topic; Wee et al. [3] considered an inventory model for deteriorating items with quantity discount, pricing, and partial backlogging, and more related articles can be referred to like Yang et al. [4], Shah et al. [5], Yang et al. [6], and so forth.

However, it is impossible for deteriorating items that the demand rate increases continuously during their full life cycle. Based on such reasons, Hill [7] proposed an inventory model with ramp-type demand rate, which extended to allow shortage proposed by Mandal and Pal [8]. Further, Wu [9] extended Mandal and Pal's model to have Weibull-distributed deterioration and time-dependent backlogging. Skouri et al. [10] considered a model with general ramp-type demand rate, partial backlogging, and Weibull deterioration rate. Tan and Weng [11] considered a discrete inventory model for deteriorating items with partial backlogging. Pal et al. [12] considered a production inventory model for deteriorating items with ramp-type demand rate under the effect of inflation and shortages under fuzziness. There are many other related literatures about such inventory model, such as Ahmed et al. [13], Agrawal et al. [14], Singh and Pattnayak [15], Wu et al. [16], and Rabbani et al. [17].

While, for some short life cycle products, the demand rate may increase up to a certain level in the early stage of marketing, then reach a stabilized period, and finally decrease to zero and reach the end of their life cycle, Cheng and Wang [18] proposed an inventory model for deteriorating items with trapezoidal type demand rate. Further, Cheng et al. [19] extended the model to be of time-dependent deteriorating items with trapezoidal type demand rate and partial backlogging. Uthayakumar and Rameswari [20] studied a model for defective items with trapezoidal type demand rate to determine the optimal product reliability. Lin [21] and

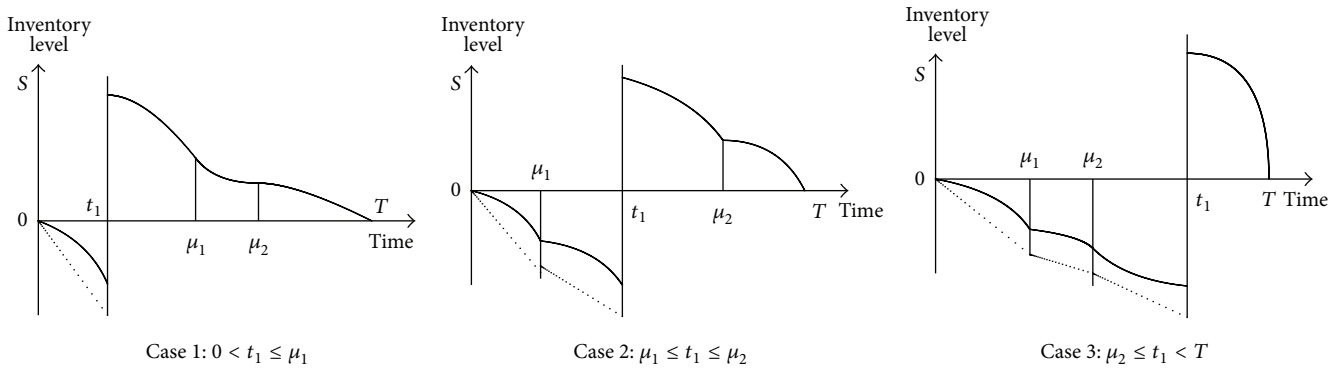



Figure 1: Graphical representation of inventory level over the cycle starting with shortage.

Lin et al. [22] considered the inventory model which was proposed by Cheng and Wang [18]. Zhao [23] studied an inventory model for Weibull-distributed deterioration items with trapezoidal type demand rate and partial backlogging in which the model begins with no-shortage. While, in real situation, customers will respond differently in case of shortages, either to wait until replenishment or to look for alternative products, in this paper, we study an inventory model beginning with shortage, Weibull-distributed deterioration items, trapezoidal type demand rate, and time-dependent partial backlogging. By analyzing the studied model, we propose an optimal replenishment strategy for it and also carry out a sensitivity analysis of the main parameters.

The remainder of the paper is organized as follows. Notations and assumptions are described in Section 2. Section 3 constructs and analyzes the inventory model starting with shortage. Some numerical examples to illustrate the solution procedure are provided; then sensitivity analysis of the major parameters is also carried out in Section 4. Finally, a conclusion is presented.

## 2. Notations and Assumptions

The fundamental notations and assumptions used in this paper are given as follows.

(i) $I(t)$ is the inventory level at time $t$, $0 \le t \le T$, where $T$ is the length of inventory cycle.

(ii) $t_1$ denotes the replenishment time, and $t_1^*$ denotes the optimal $t_1$.

(iii) $S$ is the maximum inventory level for each ordering cycle.

(iv) $Q^*$ denotes the optimal ordering quantity.

(v) $A_0$ denotes the fixed cost per order.

(vi) $c_1$ is the cost of each deteriorated item, $c_2$ is the unit inventory holding cost per unit time, $c_3$ is the unit shortage cost per unit time, and $c_4$ is the unit lost sales cost.

(vii) $C_i(t_1)$, $i = 1, 2, 3$, denotes the average total cost per unit time under different conditions, respectively.

(viii) $TC(t_1)$ denotes the average total cost per unit time.

(ix) The demand rate, $D(t)$, which is positive and consecutive, is assumed to be a trapezoidal type function of time; that is,

$$D(t) = \begin{cases} f(t), & t \le \mu_1, \\ D_0, & \mu_1 < t < \mu_2, \\ g(t), & \mu_2 \le t < T, \end{cases} \tag{1}$$

where $\mu_1$ is time point changing from the increasing demand function $f(t)$ to the constant demand $D_0$, and $\mu_2$ is time point changing from the constant demand $D_0$ to the decreasing demand function $g(t)$.

(x) The replenishment rate is infinite; that is, replenishment is instantaneous.

(xi) The time to deterioration of the item is distributed as Weibull $(\alpha, \beta)$; that is, the deterioration rate is $\theta(t) = \alpha\beta t^{\beta-1}$, where $\alpha$ is the scale parameter, $0 < \alpha < 1$, and $\beta$ is the shape parameter, $0 < \beta$.

(xii) Shortage is allowed and let $e^{-\delta t}$ be fraction of shortages backlogged, where $t$ is the waiting time up to the next replenishment and $\delta$ is a tiny positive constant; that is, $\delta \ll 1/T$.

## 3. Inventory Model Starting with Shortage

The system of model begins with shortage and until $t_1$ the shortage achieves its maximum; then replenishment occurs at time $t = t_1$. All of the shortage demand during $(0, t_1)$ is partial backlogged (Figure 1). At $t_1$, the total number of backlogged items will be satisfied by the replenishment and the inventory level up to $S$, while the deteriorating occurs during the time of $[t_1, T]$. According to the notations and

assumptions mentioned above, the model is described as shown in Figure 1:

$$\frac{dI(t)}{dt} = \begin{cases} -e^{-\delta(t_1-t)}D(t), & 0 < t < t_1, \\ -\theta(t)I(t) - D(t), & t_1 < t < T, \end{cases} \tag{2}$$

with boundary condition $I(0) = 0$, $I(t_1) = S$, $I(T) = 0$.

Based on the values of $t_1$, $\mu_1$, and $\mu_2$, three possible cases are presented as follows.

*Case 1* ($0 < t_1 \le \mu_1$). If the replenishment time $t_1 \in (0, \mu_1]$, from (2), we have

$$\frac{dI(t)}{dt} = \begin{cases} -e^{-\delta(t_1-t)}f(t), & 0 < t < t_1, \\ -\alpha\beta t^{\beta-1}I(t) - f(t), & t_1 < t < \mu_1, \\ -\alpha\beta t^{\beta-1}I(t) - D_0, & \mu_1 < t < \mu_2, \\ -\alpha\beta t^{\beta-1}I(t) - g(t), & \mu_2 < t < T. \end{cases} \tag{3}$$

Solving the differential equations (3) with $I(0) = 0$, we have

$$I(t) = \begin{cases} -\int_0^t e^{\delta(x-t_1)}f(x)\,dx, & 0 < t < t_1, \\ \int_t^{\mu_1} e^{\alpha(x^\beta - t^\beta)}f(x)\,dx + D_0\int_{\mu_1}^{\mu_2} e^{\alpha(x^\beta - t^\beta)}dx + \int_{\mu_2}^{T} e^{\alpha(x^\beta - t^\beta)}g(x)\,dx, & t_1 < t < \mu_1, \\ D_0\int_t^{\mu_2} e^{\alpha(x^\beta - t^\beta)}dx + \int_{\mu_2}^{T} e^{\alpha(x^\beta - t^\beta)}g(x)\,dx, & \mu_1 < t < \mu_2, \\ \int_t^{T} e^{\alpha(x^\beta - t^\beta)}g(x)\,dx, & \mu_2 < t < T. \end{cases} \tag{4}$$

The total replenishment quantity can be computed as

$$\begin{aligned} S &= I(t_1) \\ &= \int_{t_1}^{\mu_1} e^{\alpha(x^\beta - t_1^\beta)}f(x)\,dx + D_0\int_{\mu_1}^{\mu_2} e^{\alpha(x^\beta - t_1^\beta)}dx \\ &\quad + \int_{\mu_2}^{T} e^{\alpha(x^\beta - t_1^\beta)}g(x)\,dx. \end{aligned} \tag{5}$$

The total shortage quantity $B_T$ during the interval $[0, t_1)$ is

$$B_T = \int_0^{t_1} (t_1 - t)\, e^{\delta(t-t_1)}f(t)\,dt. \tag{6}$$

The total number of perished items during the interval $[t_1, T]$, say $D_T$, is

$$\begin{aligned} D_T &= \int_{t_1}^{\mu_1} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) f(t)\,dt \\ &\quad + D_0\int_{\mu_1}^{\mu_2} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) dt \\ &\quad + \int_{\mu_2}^{T} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) g(t)\,dt. \end{aligned} \tag{7}$$

The total number of inventories carried during the interval $[t_1, T]$, say $H_T$, is

$$\begin{aligned} H_T &= \int_{t_1}^{\mu_1}\int_t^{\mu_1} e^{\alpha(x^\beta - t^\beta)}f(x)\,dx\,dt \\ &\quad + D_0\int_{t_1}^{\mu_1}\int_{\mu_1}^{\mu_2} e^{\alpha(x^\beta - t^\beta)}dx\,dt \\ &\quad + \int_{\mu_1}^{\mu_2}\int_{\mu_2}^{T} e^{\alpha(x^\beta - t^\beta)}g(x)\,dx\,dt \\ &\quad + D_0\int_{\mu_1}^{\mu_2}\int_t^{\mu_2} e^{\alpha(x^\beta - t^\beta)}dx\,dt \\ &\quad + \int_{\mu_2}^{T}\int_t^{T} e^{\alpha(x^\beta - t^\beta)}g(x)\,dx\,dt. \end{aligned} \tag{8}$$

The total of lost sales during the interval $[0, t_1]$, say $L_T$, is

$$L_T = \int_0^{t_1} \left(1 - e^{\delta(t-t_1)}\right) f(t)\,dt. \tag{9}$$

Therefore, the average total cost per unit time under the condition $t_1 \le \mu_1$ can be given by

$$\begin{aligned} \mathrm{C}_1(t_1) &= \frac{1}{T}\left[A_0 + \mathrm{c}_1 D_T + \mathrm{c}_2 H_T + \mathrm{c}_3 B_T + \mathrm{c}_4 L_T\right] \\ &= \frac{1}{T}\left\{A_0 + \mathrm{c}_3\left[\int_0^{t_1} (t_1 - t)\, e^{\delta(t-t_1)}f(t)\,dt\right]\right. \\ &\quad + \mathrm{c}_4\left[\int_0^{t_1} \left(1 - e^{\delta(t-t_1)}\right) f(t)\,dt\right] \\ &\quad + \mathrm{c}_1\left[\int_{t_1}^{\mu_1} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) f(t)\,dt\right. \\ &\quad \left. + D_0\int_{\mu_1}^{\mu_2} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) dt \right. \end{aligned}$$

$$
\begin{aligned}
&+\int_{\mu_2}^{T}\left(e^{\alpha(t^{\beta}-t_1^{\beta})}-1\right)g(t)\,dt\Bigg]\\
&+c_2\Bigg[\int_{t_1}^{\mu_1}\int_{t}^{\mu_1}e^{\alpha(x^{\beta}-t^{\beta})}f(x)\,dx\,dt\\
&+D_0\int_{t_1}^{\mu_1}\int_{\mu_1}^{\mu_2}e^{\alpha(x^{\beta}-t^{\beta})}dx\,dt\\
&+\int_{t_1}^{\mu_2}\int_{\mu_2}^{T}e^{\alpha(x^{\beta}-t^{\beta})}g(x)\,dx\,dt\\
&+D_0\int_{\mu_1}^{\mu_2}\int_{t}^{\mu_2}e^{\alpha(x^{\beta}-t^{\beta})}dx\,dt\\
&+\int_{\mu_2}^{T}\int_{t}^{T}e^{\alpha(x^{\beta}-t^{\beta})}g(x)\,dx\,dt\Bigg]\Bigg\}.
\end{aligned}
\tag{10}
$$

Taking the first and the second order derivative of $C_1(t_1)$ with respect to $t_1$, respectively, we have

$$
\begin{aligned}
\frac{dC_1(t_1)}{dt_1}&=\frac{1}{T}\Bigg\{c_3\int_0^{t_1}(1-\delta t_1+\delta t)e^{\delta(t-t_1)}f(t)\,dt\\
&+c_4\delta\int_0^{t_1}e^{\delta(t-t_1)}f(t)\,dt-\left(c_2+c_1\alpha\beta t_1^{\beta-1}\right)\\
&\cdot\int_{t_1}^{T}e^{\alpha(t^{\beta}-t_1^{\beta})}D(t)\,dt\Bigg\},
\end{aligned}
\tag{11}
$$

$$
\begin{aligned}
\frac{d^2C_1(t_1)}{dt_1^2}&=\frac{1}{T}\Bigg\{c_3\Bigg[f(t_1)\\
&+\delta\int_0^{t_1}(\delta t_1-\delta t-2)e^{\delta(t-t_1)}f(t)\,dt\Bigg]+c_4\delta\Bigg[f(t_1)\\
&-\delta\int_0^{t_1}e^{\delta(t-t_1)}f(t)\,dt\Bigg]+\left(c_1\alpha\beta t_1^{\beta-1}+c_2\right)f(t_1)\\
&+\alpha\beta t_1^{\beta-2}\left[c_1\left(\alpha\beta t_1^{\beta}-\beta+1\right)+c_2t_1\right]\\
&\cdot\int_{t_1}^{T}e^{\alpha(t^{\beta}-t_1^{\beta})}D(t)\,dt\Bigg\}.
\end{aligned}
\tag{12}
$$

Then, we have

$$
\left.\frac{dC_1(t_1)}{dt_1}\right|_{t_1=0}=-\frac{c_2}{T}\int_0^{T}e^{\alpha t^{\beta}}D(t)\,dt<0,
\tag{13}
$$

and let

$$
\begin{aligned}
\Delta_1&=c_3\int_0^{\mu_1}(1-\delta\mu_1+\delta t)e^{\delta(t-\mu_1)}f(t)\,dt\\
&+c_4\delta\int_0^{\mu_1}e^{\delta(t-\mu_1)}f(t)\,dt-\left(c_2+c_1\alpha\beta\mu_1^{\beta-1}\right)\\
&\cdot\left(D_0\int_{\mu_1}^{\mu_2}e^{\alpha(t^{\beta}-\mu_1^{\beta})}dt+\int_{\mu_2}^{T}e^{\alpha(t^{\beta}-\mu_1^{\beta})}g(t)\,dt\right).
\end{aligned}
\tag{14}
$$

Therefore, (i) if $\Delta_1>0$ and $d^2C_1(t_1)/dt_1^2>0$, there exists a unique solution $t_1^*\in(0,\mu_1)$ satisfying $dC_1(t_1)/dt_1=0$, and $C_1(t_1)$ obtains its minimum at $t_1=t_1^*$; (ii) if $\Delta_1<0$ and $d^2C_1(t_1)/dt_1^2>0$, it means that $C_1(t_1)$ is a nonincreasing function and obtains its minimum value at point $t_1^*=\mu_1$; (iii) if the sign of $d^2C_1(t_1)/dt_1^2$ is indefinite, then the curve of $C_1(t_1)$ is W-shape, and $C_1(t_1)$ may obtain its minimum at more than one point; without loss of generality, set $t_1=t_1^*$, where $t_1^*$ is one of the solutions of $dC_1(t_1)/dt_1=0$.

Thus, the optimal value of the order level, $S=I(t_1^*)$, is

$$
\begin{aligned}
S^*&=\int_{t_1^*}^{\mu_1}e^{\alpha(x^{\beta}-t_1^{*\beta})}f(x)\,dx+D_0\int_{\mu_1}^{\mu_2}e^{\alpha(x^{\beta}-t_1^{*\beta})}dx\\
&+\int_{\mu_2}^{T}e^{\alpha(x^{\beta}-t_1^{*\beta})}g(x)\,dx,
\end{aligned}
\tag{15}
$$

and the optimal order quantity $Q^*$ is

$$
Q^*=S^*+\int_0^{t_1^*}e^{\delta(t-t_1^*)}f(t)\,dt.
\tag{16}
$$

*Case 2* ($\mu_1\le t_1\le\mu_2$). If the replenishment time $t_1\in[\mu_1,\mu_2]$, then from (2), we have

$$
\frac{dI(t)}{dt}=\begin{cases}
-e^{-\delta(t_1-t)}f(t), & 0<t<\mu_1,\\
-e^{-\delta(t_1-t)}D_0, & \mu_1<t<t_1,\\
-\alpha\beta t^{\beta-1}I(t)-D_0, & t_1<t<\mu_2,\\
-\alpha\beta t^{\beta-1}I(t)-g(t), & \mu_2<t<T.
\end{cases}
\tag{17}
$$

Solving the differential equation (17) with $I(0)=I(T)=0$, we have

$$
I(t)=\begin{cases}
-\int_0^{t}e^{\delta(x-t_1)}f(x)\,dx, & 0<t<\mu_1,\\
-\frac{D_0}{\delta}\left(e^{\delta(t-t_1)}-e^{\delta(\mu_1-t_1)}\right)-\int_0^{\mu_1}e^{\delta(x-t_1)}f(x)\,dx, & \mu_1<t<t_1,\\
D_0\int_{t}^{\mu_2}e^{\alpha(x^{\beta}-t^{\beta})}dx+\int_{\mu_2}^{T}e^{\alpha(x^{\beta}-t^{\beta})}g(x)\,dx, & t_1<t<\mu_2,\\
\int_{t}^{T}e^{\alpha(x^{\beta}-t^{\beta})}g(x)\,dx, & \mu_2<t<T.
\end{cases}
\tag{18}
$$

From (18), the total replenishment quantity can be computed:

$$S = I(t_1) = D_0 \int_{t_1}^{\mu_2} e^{\alpha(x^\beta - t_1^\beta)} dx + \int_{\mu_2}^{T} e^{\alpha(x^\beta - t_1^\beta)} g(x)\, dx. \quad (19)$$

The total shortage quantity during the interval $(0, t_1]$ is

$$\begin{aligned} B_T = &\int_0^{\mu_1} (t_1 - t) e^{\delta(t-t_1)} f(t)\, dt \\ &+ \frac{D_0}{\delta} \int_{\mu_1}^{t_1} \left(e^{\delta(t-t_1)} - e^{\delta(\mu_1 - t_1)}\right) dt. \end{aligned} \quad (20)$$

The total number of perished items in the interval $[t_1, T]$ is

$$\begin{aligned} D_T = &D_0 \int_{t_1}^{\mu_2} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) dt \\ &+ \int_{\mu_2}^{T} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) g(t)\, dt. \end{aligned} \quad (21)$$

The total number of inventories carried during the interval $[t_1, T]$ is

$$\begin{aligned} H_T = &D_0 \int_{t_1}^{\mu_2} \int_{t}^{\mu_2} e^{\alpha(x^\beta - t^\beta)} dx\, dt \\ &+ \int_{t_1}^{\mu_2} \int_{\mu_2}^{T} e^{\alpha(x^\beta - t^\beta)} g(x)\, dx\, dt \\ &+ \int_{\mu_2}^{T} \int_{t}^{T} e^{\alpha(x^\beta - t^\beta)} g(x)\, dx\, dt. \end{aligned} \quad (22)$$

The total of lost sales during the interval $[0, t_1]$ is

$$\begin{aligned} L_T = &\int_0^{\mu_1} \left(1 - e^{\delta(t-t_1)}\right) f(t)\, dt \\ &+ D_0 \int_{\mu_1}^{t_1} \left(1 - e^{\delta(t-t_1)}\right) dt. \end{aligned} \quad (23)$$

Therefore, the average total cost per unit time under the condition $\mu_1 < t_1 \le \mu_2$ can be given by

$$\begin{aligned} \mathrm{C}_2(t_1) &= \frac{1}{T}\left[A_0 + \mathrm{c}_1 D_T + \mathrm{c}_2 H_T + \mathrm{c}_3 B_T + \mathrm{c}_4 L_T\right] \\ &= \frac{1}{T}\left\{A_0 + \mathrm{c}_1\left[D_0 \int_{t_1}^{\mu_2} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) dt \right.\right. \\ &\quad \left. + \int_{\mu_2}^{T} \left(e^{\alpha(t^\beta - t_1^\beta)} - 1\right) g(t)\, dt\right] \\ &\quad + \mathrm{c}_2\left[D_0 \int_{t_1}^{\mu_2} \int_{t}^{\mu_2} e^{\alpha(x^\beta - t^\beta)} dx\, dt \right. \\ &\quad + \int_{t_1}^{\mu_2} \int_{\mu_2}^{T} e^{\alpha(x^\beta - t^\beta)} g(x)\, dx\, dt \\ &\quad \left. + \int_{\mu_2}^{T} \int_{t}^{T} e^{\alpha(x^\beta - t^\beta)} g(x)\, dx\, dt\right] \\ &\quad + \mathrm{c}_4\left[\int_0^{\mu_1} \left(1 - e^{\delta(t-t_1)}\right) f(t)\, dt \right. \\ &\quad \left. + D_0 \int_{\mu_1}^{t_1} \left(1 - e^{\delta(t-t_1)}\right) dt\right] \\ &\quad + \mathrm{c}_3\left[\int_0^{\mu_1} (t_1 - t) e^{\delta(t-t_1)} f(t)\, dt \right. \\ &\quad \left.\left. + \frac{D_0}{\delta} \int_{\mu_1}^{t_1} \left(e^{\delta(t-t_1)} - e^{\delta(\mu_1 - t_1)}\right) dt\right]\right\}. \end{aligned} \quad (24)$$

Taking the first and the second order derivative of $\mathrm{C}_2(t_1)$ with respect to $t_1$, respectively, we have

$$\begin{aligned} \frac{d\mathrm{C}_2(t_1)}{dt_1} = &\frac{1}{T}\left\{\mathrm{c}_4\delta\left[\int_0^{\mu_1} e^{\delta(t-t_1)} f(t)\, dt \right.\right. \\ &\left. + D_0 \int_{\mu_1}^{t_1} e^{\delta(t-t_1)} dt\right] - \left(\mathrm{c}_1\alpha\beta t_1^{\beta-1} + \mathrm{c}_2\right) \\ &\cdot \left[D_0 \int_{t_1}^{\mu_2} e^{\alpha(t^\beta - t_1^\beta)} dt + \int_{\mu_2}^{T} e^{\alpha(t^\beta - t_1^\beta)} g(t)\, dt\right] \\ &+ \mathrm{c}_3\left[\int_0^{\mu_1} \left(1 - \delta t_1 + \delta t\right) e^{\delta(t-t_1)} f(t)\, dt \right. \\ &\left.\left. - D_0 (\mu_1 - t_1) e^{\delta(\mu_1 - t_1)}\right]\right\}, \end{aligned} \quad (25)$$

$$\begin{aligned} \frac{d^2\mathrm{C}_2(t_1)}{dt_1^2} = &\frac{1}{T}\left\{\alpha\beta t_1^{\beta-2}\left(\mathrm{c}_1\left(\alpha\beta t_1^\beta - \beta + 1\right) + \mathrm{c}_2 t_1\right)\right. \\ &\cdot \left[D_0 \int_{t_1}^{\mu_2} e^{\alpha(t^\beta - t_1^\beta)} dt + \int_{\mu_2}^{T} e^{\alpha(t^\beta - t_1^\beta)} g(t)\, dt\right] \\ &+ \left(\mathrm{c}_1\alpha\beta t_1^{\beta-1} + \mathrm{c}_2\right) D_0 + \mathrm{c}_4\delta\left[D_0 e^{\delta(\mu_1 - t_1)} \right. \\ &\left. - \delta \int_0^{\mu_1} e^{\delta(t-t_1)} f(t)\, dt\right] \\ &+ \mathrm{c}_3\left[\int_0^{\mu_1} \left(\delta^2 t_1 - \delta^2 t - 2\delta\right) e^{\delta(t-t_1)} f(t)\, dt \right. \\ &\left.\left. + D_0 \left(1 + \delta\mu_1 - \delta t_1\right) e^{\delta(\mu_1 - t_1)}\right]\right\}. \end{aligned} \quad (26)$$

Let

$$\begin{aligned} \Delta_2 = &\mathrm{c}_3\left[\int_0^{\mu_1} \left(1 - \delta\mu_2 + \delta t\right) e^{\delta(t-\mu_2)} f(t)\, dt \right. \\ &\left. + D_0 (\mu_2 - \mu_1) e^{\delta(\mu_1 - \mu_2)}\right] \\ &+ \mathrm{c}_4\delta\left[\int_0^{\mu_1} e^{\delta(t-\mu_2)} f(t)\, dt + D_0 \int_{\mu_1}^{\mu_2} e^{\delta(t-\mu_2)} dt\right] \\ &- \left(\mathrm{c}_2 + \mathrm{c}_1\alpha\beta\mu_2^{\beta-1}\right) \int_{\mu_2}^{T} e^{\alpha(t^\beta - \mu_2^\beta)} g(t)\, dt. \end{aligned} \quad (27)$$

Therefore, (i) if $\Delta_1 < 0$, $\Delta_2 > 0$, and $d^2C_2(t_1)/dt_1^2 > 0$, there exists a unique solution $t_1^* \in (\mu_1, \mu_2)$ satisfying $dC_2(t_1)/dt_1 = 0$, and $C_2(t_1)$ obtains its minimum at $t_1 = t_1^*$; (ii) if $\Delta_1 < 0$, $\Delta_2 \le 0$, and $d^2C_2(t_1)/dt_1^2 > 0$, it means that $C_2(t_1)$ is a nonincreasing function and obtains its minimum at point $t_1^* = \mu_2$; (iii) if $\Delta_1 \ge 0$, $\Delta_2 > 0$, and $d^2C_2(t_1)/dt_1^2 > 0$, it means that $C_2(t_1)$ is a nondecreasing function and obtains its minimum at point $t_1^* = \mu_1$; (iv) if the sign of $d^2C_2(t_1)/dt_1^2$ is indefinite, then the curve of $C_2(t_1)$ is W-shape, and $C_2(t_1)$ may obtain its minimum value at more than one point; without loss of generality, set $t_1 = t_1^*$, where $t_1^*$ is one of the solutions of $dC_2(t_1)/dt_1 = 0$.

Thus, the optimal value of the order level, $S = I(t_1^*)$, is

$$S^* = D_0 \int_{t_1^*}^{\mu_2} e^{\alpha(t^\beta - t_1^{*\beta})} dt + \int_{\mu_2}^{T} e^{\alpha(t^\beta - t_1^{*\beta})} g(t)\, dt, \tag{28}$$

and the optimal order quantity $Q^*$ is

$$Q^* = S^* + \int_0^{\mu_1} e^{\delta(t-t_1^*)} f(t)\, dt + D_0 \int_{\mu_1}^{t_1^*} e^{\delta(t-t_1^*)} dt. \tag{29}$$

*Case 3* ($\mu_2 \le t_1 < T$). If the replenishment time $t_1 \in [\mu_2, T)$, from (2), we have

$$\frac{dI(t)}{dt} = \begin{cases} -e^{-\delta(t_1-t)} f(t), & 0 < t < \mu_1, \\ -e^{-\delta(t_1-t)} D_0, & \mu_1 < t < \mu_2, \\ -e^{-\delta(t_1-t)} g(t), & \mu_2 < t < t_1, \\ -\alpha\beta t^{\beta-1} I(t) - g(t), & t_1 < t < T. \end{cases} \tag{30}$$

Solving (30) with $I(0) = I(T) = 0$, we have

$$I(t) = \begin{cases} -\int_0^t e^{\delta(x-t_1)} f(x)\, dx, & 0 < t < \mu_1, \\ \frac{D_0}{\delta}\left(e^{\delta(\mu_1-t_1)} - e^{\delta(t-t_1)}\right) - \int_0^{\mu_1} e^{\delta(x-t_1)} f(x)\, dx, & \mu_1 < t < \mu_2, \\ -\int_{\mu_2}^t e^{\delta(x-t_1)} g(x)\, dx + \frac{D_0}{\delta}\left(e^{\delta(\mu_1-t_1)} - e^{\delta(\mu_2-t_1)}\right) - \int_0^{\mu_1} e^{\delta(x-t_1)} f(x)\, dx, & \mu_2 < t < t_1, \\ \int_t^T e^{\alpha(x^\beta - t^\beta)} g(x)\, dx, & t_1 < t < T. \end{cases} \tag{31}$$

From the last equation of (31), the total replenishment quantity can be computed:

$$S = I(t_1) = \int_{t_1}^{T} e^{\alpha(x^\beta - t_1^\beta)} g(x)\, dx. \tag{32}$$

The total shortage quantity during the interval $(0, t_1]$ is

$$\begin{aligned} B_T = {} & \int_0^{\mu_1} (t_1 - t) e^{\delta(t-t_1)} f(t)\, dt \\ & + \int_{\mu_2}^{t_1} (t_1 - t) e^{\delta(t-t_1)} g(t)\, dt \\ & + \frac{D_0}{\delta^2}\left(e^{\delta(\mu_1-t_1)} - e^{\delta(\mu_2-t_1)}\right) \\ & + \frac{D_0}{\delta}\left[(t_1 - \mu_1) e^{\delta(\mu_1-t_1)} - (t_1 - \mu_2) e^{\delta(\mu_2-t_1)}\right]. \end{aligned} \tag{33}$$

The total number of perished items in the interval $[t_1, T]$ is

$$D_T = \int_{t_1}^{T} \left(e^{\alpha(x^\beta - t_1^\beta)} - 1\right) g(x)\, dx. \tag{34}$$

The total number of inventories carried during the interval $[t_1, T]$ is

$$H_T = \int_{t_1}^{T} \int_t^T e^{\alpha(x^\beta - t^\beta)} g(x)\, dx\, dt. \tag{35}$$

The total of lost sales during the interval $[0, t_1]$ is

$$\begin{aligned} L_T = {} & \int_0^{\mu_1} \left(1 - e^{\delta(t-t_1)}\right) f(t)\, dt \\ & + D_0 \int_{\mu_1}^{\mu_2} \left(1 - e^{\delta(t-t_1)}\right) dt \\ & + \int_{\mu_2}^{t_1} \left(1 - e^{\delta(t-t_1)}\right) g(t)\, dt. \end{aligned} \tag{36}$$

Therefore, the average total cost per unit time under the condition $\mu_2 < t_1 \le T$ can be given by

$$\begin{aligned} C_3(t_1) &= \frac{1}{T}\left[A_0 + c_1 D_T + c_2 H_T + c_3 B_T + c_4 L_T\right] \\ &= \frac{1}{T}\left\{A_0 + c_1 \int_{t_1}^{T} \left(e^{\alpha(x^\beta - t_1^\beta)} - 1\right) g(x)\, dx \right. \\ &\quad + c_2 \int_{t_1}^{T} \int_t^T e^{\alpha(x^\beta - t^\beta)} g(x)\, dx\, dt \\ &\quad + c_4 \left[\int_0^{\mu_1} \left(1 - e^{\delta(t-t_1)}\right) f(t)\, dt \right. \\ &\quad \left. \left. + D_0 \int_{\mu_1}^{\mu_2} \left(1 - e^{\delta(t-t_1)}\right) dt \right. \right. \end{aligned}$$

$$+\int_{\mu_2}^{t_1}\left(1-e^{\delta(t-t_1)}\right)g(t)\,dt\Bigg]$$
$$+c_3\left[\int_0^{\mu_1}(t_1-t)\,e^{\delta(t-t_1)}f(t)\,dt\right.$$
$$+\int_{\mu_2}^{t_1}(t_1-t)\,e^{\delta(t-t_1)}g(t)\,dt$$
$$+\frac{D_0}{\delta^2}\left(e^{\delta(\mu_1-t_1)}-e^{\delta(\mu_2-t_1)}\right)$$
$$\left.+\frac{D_0}{\delta}\left((t_1-\mu_1)\,e^{\delta(\mu_1-t_1)}-(t_1-\mu_2)\,e^{\delta(\mu_2-t_1)}\right)\right]\Bigg\}. \tag{37}$$

Taking the first and the second order derivative of $C_3(t_1)$ with respect to $t_1$, respectively, we have

$$\frac{dC_3(t_1)}{dt_1}=\frac{1}{T}\left\{-\left(c_2+c_1\alpha\beta t_1^{\beta-1}\right)\int_{t_1}^{T}e^{\alpha(t^\beta-t_1^\beta)}g(t)\,dt\right.$$
$$+c_4\delta\left[\int_0^{\mu_1}e^{\delta(t-t_1)}f(t)\,dt+D_0\int_{\mu_1}^{\mu_2}e^{\delta(t-t_1)}dt\right.$$
$$\left.+\int_{\mu_2}^{t_1}e^{\delta(t-t_1)}g(t)\,dt\right]$$
$$+c_3\left[\int_0^{\mu_1}(1+\delta t-\delta t_1)\,e^{\delta(t-t_1)}f(t)\,dt\right. \tag{38}$$
$$+\int_{\mu_2}^{t_1}(1+\delta t-\delta t_1)\,e^{\delta(t-t_1)}g(t)\,dt$$
$$\left.+D_0\left((t_1-\mu_1)\,e^{\delta(\mu_1-t_1)}-(t_1-\mu_2)\,e^{\delta(\mu_2-t_1)}\right)\right]\Bigg\},$$

$$\frac{d^2C_3(t_1)}{dt_1^2}=\frac{1}{T}\left\{\alpha\beta t_1^{\beta-2}\left[c_1\left(\alpha\beta t_1^\beta-\beta+1\right)+c_2t_1\right]\right.$$
$$\cdot\int_{t_1}^{T}e^{\alpha(t^\beta-t_1^\beta)}g(t)\,dt+\left(c_2+c_1\alpha\beta t_1^{\beta-1}\right)g(t_1)$$
$$+c_4\delta\left[\delta\int_0^{\mu_1}e^{\delta(t-t_1)}f(t)\,dt+D_0\delta\int_{\mu_1}^{\mu_2}e^{\delta(t-t_1)}dt\right.$$
$$\left.+\delta\int_{\mu_2}^{t_1}e^{\delta(t-t_1)}g(t)\,dt+g(t_1)\right]$$
$$+c_2\left[\alpha\beta t_1^{\beta-1}\int_{t_1}^{T}e^{\alpha(t^\beta-t_1^\beta)}g(t)\,dt+g(t_1)\right]$$
$$+c_3\left[\int_0^{\mu_1}\left(\delta^2t_1-\delta^2t-2\delta\right)e^{\delta(t-t_1)}f(t)\,dt\right.$$
$$+\int_{\mu_2}^{t_1}\left(\delta^2t_1-\delta^2t-2\delta\right)e^{\delta(t-t_1)}g(t)\,dt+g(t_1)$$
$$+D_0\left((1+\delta\mu_1-\delta t_1)\,e^{\delta(\mu_1-t_1)}\right.$$
$$\left.\left.-\left(1+\delta\mu_2-\delta t_1\right)e^{\delta(\mu_2-t_1)}\right)\right]\Bigg\}. \tag{39}$$

We know that

$$\left.\frac{dC_3(t_1)}{dt_1}\right|_{t_1=T}=\frac{1}{T}\left\{c_4\delta\left[\int_0^{\mu_1}e^{\delta(t-T)}f(t)\,dt\right.\right.$$
$$\left.+D_0\int_{\mu_1}^{\mu_2}e^{\delta(t-T)}dt+\int_{\mu_2}^{T}e^{\delta(t-T)}g(t)\,dt\right]$$
$$+c_3\left[\int_0^{\mu_1}(1+\delta t-\delta T)\,e^{\delta(t-T)}f(t)\,dt\right. \tag{40}$$
$$+\int_{\mu_2}^{T}(1+\delta t-\delta T)\,e^{\delta(t-T)}g(t)\,dt$$
$$\left.\left.+D_0\left((T-\mu_1)\,e^{\delta(\mu_1-T)}-(T-\mu_2)\,e^{\delta(\mu_2-T)}\right)\right]\right\}$$
$$>0.$$

Therefore, (i) if $\Delta_2<0$ and $d^2C_3(t_1)/dt_1^2>0$, there exists a unique solution $t_1^*\in(\mu_2,T)$ satisfying $dC_3(t_1)/dt_1=0$, and $C_3(t_1)$ obtains its minimum at $t_1=t_1^*$; (ii) if $\Delta_2\geq 0$ and $d^2C_3(t_1)/dt_1^2>0$, it means that $C_3(t_1)$ is a nondecreasing function and obtains its minimum value at point $t_1^*=\mu_2$; (iii) if the sign of $d^2C_3(t_1)/dt_1^2$ is indefinite, then the curve of $C_3(t_1)$ is W-shape, and $C_3(t_1)$ may obtain its minimum value at more than one point; without loss of generality, set $t_1=t_1^*$, where $t_1^*$ is one of the solutions of $dC_3(t_1)/dt_1=0$.

Therefore, the optimal value of the order level, $S=I(t_1^*)$, is

$$S^*=\int_{t_1^*}^{T}e^{\alpha(x^\beta-t_1^{*\beta})}g(x)\,dx, \tag{41}$$

and the optimal order quantity $Q^*$ is

$$Q^*=S^*+\int_0^{\mu_1}e^{\delta(t-t_1^*)}f(t)\,dt+D_0\int_{\mu_1}^{\mu_2}e^{\delta(t-t_1^*)}dt$$
$$+\int_{\mu_2}^{t_1^*}e^{\delta(t-t_1^*)}g(t)\,dt. \tag{42}$$

From the above analysis, we obtain that the total average cost of the system over the interval $[0,T]$ is

$$\mathrm{TC}(t_1)=\begin{cases}C_1(t_1), & 0<t_1\leq\mu_1,\\ C_2(t_1), & \mu_1<t_1\leq\mu_2,\\ C_3(t_1), & \mu_2<t_1<T,\end{cases} \tag{43}$$

where $C_1(t_1)$, $C_2(t_1)$, and $C_3(t_1)$ are obtained from (10), (24), and (37), respectively.

Basing on the analyzed results above, we provide the results which ensure the existence of a unique $t_1$, say $t_1^*$, to minimize the total average cost for the model system starting with shortages. The procedure is as follows:

*Step 0.* Input the parameters, $\mu_1$, $\mu_2$, $T$, $a_1$, $a_2$, $b_1$, $\alpha$, $\beta$, $\delta$, $D_0$, and $c_i$, $i = 1, 2, 3, 4$.

*Step 1.* Calculate $\Delta_1$ and $d^2C_1(t_1)/dt_1^2$.

*Step 1.1.* If $\Delta_1 > 0$ and $d^2C_1(t_1)/dt_1^2 > 0$, solve a unique $t_1^*$ from $dC_1(t_1)/dt_1 = 0$, and the total average cost and the order quantity can be obtained by (10) and (16), respectively.

*Step 1.2.* If $\Delta_1 \leq 0$ and $d^2C_1(t_1)/dt_1^2 > 0$, then the optimal average cost and the optimal order quantity can be obtained by (10) and (16) at $t_1^* = \mu_1$, respectively.

*Step 1.3.* If the sign of $d^2C_1(t_1)/dt_1^2$ is indefinite, then solve equation $dC_1(t_1)/dt_1 = 0$. If the solution is not unique, by comparing their average costs $C_1(t_1)$ to obtain the optimal replenishment point, compute the order quantity by (16).

*Step 2.* Calculate $\Delta_2$ and $d^2C_2(t_1)/dt_1^2$.

*Step 2.1.* If $\Delta_1 < 0$, $\Delta_2 > 0$, and $d^2C_2(t_1)/dt_1^2 > 0$, then seek a unique $t_1^*$ from equation $dC_2(t_1)/dt_1 = 0$, and the optimal total average cost and the optimal order quantity can be obtained by (24) and (29) at $t_1 = t_1^*$, respectively.

*Step 2.2.* If $\Delta_1 > 0$, $\Delta_2 > 0$, and $d^2C_2(t_1)/dt_1^2 > 0$, then $C_2(t_1)$ obtains its minimum at $t_1^* = \mu_1$, and compute the order quantity by (29).

*Step 2.3.* If $\Delta_1 < 0$, $\Delta_2 \leq 0$, and $d^2C_2(t_1)/dt_1^2 > 0$, then $C_2(t_1)$ obtains its minimum at $t_1^* = \mu_2$, and compute the order quantity by (29).

*Step 2.4.* If the sign of $d^2C_2(t_1)/dt_1^2$ is indefinite, then solve equation $dC_2(t_1)/dt_1 = 0$. If the solution is not unique, by comparing their average costs $C_2(t_1)$ to obtain the optimal replenishment point, compute the optimal order quantity by (29).

*Step 3.* Calculate $d^2C_3(t_1)/dt_1^2$.

*Step 3.1.* If $\Delta_2 < 0$ and $d^2C_3(t_1)/dt_1^2 > 0$, seek the unique $t_1^*$ from equation $dC_3(t_1)/dt_1 = 0$, and the total average cost and the order quantity can be obtained by (37) and (42), respectively.

*Step 3.2.* If $\Delta_2 \geq 0$ and $d^2C_3(t_1)/dt_1^2 > 0$, then $C_3(t_1)$ obtains its minimum at $t_1^* = \mu_2$, and compute the order quantity by (42).

*Step 3.3.* If the sign of $d^2C_3(t_1)/dt_1^2$ is indefinite, then solve equation $dC_3(t_1)/dt_1 = 0$. If the solution is not unique, by comparing their average costs $C_3(t_1)$ to obtain the optimal replenishment point, compute the order quantity by (42).

*Step 4.* Find $TC(t_1^*) = \min\{C_1(t_1^*), C_2(t_1^*), C_3(t_1^*)\}$ and accordingly select the optimum $t_1^*$, and then obtain corresponding optimal order quantity $Q^*$.

## 4. Numerical Examples and Sensitivity Analysis

In order to demonstrate the procedure to obtain the optimal solution of the model, we present four examples for the model. Examples are based on piece-wise demand rate, such as $f(t) = a_1 + b_1 t$ and $g(t) = a_2 e^{-b_2 t}$.

*Example 1.* The parameter values of an inventory system are given as follows: $T = 12$ weeks, $\mu_1 = 4$ weeks, $\mu_2 = 8$ weeks, $\alpha = 0.005$, $\beta = 2$, $\delta = 0.04$, $a_1 = 30$ units, $b_1 = 5$ units, $a_2 = 100$ units, $A_0 = \$500$, $c_1 = \$2$, $c_2 = \$3$, $c_3 = \$12$, $c_4 = \$8$.

Since $\Delta_1 = 18.809 > 0$, $\Delta_2 = 222.869 > 0$, $d^2C_1(t_1)/dt_1^2 > 0$ for all $t_1 \in (0, \mu_1)$, $d^2C_2(t_1)/dt_1^2 > 0$ for all $t_1 \in [\mu_1, \mu_2)$, and $d^2C_3(t_1)/dt_1^2 > 0$ for all $t_1 \in [\mu_2, T)$, then we get local minimum value $C_2(\mu_1) = 763.891$ in $[\mu_1, \mu_2)$ and $C_3(u_2) = 1263.58$ in $[\mu_2, T)$. By solving equation $dC_1(t_1)/dt_1 = 0$, we have $t_1^* = 3.664$. From (10), we obtain $C_1(t_1^*) = 762.759$. Therefore, the optimal ordering quantity is $Q^* = 637.839$ and the total average cost is $TC(t_1^*) = 762.759$.

*Example 2.* The parameter values of an inventory system are given as follows: $T = 12$ weeks, $\mu_1 = 4$ weeks, $\mu_2 = 8$ weeks, $\alpha = 0.005$, $\beta = 2$, $\delta = 0.02$, $a_1 = 30$ units, $b_1 = 5$ units, $a_2 = 100$ units, $A_0 = \$500$, $c_1 = \$5$, $c_2 = \$10$, $c_3 = \$12$, $c_4 = \$8$.

Since $\Delta_1 = -257.889 < 0$, $\Delta_2 = 139.269 > 0$, $d^2C_1(t_1)/dt_1^2 > 0$ for all $t_1 \in (0, \mu_1]$, $d^2C_2(t_1)/dt_1^2 > 0$ for all $t_1 \in [\mu_1, \mu_2)$, and $d^2C_3(t_1)/dt_1^2 > 0$ for all $t_1 \in (\mu_2, T)$, then we get local minimum value $C_1(\mu_1) = 1805.883$ in $(0, \mu_1]$ and $C_3(u_2) = 1582.32$ in $[\mu_2, T)$. By solving equation $dC_2(t_1)/dt_1 = 0$, we have $t_1^* = 6.559$. From (24), we obtain $C_2(t_1^*) = 1480.324$. Therefore, the optimal ordering quantity is $Q^* = 571.827$ and the total average cost is $TC(t_1^*) = 1480.324$.

*Example 3.* The parameter values of an inventory system are given as follows: $T = 12$ weeks, $\mu_1 = 4$ weeks, $\mu_2 = 6$ weeks, $\alpha = 0.005$, $\beta = 2$, $\delta = 0.06$, $a_1 = 30$ units, $b_1 = 5$ units, $a_2 = 100$ units, $A_0 = \$500$, $c_1 = \$5$, $c_2 = \$10$, $c_3 = \$12$, $c_4 = \$8$.

Since $\Delta_1 = -249.329 < 0$, $\Delta_2 = -121.869 < 0$, $d^2C_1(t_1)/dt_1^2 > 0$ for all $t_1 \in (0, \mu_1]$, $d^2C_2(t_1)/dt_1^2 > 0$ for all $t_1 \in (\mu_1, \mu_2]$, and $d^2C_3(t_1)/dt_1^2 > 0$ for all $t_1 \in [\mu_2, T)$, then we get local minimum value $C_1(\mu_1) = 1417.134$ in $(0, \mu_1]$ and $C_2(u_2) = 1052.016$ in $(\mu_1, \mu_2]$. By solving equation $dC_3(t_1)/dt_1 = 0$, we have $t_1^* = 9.120$. From (37), we obtain $C_3(t_1^*) = 889.767$. Therefore, the optimal ordering quantity is $Q^* = 278.960$ and the total average cost is $TC(t_1^*) = 889.767$.

*Example 4.* The parameter values of an inventory system are given as follows: $T = 12$ weeks, $\mu_1 = 4$ weeks, $\mu_2 = 6$ weeks, $\alpha = 0.005$, $\beta = 1.6$, $\delta = 0.02$, $a_1 = 30$ units, $b_1 = 5$ units, $a_2 = 100$ units, $A_0 = \$500$, $c_1 = \$5$, $c_2 = \$10$, $c_3 = \$12$, $c_4 = \$8$.

Since $\Delta_1 = -197.263 < 0$, $\Delta_2 = -86.675 < 0$, $d^2C_4(t_1)/dt_1^2 > 0$ for all $t_1 \in (0, \mu_1]$, $d^2C_2(t_1)/dt_1^2 > 0$

Table 1: The sensitivity of $\delta$ for the model in Example 1.

| $\delta$ | 0 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 |
|---|---|---|---|---|---|---|---|---|---|
| $t_1^*$ | 3.4366 | 3.4905 | 3.5463 | 3.6042 | 3.6642 | 3.7263 | 3.7907 | 3.8574 | 3.9265 |
| $S^*$ | 515.0829 | 515.0829 | 511.4210 | 507.5965 | 503.6066 | 499.4433 | 495.0985 | 490.5641 | 485.8321 |
| $Q^*$ | 648.1034 | 648.1034 | 644.8369 | 641.4184 | 637.8392 | 634.0900 | 630.1614 | 626.0436 | 621.7267 |
| $\mathrm{TC}_2(t_1^*)$ | 775.2147 | 771.8437 | 768.4699 | 764.9983 | 761.4252 | 757.7473 | 753.9611 | 750.063 | 746.0498 |

Table 2: The sensitivity of $\alpha$ for the models in Example 1.

| $\alpha$ | 0 | 0.001 | 0.002 | 0.003 | 0.004 | 0.005 | 0.006 | 0.007 | 0.008 |
|---|---|---|---|---|---|---|---|---|---|
| $t_1^*$ | 3.0141 | 3.1325 | 3.2569 | 3.3870 | 3.5229 | 3.6642 | 3.8105 | 3.9614 | 4.1171 |
| $S^*$ | 415.8882 | 432.0629 | 448.9624 | 466.5543 | 484.7907 | 503.6066 | 522.9206 | 542.6345 | 575.5923 |
| $Q^*$ | 522.8982 | 543.9279 | 565.9745 | 589.0106 | 612.9883 | 637.8392 | 663.4729 | 689.7782 | 723.0000 |
| $\mathrm{TC}_2(t_1^*)$ | 638.5314 | 661.4087 | 685.1585 | 709.7639 | 735.1984 | 761.4252 | 788.3972 | 816.0570 | 845.0147 |

Table 3: The sensitivity of $\beta$ for the models in Example 1.

| $\beta$ | 1.4 | 1.5 | 1.6 | 1.7 | 1.8 | 1.9 | 2.0 | 2.1 | 2.2 |
|---|---|---|---|---|---|---|---|---|---|
| $t_1^*$ | 3.1114 | 3.1995 | 3.2513 | 3.3176 | 3.40332 | 3.5148 | 3.6613 | 3.8551 | 4.1128 |
| $S^*$ | 427.5053 | 440.1612 | 447.1691 | 456.2029 | 467.9084 | 483.1709 | 503.6066 | 529.5334 | 564.1669 |
| $Q^*$ | 538.4988 | 554.7959 | 563.9494 | 575.7494 | 591.0495 | 611.0248 | 637.8392 | 672.0262 | 717.9964 |
| $\mathrm{TC}_2(t_1^*)$ | 644.4596 | 670.3559 | 680.1679 | 692.9359 | 709.5918 | 731.4743 | 761.4252 | 797.6508 | 845.0147 |

for all $t_1 \in (\mu_1, \mu_2]$, and $d^2\mathrm{C}_3(t_1)/dt_1^2 > 0$ for all $t_1 \in [\mu_2, T)$, then we get local minimum value $\mathrm{C}_1(\mu_1) = 1252.59$ in $(0, \mu_1]$ and $\mathrm{C}_2(u_2) = 1052.016$ in $(\mu_1, \mu_2]$. By solving equation $d\mathrm{C}_3(t_1)/dt_1 = 0$, we have $t_1^* = 8.627$. From (37), we obtain $\mathrm{TC}(t_1^*) = 842.935$. Therefore, the optimal ordering quantity is $Q^* = 241.524$ obtained from (42), and the minimum cost $\mathrm{TC}(t_1^*) = 842.935$.

In order to clearly indicate the effects of parameters such as $\delta$, $\alpha$, and $\beta$ on $S^*$, $Q^*$, and $\mathrm{TC}_i(t_1^*)$ $(i = 1, 2)$, respectively, we study the sensitivity of different parameters under the studied inventory model. The sensitivity analysis is performed on the base of Example 1, and the results are shown in Tables 1–3, respectively.

From Table 1, we observe that inventory level $S^*$, order quantity $Q^*$, and the total average cost $\mathrm{TC}_i(t_1^*)$ $(i = 1, 2)$ gradually decrease as the shortage parameter $\delta$ increases for the model, and the replenishment time $t_1^*$ increases as $\delta$ increases for the model starting with shortage. Also, we observe that the values of $t_1^*$, $S^*$, $Q^*$, and $\mathrm{TC}(t_1^*)$ are slightly sensitive to the changes of $\delta$.

From Table 2, we find that when the deterioration parameter $\alpha$ increases, then $S^*$, $Q^*$, and $\mathrm{TC}(t_1^*)$ increase for both models, and the replenishment time $t_1^*$ increases for the model starting with shortage. We see that the values of $t_1^*$, $S^*$, $Q^*$, and $\mathrm{TC}(t_1^*)$ all are temperately sensitive to the changes of $\alpha$ for the model.

From Table 3, we find that $S^*$, $Q^*$, and $\mathrm{TC}(t_1^*)$ increase as the deterioration parameter $\beta$ increases for both models. Also, we find that the replenishment time $t_1^*$ increases as $\beta$ increases for the model starting with shortage. We see that the values of $t_1^*$, $S^*$, $Q^*$, and $\mathrm{TC}(t_1^*)$ all are moderately sensitive to the changes of $\beta$ for the model.

## 5. Conclusion

This paper studies an inventory model for Weibull-distributed deterioration with trapezoidal type demand rate and partial backlogging. First, an optimal replenishment strategy for deteriorating items with the model starting with shortage is proposed. Then, numerical examples are provided to illustrate the theoretical results. From the analysis of Tables 1–3, it can be found that the replenishment time point $t_1^*$, order quantity $Q^*$, and the total average cost $\mathrm{TC}(t_1^*)$ are moderately sensitive to the changes of $\alpha$ and $\beta$ and lowly sensitive to the changes of $\delta$, respectively. However, if the order cycle ($T$) is a decision variable, such inventory model will be of more practical significance, which provides us an interesting topic in future.

## Competing Interests

The author declares that there are no competing interests.

## Acknowledgments

The research is supported partly by National Natural Science Foundation of China (Grant no. 71502100).

## References

[1] P. M. Ghare and G. F. Schrader, "A model for exponentially decaying inventories," *Journal of Industrial Engineering*, vol. 14, no. 3, pp. 238–243, 1963.

[2] R. P. Covert and G. C. Philip, "An EOQ model for items with weibull distribution deterioration," *AIIE Transaction*, vol. 5, no. 4, pp. 323–326, 1973.

[3] H.-M. Wee, J. C. P. Yu, and S. T. Law, "Two-warehouse inventory model with partial backordering and Weibull distribution deterioration under inflation," *Journal of the Chinese Institute of Industrial Engineers*, vol. 22, no. 6, pp. 451–462, 2005.

[4] G. K. Yang, R. Lin, J. Lin, K.-C. Hung, P. Chu, and W. Chouhuang, "Note on inventory models with Weibull distribution deterioration," *Production Planning & Control*, vol. 22, no. 4, pp. 437–444, 2011.

[5] N. H. Shah, H. N. Soni, and K. A. Patel, "Optimizing inventory and marketing policy for non-instantaneous deteriorating items with generalized type deterioration and holding cost rates," *Omega*, vol. 41, no. 2, pp. 421–430, 2013.

[6] C.-T. Yang, L.-Y. Ouyang, H.-F. Yen, and K.-L. Lee, "Joint pricing and ordering policies for deteriorating item with retail price-dependent demand in response to announced supply price increase," *Journal of Industrial and Management Optimization*, vol. 9, no. 2, pp. 437–454, 2013.

[7] R. M. Hill, "Inventory models for increasing demand followed by level demand," *Journal of the Operational Research Society*, vol. 46, no. 10, pp. 1250–1259, 1995.

[8] B. Mandal and A. K. Pal, "Order level inventory system with ramp type demand rate for deteriorating items," *Journal of Interdisciplinary Mathematics*, vol. 1, no. 1, pp. 49–66, 1998.

[9] K.-S. Wu, "An EOQ inventory model for items with Weibull distribution deterioration, ramp type demand rate and partial backlogging," *Production Planning & Control*, vol. 12, no. 8, pp. 787–793, 2001.

[10] K. Skouri, I. Konstantaras, S. Papachristos, and I. Ganas, "Inventory models with ramp type demand rate, partial backlogging and Weibull deterioration rate," *European Journal of Operational Research*, vol. 192, no. 1, pp. 79–92, 2009.

[11] Y. Tan and M. X. Weng, "A discrete-in-time deteriorating inventory model with time-varying demand, variable deterioration rate and waiting-time-dependent partial backlogging," *International Journal of Systems Science*, vol. 44, no. 8, pp. 1483–1493, 2013.

[12] S. Pal, G. S. Mahapatra, and G. P. Samanta, "A production inventory model for deteriorating item with ramp type demand allowing inflation and shortages under fuzziness," *Economic Modelling*, vol. 46, no. 2, pp. 334–345, 2015.

[13] M. A. Ahmed, T. A. Al-Khamis, and L. Benkherouf, "Inventory models with ramp type demand rate, partial backlogging and general deterioration rate," *Applied Mathematics and Computation*, vol. 219, no. 9, pp. 4288–4307, 2013.

[14] S. Agrawal, S. Banerjee, and S. Papachristos, "Inventory model with deteriorating items, ramp-type demand and partially backlogged shortages for a two warehouse system," *Applied Mathematical Modelling*, vol. 37, no. 20-21, pp. 8912–8929, 2013.

[15] T. Singh and H. Pattnayak, "An EOQ inventory model for deteriorating items with varying trapezoidal type demand rate and Weibull distribution deterioration," *Journal of Information and Optimization Sciences*, vol. 34, no. 6, pp. 341–360, 2013.

[16] J. Wu, L.-Y. Ouyang, L. E. Cárdenas-Barrón, and S. K. Goyal, "Optimal credit period and lot size for deteriorating items with expiration dates under two-level trade credit financing," *European Journal of Operational Research*, vol. 237, no. 3, pp. 898–908, 2014.

[17] M. Rabbani, N. P. Zia, and H. Rafiei, "Coordinated replenishment and marketing policies for non-instantaneous stock deterioration problem," *Computers & Industrial Engineering*, vol. 88, pp. 49–62, 2015.

[18] M. B. Cheng and G. Q. Wang, "A note on the inventory model for deteriorating items with trapezoidal type demand rate," *Computers & Industrial Engineering*, vol. 56, no. 4, pp. 1296–1300, 2009.

[19] M. B. Cheng, B. X. Zhang, and G. Q. Wang, "Optimal policy for deteriorating items with trapezoidal type demand and partial backlogging," *Applied Mathematical Modelling*, vol. 35, no. 7, pp. 3552–3560, 2011.

[20] R. Uthayakumar and M. Rameswari, "An economic production quantity model for defective items with trapezoidal type demand rate," *Journal of Optimization Theory and Applications*, vol. 154, no. 3, pp. 1055–1079, 2012.

[21] K.-P. Lin, "An extended inventory models with trapezoidal type demands," *Applied Mathematics and Computation*, vol. 219, no. 24, pp. 11414–11419, 2013.

[22] J. Lin, K.-C. Hung, and P. Julian, "Technical note on inventory model with trapezoidal type demand," *Applied Mathematical Modelling*, vol. 38, no. 19-20, pp. 4941–4948, 2014.

[23] L. Zhao, "An inventory model under trapezoidal type demand, Weibull-distributed deterioration, and partial backlogging," *Journal of Applied Mathematics*, vol. 2014, Article ID 747419, 10 pages, 2014.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 1475148, 9 pages
http://dx.doi.org/10.1155/2016/1475148

*Research Article*

# An Analysis of Bank Service Satisfaction Based on Quantile Regression and Grey Relational Analysis

**Wen-Tsao Pan and Yungho Leu**

*Department of Information Management, National Taiwan University of Science and Technology, Taipei 10607, Taiwan*

Correspondence should be addressed to Yungho Leu; yhl@cs.ntust.edu.tw

Received 6 January 2016; Accepted 13 March 2016

Academic Editor: Feng Yang

Copyright © 2016 W.-T. Pan and Y. Leu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bank service satisfaction is vital to the success of a bank. In this paper, we propose to use the grey relational analysis to gauge the levels of service satisfaction of the banks. With the grey relational analysis, we compared the effects of different variables on service satisfaction. We gave ranks to the banks according to their levels of service satisfaction. We further used the quantile regression model to find the variables that affected the satisfaction of a customer at a specific quantile of satisfaction level. The result of the quantile regression analysis provided a bank manager with information to formulate policies to further promote satisfaction of the customers at different quantiles of satisfaction level. We also compared the prediction accuracies of the regression models at different quantiles. The experiment result showed that, among the seven quantile regression models, the median regression model has the best performance in terms of RMSE, RTIC, and CE performance measures.

## 1. Introduction

In response to the requirement of opening market and to promote the economic growth, Taiwanese government has actively formulated policies to promote the financial liberalization in Taiwan. As a result, many mergers of different banks have occurred in the past few years. Bank mergers may change the perception of a customer on the service quality of a bank. Therefore, a bank manager should improve the service quality of his clerks in order to promote the customer's satisfaction. High customer satisfaction on a bank can attract customers to continue doing their business with the bank and thus establishes the customers' loyalty to the bank. To effectively improve the service quality of a bank so as to promote the customer's satisfaction has become an important issue of the banking industry today.

Different from the existing studies on bank service satisfaction [1, 2], we use the data on service satisfaction of the public and private banks in Taiwan in this study. The questionnaire for data collection was designed by the Louis Harris International in 1995. In this study, we first collected the data on service satisfaction using the questionnaire. Then, we used the grey relational analysis to examine the customer satisfaction on the service of the banks, including the public and the private banks in Taiwan. With the grey relational analysis, we found the variables that affected the levels of service satisfaction of the public banks and the private banks. We also ranked the banks according to their levels of service satisfaction. Finally, we conducted a quantile regression analysis on the 36 banks of the dataset to find the variables that affected the levels of satisfaction for customers at different quantiles of service satisfaction. The results can be used by a bank manager to formulate different policies to further promote the satisfaction of the customers at different quantile of service satisfaction. Finally, we compared the performance of different quantile regressions in terms of three different measures of performance evaluation.

The remainder of this paper is organized as follows. Section 1 introduces the research motivation and purpose. Section 2 reviews the literature on grey theory and quantile regression. Section 3 discusses the dataset and results of the empirical study. Section 4 gives the conclusion and the suggestion of this paper.

## 2. Research Method

*2.1. Grey Relational Analysis.* First proposed by Deng [3], the grey relational analysis can be used to measure the similarity

Table 1: An example dataset.

| Sequence | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ | $k = 6$ |
|---|---|---|---|---|---|---|
| $x_0$ | 1.000 | 1.100 | 2.000 | 2.250 | 3.000 | 4.000 |
| $x_1$ | 1.000 | 1.166 | 1.834 | 2.000 | 2.340 | 3.000 |
| $x_2$ | 1.000 | 1.125 | 1.075 | 1.375 | 1.625 | 1.750 |
| $x_3$ | 1.000 | 1.000 | 0.700 | 0.800 | 0.900 | 1.200 |

between two sequences of factors. Assume that the state of a system can be modeled by a sequence of factors. Two system states are said to be similar if the values of their corresponding factors are similar. In a real life application, a standard sequence is selected to calculate the relational grades of a set of inspected sequences. Usually, the standard sequence is composed of the maximal values of each of the factors in the current dataset. By referring to the standard sequence, the grey relational analysis gives each inspected sequence a grey relational grade (GRG). The more an inspected sequence is similar to the standard sequence, the larger the GRG of the inspected sequence is. Since the standard sequence usually represents a state of the system with maximum performance, a sequence with a large GRG is therefore better than a sequence with a small GRG. The grey relational analysis has been widely used in performance evaluation [4–6].

The grey relational grade of an inspected sequence is calculated through the following steps.

*Step 1* (define the sequences). To perform the grey relational analysis, we need to define the sequences for comparison. To illustrate, we use Table 1 as an example. Each column in Table 1 represents a factor of the system state while each row in Table 1 represents a sequence of the system state. Among all the sequences of the system state, we choose or create a sequence, called the standard sequence, to measure the performance of the other sequences which are called the inspected sequences. A sequence $x_i$ is denoted by a set $\{x_i(k) \mid k = 1, \ldots, n\}$. Without loss of generality, we choose $x_0$ as the standard sequence.

*Step 2* (standardization). Since different factors may have different ranges of values which are not comparable, we need to standardize the values of all factors so that values from different factors are comparable. One way of standardization is to divide all the values of a factor by the maximum value of the factor. By doing so, values of all factors are all less than one and greater than zero.

*Step 3* (calculate the grey relational coefficients). The grey relational coefficient of the $k$th factor of sequence $x_i$ is calculated by the following equation:

$$\xi_i(k) = \frac{\min_i \min_k \left|x_0(k) - x_i(k)\right| + \rho \max_i \max_k \left|x_0(k) - x_i(k)\right|}{\left|x_0(k) - x_i(k)\right| + \rho \max_i \max_k \left|x_0(k) - x_i(k)\right|}. \quad (1)$$

Let $\Delta_{0i}(k)$ denote $|x_0(k) - x_i(k)|$. Equation (1) can be written as

$$\xi_i(k) = \frac{\min_i \min_k \Delta_{0i}(k) + \rho \max_i \max_k \Delta_{0i}(k)}{\left|x_0(k) - x_i(k)\right| + \rho \max_i \max_k \Delta_{0i}(k)}. \quad (2)$$

Note that to find the value of $\min_i \min_k \Delta_{0i}(k)$, we find, for inspected sequence $x_i$, the minimum value of the differences between factor $k$ of $x_0$ and factor $k$ of $x_i$, $k = 1$ to $n$. Then, we find the minimum value among the minimum values of all the inspected sequences. The value of $\max_i \max_k |x_0(k) - x_i(k)|$ can be found in a similar way. The rationale behind (2) is that we use the global minimum and the global maximum of the componentwise differences between the standard sequence and all the inspected sequences to rescale the difference between a factor of an inspected sequence and the same factor of the standard sequence. And then, we use the rescaled differences, also called the grey relational coefficients, to replace the values of the original factors. The value of $\rho$ belongs to $(0, \infty)$ and is usually set to 0.5.

*Step 4* (find the grey relational grade). The grey relational grade $r_i$ of inspected sequence $x_i$ is defined to be the average of all its grey relational coefficients. The grey relational grade $r_i$ of $x_i$ is calculated according to the following equation:

$$r_i = \frac{1}{n} \sum_{k=1}^{n} \xi_i(k). \quad (3)$$

To illustrate how to calculate the grey relational grade of an inspected sequence, we use the dataset in Table 1.

*Step 1* (define the sequence). We assume that $x_0$ is the standard sequence.

*Step 2* (standardization). We assume that the factors of the dataset in Table 1 are comparable. No standardization is required.

*Step 3* (calculate the grey relational coefficients). The componentwise differences between the standard sequence and all the inspected sequences are shown in the following:

(1) Componentwise differences between $x_0$ and $x_1$:

$$\begin{aligned} \Delta_{01}(1) &= 0.000, \\ \Delta_{01}(2) &= 0.066, \\ \Delta_{01}(3) &= 0.166, \\ \Delta_{01}(4) &= 0.250, \\ \Delta_{01}(5) &= 0.660, \\ \Delta_{01}(6) &= 1.000. \end{aligned} \quad (4)$$

(2) Componentwise differences between $x_0$ and $x_2$:

$$\begin{aligned}
&\Delta_{02}(1) = 0.000,\\
&\Delta_{02}(2) = 0.025,\\
&\Delta_{02}(3) = 0.925,\\
&\Delta_{02}(4) = 0.875,\\
&\Delta_{02}(5) = 1.375,\\
&\Delta_{02}(6) = 2.250.
\end{aligned} \tag{5}$$

(3) Componentwise differences between $x_0$ and $x_3$:

$$\begin{aligned}
&\Delta_{03}(1) = 0.000,\\
&\Delta_{03}(2) = 0.100,\\
&\Delta_{03}(3) = 1.300,\\
&\Delta_{03}(4) = 1.450,\\
&\Delta_{03}(5) = 2.100,\\
&\Delta_{03}(6) = 2.800.
\end{aligned} \tag{6}$$

We find the following vectors of differences:

(1) $\Delta_{01} = (0.000, 0.066, 0.166, 0.250, 0.660, 1.000)$;
(2) $\Delta_{02} = (0.000, 0.025, 0.925, 0.875, 1.375, 2.250)$;
(3) $\Delta_{03} = (0.000, 0.100, 1.300, 1.450, 2.100, 2.800)$.

The global minimum and the global maximum of all the differences are 0.000 and 2.800, respectively. Let $\rho = 0.5$. We write

$$\xi_i(k) = \frac{0 + (0.5) \times (2.8)}{\Delta_{0i}(k) + (0.5) \times (2.8)}. \tag{7}$$

We then calculate the grey relational coefficients of the inspected sequence $x_1$ in the following:

$$\begin{aligned}
&\xi_1(1) = 1.000,\\
&\xi_1(2) = 0.9549,\\
&\xi_1(3) = 0.8939,\\
&\xi_1(4) = 0.8484,\\
&\xi_1(5) = 0.6796,\\
&\xi_1(6) = 0.5833.
\end{aligned} \tag{8}$$

Finally, we find the grey relational grade (GRG) of the inspected sequence $x_1$ in the following:

$$\begin{aligned}
r_1 = \frac{1}{6}(1 + 0.9549 + 0.8939 + 0.8484 + 0.6796\\
+ 0.5833) = 0.8266.
\end{aligned} \tag{9}$$

The grey relational grades of sequences $x_2$ and $x_3$ can be found in the same way.

*2.2. Quantile Regression.* Koenker and Bassett [7] proposed the quantile regression method. Quantile regression is used to explore the effects of explanatory variables ($X$) on the explained variable ($Y$) at different quantiles of the explained variable. Different from the conventional linear models (such as the least squares method) that predict the mean of the explained variable given a specific value of each explanatory variable, a quantile regression model predicts the value of the explained variable at a specific quantile of the explained variable giving a specific value of each explanatory variable. Therefore, the quantile regression method facilitates the study of the effects of the explanatory variables on different quantiles of the explained variable. The quantile regression method has been widely used in many applications [8–10].

## 3. Empirical Analysis

*3.1. Sample Dataset and Variables.* Based on a questionnaire designed by the Louis Harris International Taiwan, we conducted a survey on the bank service quality of Taiwan in 2013. We analyzed the satisfaction of bank service in Taiwan by using the collected questionnaire data. The dataset consists of 8 public banks and 28 private banks. The variables designed for bank service satisfaction include "clerk service attitude" ($X1$), "customer-oriented service" ($X2$), "flexibility in handling customer inquiries" ($X3$), "service efficiency" ($X4$), "interest rate favoring customers" ($X5$), "transaction errors" ($X6$), "convenience of branch locations" ($X7$), "service traffic flow" ($X8$), and "service counter design" ($X9$). The value of each variable of a bank is obtained by averaging the scores of the same variable on all the questionnaires for the bank. The dataset for the public bank is shown in Table 2. The descriptive statistics of the variables are shown in Table 3. The source dataset is divided into three datasets. The first dataset contains all the 36 banks including the private and public banks; the second dataset contains only the 28 private banks; the third dataset contains the 8 public banks whose data are shown in Table 2.

*3.2. Grey Relational Analysis on Bank Service Satisfaction.* This study applied the Matlab tool box for grey relational analysis [11] to obtain a grey relational grade (GRG) for a bank to represent the level of satisfaction of the bank. In this study, we performed two analyses using the grey relational analysis. In the first analysis, we used the grey relational analysis to find the satisfaction levels of different variables. To do that, we first transposed the dataset so that the samples of the dataset become the variables and the variables become the samples of the transposed dataset. Figures 1(a) and 1(b) show the results of grey relational analysis for the eight public banks and the 28 private banks, respectively. The bold dotted lines represent the standard sequences and the thin lines represent the inspected sequences, which are samples to be evaluated. The more an inspected sequence is close to the standard sequence, the larger its satisfaction grade is. According to Figure 1(a), the top three most satisfied variables among the nine surveyed variables for the public banks are "convenience of branch locations" (GRG = 0.8775),

Table 2: The dataset for public banks.

| Bank names | $X1$ | $X2$ | $X3$ | $X4$ | $X5$ | $X6$ | $X7$ | $X8$ | $X9$ |
|---|---|---|---|---|---|---|---|---|---|
| Farmers Bank of China | 4.31 | 3.63 | 3.67 | 3.89 | 3.31 | 4.52 | 3.82 | 3.7 | 3.84 |
| Land Bank of Taiwan | 3.98 | 3.44 | 3.52 | 3.74 | 3.77 | 4.62 | 2.67 | 2.87 | 3.65 |
| First Commercial Bank | 4.04 | 3.52 | 3.31 | 3.73 | 3.41 | 4.63 | 3.61 | 3.29 | 3.31 |
| Taiwan Business Bank | 3.69 | 3.14 | 3.21 | 3.38 | 3.28 | 4.81 | 3.64 | 3.09 | 4 |
| Bank of Taiwan | 3.86 | 3.24 | 3.44 | 3.67 | 3.16 | 4.67 | 3.08 | 3.19 | 3.95 |
| Hua Nan Bank | 3.83 | 3.12 | 3.28 | 3.63 | 2.92 | 4.57 | 3.92 | 3.25 | 3.35 |
| Chang Hwa Bank | 3.53 | 2.88 | 3.01 | 3.53 | 2.72 | 4.78 | 3.6 | 3.25 | 3.8 |
| Taiwan Cooperative Bank | 3.41 | 2.85 | 3.1 | 3.03 | 3.06 | 4.46 | 3.57 | 3.12 | 2.89 |

Table 3: Descriptive statistics of the variables in three different groups of data.

| | Variables | $X1$ | $X2$ | $X3$ | $X4$ | $X5$ | $X6$ | $X7$ | $X8$ | $X9$ |
|---|---|---|---|---|---|---|---|---|---|---|
| First dataset (36) | Max. | 4.65 | 4.02 | 4.06 | 4.29 | 4.12 | 4.92 | 3.98 | 4.16 | 4.63 |
| | Min. | 3.10 | 2.68 | 2.96 | 2.80 | 2.72 | 3.28 | 2.67 | 2.87 | 2.89 |
| | Avg. | 3.97 | 3.38 | 3.46 | 3.66 | 3.39 | 4.41 | 3.45 | 3.46 | 3.80 |
| | Std. | 0.34 | 0.28 | 0.23 | 0.33 | 0.29 | 0.46 | 0.33 | 0.30 | 0.39 |
| Second dataset (28) | Max. | 4.65 | 4.02 | 4.06 | 4.29 | 4.12 | 4.92 | 3.98 | 4.16 | 4.63 |
| | Min. | 3.10 | 2.68 | 2.96 | 2.80 | 3.00 | 3.28 | 2.81 | 2.94 | 3.02 |
| | Avg. | 4.01 | 3.42 | 3.50 | 3.69 | 3.44 | 4.34 | 3.44 | 3.53 | 3.85 |
| | Std. | 0.34 | 0.27 | 0.22 | 0.34 | 0.27 | 0.50 | 0.32 | 0.29 | 0.37 |
| Third dataset (8) | Max. | 4.31 | 3.63 | 3.67 | 3.89 | 3.77 | 4.81 | 3.92 | 3.70 | 4.00 |
| | Min. | 3.41 | 2.85 | 3.01 | 3.03 | 2.72 | 4.46 | 2.67 | 2.87 | 2.89 |
| | Avg. | 3.83 | 3.23 | 3.32 | 3.58 | 3.20 | 4.63 | 3.49 | 3.22 | 3.60 |
| | Std. | 0.27 | 0.27 | 0.20 | 0.25 | 0.30 | 0.11 | 0.39 | 0.22 | 0.36 |

"clerk service attitude" (GRG = 0.5607), and "service traffic flow" (GRG = 0.509), while the three least satisfied variables are "customer-oriented service" (GRG = 0.4364), "service counter design" (GRG = 0.434), and "transaction errors" (GRG = 0.4338).

In contrast, Figure 1(b) shows that, for the 28 private banks, the top three most satisfied variables are "convenience of branch locations" (GRG = 0.94), "clerk service attitude" (GRG = 0.7302), and "service traffic flow" (GRG = 0.6594), while the three least satisfied variables are "transaction errors" (GRG = 0.5266), "service efficiency" (GRG = 0.5207), and "customer-oriented service" (GRG = 0.5038). Therefore, in regard to their advantages, both the private and public banks enjoyed "branch locations," "clerk service attitude," and "service traffic flow." In regard to their disadvantages, both public and private banks need to improve their "customer-oriented service" and "transaction errors." Furthermore, the "service efficiency" for the private banks and the "counter design" for the public banks need to be improved too.

In the second analysis, we compare the service satisfaction of different banks. Figure 2(a) shows the result of the grey relational analysis for the eight public banks. It shows that the top three banks in terms of service satisfaction are "Farmers Bank of China" (GRG = 0.8857), "Land Bank of Taiwan" (GRG = 0.6933), and "First Commercial Bank" (GRG = 0.6841); the bottom three banks are "Hua Nan Bank" (GRG = 0.6307), "Chang Hwa Bank" (GRG = 0.5996), and "Taiwan Cooperative Bank" (GRG = 0.4932). Figure 2(b) shows the grey relational analysis for the 28 private banks. It shows that the top three banks in terms of service satisfaction are "E.SUN Commercial Bank" (GRG = 0.8586), "Union Bank of Taiwan" (GRG = 0.8138), and "Yuanta Commercial Bank" (GRG = 0.7916); the bottom three banks included "Bank of Taipei" (GRG = 0.4669), "Jih Sun International Commercial Bank" (GRG = 0.4521), and "Mega International Commercial Bank" (GRG = 0.4342).

Lastly, we perform grey relational analysis on all the 36 banks, including the public and private banks. Figure 2(c) shows that the top three banks, among both public and private banks, are "E.SUN Commercial Bank" (GRG = 0.8643), "Union Bank of Taiwan" (GRG = 0.8217), and "Yuanta Commercial Bank" (GRG = 0.7993). Note that all of them are private banks. On the other hand, the bottom three banks include one public bank and two private banks which are "Taiwan Cooperative Bank" (GRG = 0.4773), "Jih Sun International Commercial Bank" (GRG = 0.4666), and "Mega International Commercial Bank" (GRG = 0.4483). Therefore, according to the above analysis, private banks usually receive better service satisfaction from their customers than public banks do.

*3.3. Quantile Regression Analysis of Different Factors on Service Satisfaction.* In addition to the grey relational analysis, in this study we perform quantile regression analysis to explore the factors that influence service satisfaction for

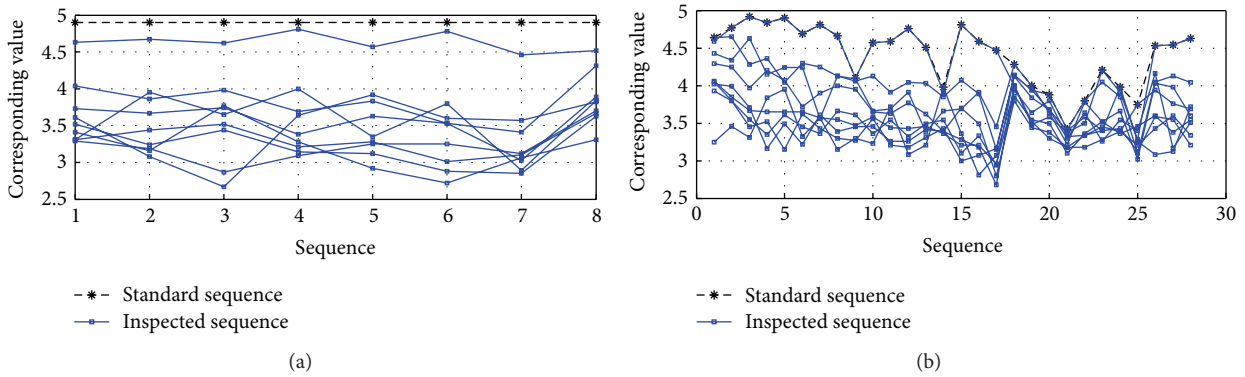



Figure 1: Results of grey relational analysis for different surveyed variables.

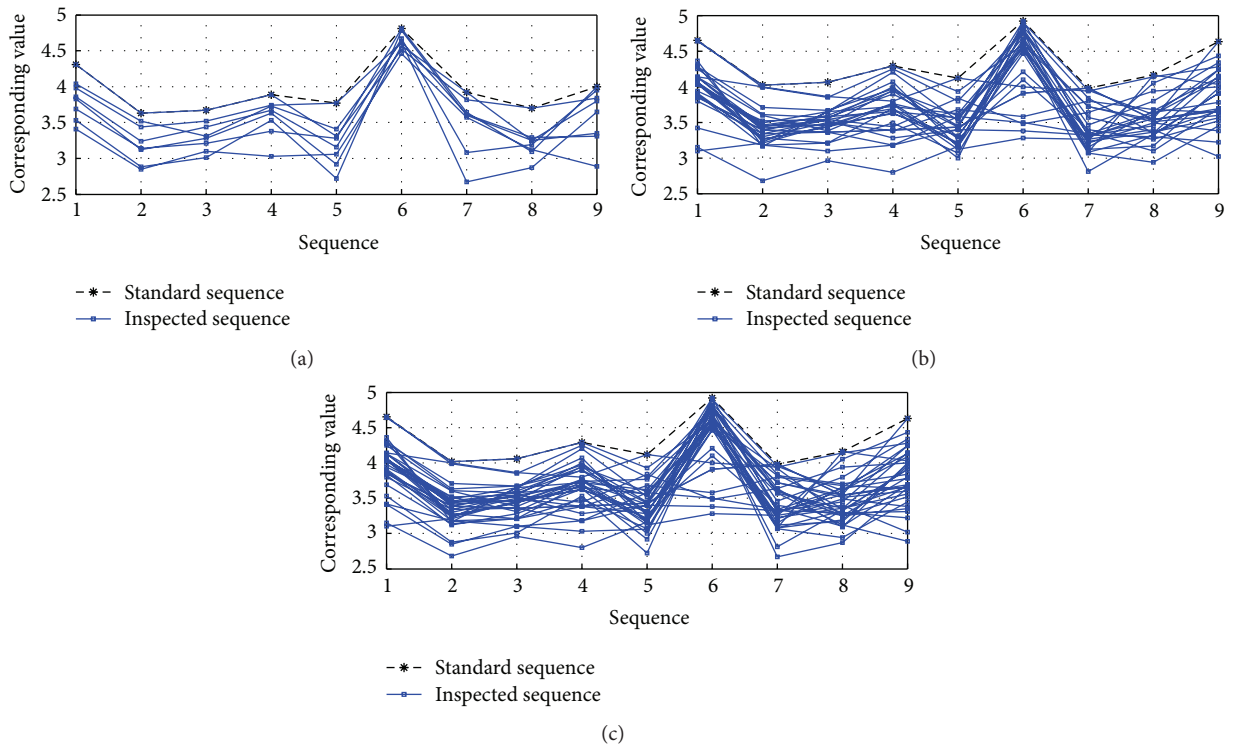



Figure 2: Grey relational analysis on the satisfaction of eight public banks (a), 28 private banks (b), and all 36 banks (c).

those samples with low-, medium-, and high-levels of service satisfaction. To do that, we discretize the values of all the variables into five intervals. Each value of a variable is transformed into its corresponding label of the interval. In this study, the top 20 percent values of a variable are given the largest label, which is 5, and so on. The transformed dataset for the public banks is shown in Table 4. Note that, in Table 4, GRG denotes the grey relational grade of a sample bank.

In this study, we use the STATA software to conduct the 0.25 quantile, median, and 0.75 quantile regressions on the dataset of all the 36 banks. By using grey relational grades of the above grey relational analysis as the dependent variables ($Y$) and nine bank service satisfaction-related variables as the independent variables ($X1$–$X9$) for quantile regression analysis, this study determined which questionnaire survey question items (variables) affected the bank service satisfaction performance for the customers at a specific quantile of service satisfaction. The quantile regression analysis is conducted mainly for three different quantiles including the 0.25 quantile, median, and 0.75 quantile. According to the analysis results shown in Table 5, "bank clerk service attitude" ($X1$), "customer-oriented service" ($X2$), and "flexibility in handling customer inquiries" ($X3$) had no effects on the bank service satisfaction performance at different quantiles of service satisfaction. However, for "service efficiency" ($X4$), its coefficient is significantly different from zero at 10% significance level for all different quantile regression models. This result indicates that the waiting time for bank service would affect the levels of service satisfaction for customers with different levels of service satisfaction. Regarding

Table 4: The transformed dataset of the public bank dataset.

| Bank names | $X1$ | $X2$ | $X3$ | $X4$ | $X5$ | $X6$ | $X7$ | $X8$ | $X9$ | GRG |
|---|---|---|---|---|---|---|---|---|---|---|
| Farmers Bank of China | 5 | 5 | 5 | 5 | 3 | 1 | 5 | 5 | 5 | 0.8857 |
| Land Bank of Taiwan | 4 | 4 | 4 | 5 | 5 | 3 | 1 | 1 | 4 | 0.6933 |
| First Commercial Bank | 4 | 5 | 3 | 5 | 4 | 3 | 4 | 3 | 2 | 0.6841 |
| Taiwan Business Bank | 2 | 2 | 2 | 3 | 3 | 5 | 4 | 2 | 5 | 0.6525 |
| Bank of Taiwan | 3 | 3 | 4 | 4 | 3 | 3 | 2 | 2 | 5 | 0.6631 |
| Hua Nan Bank | 3 | 2 | 3 | 4 | 1 | 2 | 5 | 3 | 3 | 0.6307 |
| Chang Hwa Bank | 1 | 1 | 1 | 3 | 1 | 5 | 4 | 3 | 5 | 0.5996 |
| Taiwan Cooperative Bank | 1 | 1 | 1 | 1 | 2 | 4 | 4 | 2 | 1 | 0.4932 |

Table 5: Quantile regression coefficients at different quantiles.

| Var. | Model | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | OLS Adj. $R^2 = 0.9528$ | | | Q0.25 Pseudo $R^2 = 0.8246$ | | | Q0.5 Pseudo $R^2 = 0.8135$ | | | Q0.75 Pseudo $R^2 = 0.8304$ | | |
| | Coef. | $T$ | Sig. | Coef. | $T$ | Sig. | Coef. | $T$ | Sig. | Coef. | $T$ | Sig. |
| $X1$ | −0.042 | −1.62 | | −0.038 | −0.75 | | −0.044 | −1.33 | | −0.0492 | −0.96 | |
| $X2$ | 0.072 | 2.58 | ∗ | 0.068 | 1.46 | | 0.062 | 1.52 | | 0.075 | 1.14 | |
| $X3$ | 0.064 | 1.49 | | 0.102 | 1.55 | | 0.057 | 0.69 | | −0.003 | −0.03 | |
| $X4$ | 0.059 | 2.5 | ∗ | 0.054 | 1.88 | ∧ | 0.051 | 2.01 | ∧ | 0.085 | 2.02 | ∧ |
| $X5$ | 0.085 | 3.37 | ∗∗ | 0.073 | 3.98 | ∗∗ | 0.071 | 2.53 | ∗ | 0.128 | 2.89 | ∗∗ |
| $X6$ | 0.071 | 4.98 | ∗∗ | 0.062 | 3.62 | ∗∗ | 0.065 | 3.27 | ∗∗ | 0.083 | 3.23 | ∗∗ |
| $X7$ | 0.060 | 5.05 | ∗∗ | 0.065 | 7.32 | ∗∗ | 0.063 | 4.53 | ∗∗ | 0.057 | 2.24 | ∗ |
| $X8$ | 0.055 | 2.62 | ∗ | 0.044 | 2.38 | ∗ | 0.041 | 1.5 | | 0.044 | 0.94 | |
| $X9$ | 0.050 | 3.39 | ∗∗ | 0.042 | 2.94 | ∗∗ | 0.056 | 3.4 | ∗∗ | 0.038 | 1.19 | |

Note: ∧ means that it is significant under 10% significance level; ∗ means that it is significant under 5% significance level; ∗∗ means that it is significant under 1% significance level.

the variable of "interest rate favoring customers" ($X5$), according to Table 5, its coefficient is different from zero at 1% significance level in all the three quantile regression models. Figure 3 shows that, for the customers at the 0.75 quantile of service satisfaction, the ordinary least squares (OLS) (the dash line on the plot) regression tends to underestimate the effect of "interest rate favoring customers" ($X5$) on service satisfaction. Therefore, customers with a high level of service satisfaction will be especially concerned about whether the bank can adjust the interest rates to their advantages. This signifies that a bank manager needs to pay more attention to the policy of interest rate adjustment to further promote the service satisfaction of his customers who already have a high level of service satisfaction.

For the variable of "transaction errors" ($X6$), Table 5 shows that its coefficients are different from zero for all the three different quantile regression models at 1% significance level. Furthermore, Figure 3 shows that for high quantiles of service satisfaction the ordinary least squares regression model tends to underestimate the effect of $X6$. Figure 4 shows the results of 0.25, 0.5, and 0.75 quantile regression models in terms of $X6$ and $(X6)^2$ (stands for $X6$ squared). In the plot, the satisfaction levels are calculated according to the regressed quadratic equations of variable $X6$. The trends of the curves indicate that fewer transaction errors would result in higher satisfaction levels (in terms of GRG). Note that a large value of $X6$ in Figure 4 represents a case with fewer transaction errors and, therefore, receives a higher satisfaction level from the customers. Also note that, for the median regression model, the level of service satisfaction increases as the number of transaction errors decreases. However, the trend is reversed as the number of transaction errors is further reduced to be below a certain number, a case that is counter to our intuition. This represents that the quadric equation of $X6$ is not adequate to capture the relationship between the service satisfaction levels and $X6$.

Table 5 shows that the coefficient of $X7$ (convenience of branch location) is significantly different from zero at 1% significance level for both 0.25 quantile and median regression models. For the 0.75 quantile regression model, its coefficient is significantly different from zero at 5% significance level. The results show that the convenience of branch location can affect the service satisfaction, regardless of different quantiles of service satisfaction. In addition, Figure 3 shows that the ordinary least squares regression overestimates the effect of $X7$ on service satisfaction for high quantiles of service satisfaction. As to the variable of "service traffic flow" ($X8$), Table 5 shows that its coefficient is different from zero only for the 0.25 quantile regression model at 5% significance level. In other words, improving the service traffic flow can promote the service satisfaction for customers with low levels of service satisfaction. Finally, regarding "service counter

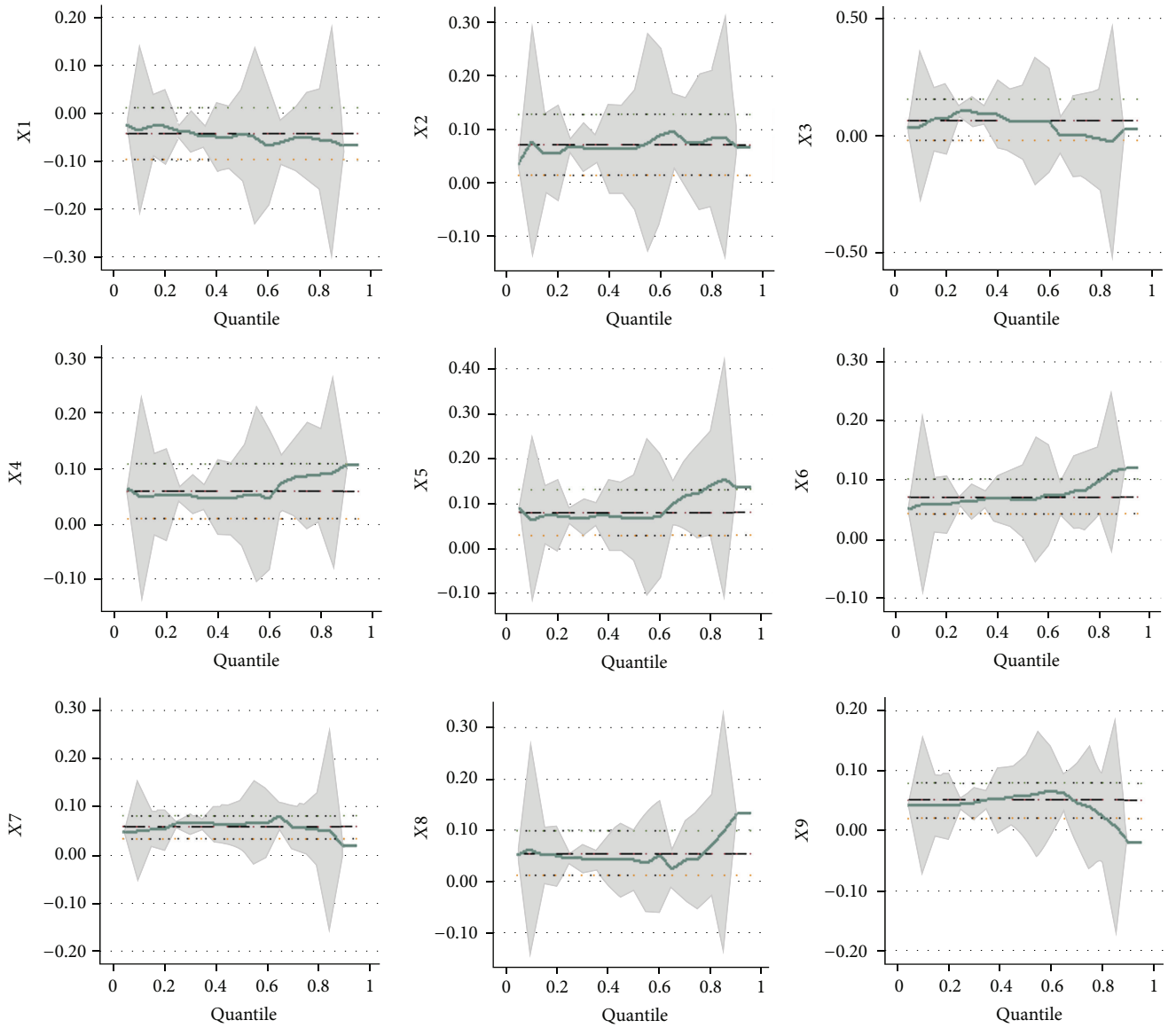



Figure 3: Satisfaction performance results of all 36 banks.

design" ($X9$), its coefficients are different from zero for the 0.25 and 0.5 quantile regression at 1% significance level. Therefore, customers with low or medium levels of service satisfaction care more about the "service counter design" and to improve the design of the service counter may further promote the service satisfaction of the customers with low to medium levels of service satisfaction. The comprehensive analysis of the effects of different variables in this study could help a bank manager to promote the service satisfaction of his customers with different levels of service satisfaction.

*3.4. Performance Comparison of Different Quantile Regression Models.* Finally, by using the grey relational grade as the dependent variables ($Y$) and the nine bank service satisfaction-related variables as the independent variables ($X1$–$X9$), this study built seven forecasting models, including Q15, Q25, Q35, Q50, Q65, Q75, and Q85. 5-fold cross-validation is used to find the values of performance measures of the seven models. To perform 5-fold cross-validation, thirty-six samples of the dataset are divided into five groups. Four groups of them are used to build a model and the rest is used as a test dataset to calculate the values of different performance measures. This procedure is repeated for five times each with a different group of the dataset as the test dataset. The five values of a performance measure are averaged to render the reported value of the performance measure. The performance measures for this study include RMSE, RTIC, and CE [12]. The equations for these measures are as follows.

The equation for RMSE (Root Mean Squared Error) is

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{N} \left(x_t - \widehat{x_t}\right)^2}{N}}. \tag{10}$$

Table 6: Performance of the seven quantile regression models.

| Index | Q15 | Q25 | Q35 | Q50 | Q65 | Q75 | Q85 |
|---|---|---|---|---|---|---|---|
| RMSE | 0.04227 | 0.03610 | 0.12830 | 0.03494 | 0.03906 | 0.04014 | 0.08812 |
| RTIC | 0.00519 | 0.00372 | 0.03859 | 0.00338 | 0.00412 | 0.00434 | 0.02032 |
| CE | 0.74438 | 0.78248 | 0.12402 | 0.84030 | 0.83163 | 0.82688 | 0.18862 |

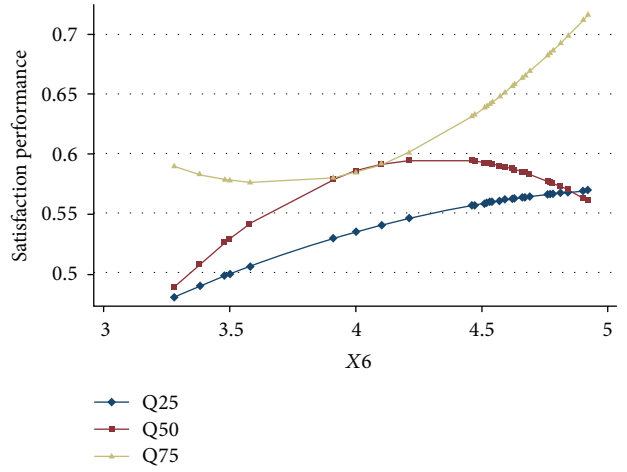



Figure 4: Regressed curves for different quantiles in terms of $X6$.

The equation for RTIC (Revision Theil Inequality Coefficient) is

$$\mathrm{RTIC} = \sqrt{\frac{\sum_{t=1}^{N} \left(x_t - \widehat{x}_t\right)^2}{\sum_{t=1}^{N} x_t{}^2}}. \tag{11}$$

The equation for CE (coefficient of efficiency) is

$$\mathrm{CE} = 1 - \frac{\sum_{t=1}^{N} \left(x_t - \widehat{x}_t\right)^2}{\sum_{t=1}^{N} \left(x_t - \overline{x}_t\right)^2}. \tag{12}$$

Note that a model with small values of RMSE and RITC performs better than a model with large values of RMSE and RITC. For CE, a good forecasting model has a CE value close to 1. The performance of the seven quantile regression models is shown in Table 6.

Table 6 shows that the median regression model has the smallest RMSE and RTIC values and the largest CE value. Therefore, the median regression model has the best prediction accuracy among all the seven quantile regression models.

## 4. Conclusion and Suggestion

The major contribution of this paper is to explore bank service satisfaction performance based on the 2013 service satisfaction survey data of 36 public and private banks in Taiwan. This paper proposes to use the grey relational analysis to gauge the service satisfaction of a customer based on nine question items in a questionnaire. With the grey relational analysis, we have found some variables that contribute more on the service satisfaction of customers than the other variables. We also ranked the banks according to their grey relational grades of satisfaction levels. Furthermore, the quantile regression analysis was used to further explore the determinant factors of service satisfaction for the customers at different quantiles of service satisfactions. With regressions on different quantiles, the manager of a bank can find the factors that are more concerned by their customers at a specific quantile of service satisfaction. As a result, the manager can formulate different policies to promote the service satisfaction of customers at different quantiles of service satisfaction. Finally, this study examined the performance of different regression models. The experimental result showed that, among the seven quantile regression models, the median regression model has the best performance in terms of RMSE, RTIC, and CE performance measures.

## Competing Interests

The authors declare that they have no competing interests.

## References

[1] P. Gerrard and B. Cunningham, "Bank service quality: a comparison between a publicly quoted bank and a government bank in Singapore," *Journal of Financial Services Marketing*, vol. 6, no. 1, pp. 50–66, 2001.

[2] C. T. Ennew and M. R. Binks, "The impact of service quality and service characteristics on customer retention: small businesses and their banks in the UK," *British Journal of Management*, vol. 7, no. 3, pp. 219–230, 1996.

[3] J. Deng, "The control problems of grey system," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.

[4] N. Tosun, "Determination of optimum parameters for multi-performance characteristics in drilling by using grey relational analysis," *International Journal of Advanced Manufacturing Technology*, vol. 28, no. 5-6, pp. 450–455, 2006.

[5] G. Nagpal, M. Uddin, and A. Kaur, "A hybrid technique using grey relational analysis and regression for software effort estimation using feature selection," *International Journal of Soft Computing and Engineering*, vol. 1, no. 6, pp. 20–27, 2012.

[6] H. Hasani, S. A. Tabatabaei, and G. Amiri, "Grey relational analysis to determine the optimum process parameters for open-end spinning yarns," *Journal of Engineered Fibers and Fabrics*, vol. 7, no. 2, pp. 81–86, 2012.

[7] R. Koenker and J. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.

[8] L. Meligkotsidou, I. D. Vrontos, and S. D. Vrontos, "Quantile regression analysis of hedge fund strategies," *Journal of Empirical Finance*, vol. 16, no. 2, pp. 264–279, 2009.

[9] R. Alhamzawi and K. Yu, "Variable selection in quantile regression via Gibbs sampling," *Journal of Applied Statistics*, vol. 39, no. 4, pp. 799–813, 2012.

[10] M. Harding and C. Lamarche, "A quantile regression approach for estimating panel data models using instrumental variables," *Economics Letters*, vol. 104, no. 3, pp. 133–135, 2009.

[11] K. L. Wen, S. K. Chang-Chien, C. K. Yeh, C. W. Wang, and H. S. Lin, *Apply MATLAB in Grey System Theory*, Chuan Hwa Book Co., 2006.

[12] W.-T. Pan, "Mixed modified fruit fly optimization algorithm with general regression neural network to build oil and gold prices forecasting model," *Kybernetes*, vol. 43, no. 7, pp. 1053–1063, 2014.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 9480769, 10 pages
http://dx.doi.org/10.1155/2016/9480769

*Research Article*

# A Robust Service Selection Method Based on Uncertain QoS

**Yanping Chen,[1] Lu Jiang,[2] Jianke Zhang,[3] and Xiaoxiao Dong[1]**

[1]*School of Computer Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China*
[2]*ZTE Corporation, Xi'an 710114, China*
[3]*School of Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China*

Correspondence should be addressed to Yanping Chen; chenyp@xupt.edu.cn

Received 14 October 2015; Revised 11 January 2016; Accepted 27 January 2016

Academic Editor: Alireza Amirteimoori

Copyright © 2016 Yanping Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nowadays, the number of Web services on the Internet is quickly increasing. Meanwhile, different service providers offer numerous services with the similar functions. Quality of Service (QoS) has become an important factor used to select the most appropriate service for users. The most prominent QoS-based service selection models only take the certain attributes into account, which is an ideal assumption. In the real world, there are a large number of uncertain factors. In particular, at the runtime, QoS may become very poor or unacceptable. In order to solve the problem, a global service selection model based on uncertain QoS was proposed, including the corresponding normalization and aggregation functions, and then a robust optimization model adopted to transform the model. Experiment results show that the proposed method can effectively select services with high robustness and optimality.

## 1. Introduction

With the wide use of the Web service in various fields, more and more services have been developed rapidly. But single service can only implement single function; service composition becomes the primary measure to provide user with a high-quality and personalized complex service. And the Web service selection is the foundation and precondition of service composition. With the number of Web services increasing gradually, a key issue arises immediately; there are many services providing same functions but differing in the Quality of Service (QoS). Therefore, in the process of service composition, we need to select service, not only depending on the functions but also according to the user's QoS requirement.

Web service selection based on QoS is a combinatorial optimization problem. Traditional QoS-based Web service selection algorithms mainly use linear programming [1] and Greedy Algorithm [2]. Currently more popular service selection algorithms are mainly using smart algorithms, such as Genetic Algorithms [3], Particle Swarm Optimization [4], and Ant Colony Optimization Algorithm [5]. These algorithms are often designed for certain QoS attributes and assumed that, after the service provider has provided services, the services would run as the described state without change. But in fact, this ideal situation does not exist. Web service is provided as a distributed, loosely coupling API on the Web, so QoS is uncertain factor; for example, the response time would fluctuate with the change of the network. It will fluctuate the QoS of the composed services which are solved by the deterministic algorithms at the running time and cannot satisfy the QoS requirements of users or is unusable. Therefore, when dealing with the service selection problem, the uncertain factors must be considered.

Uncertainty is the most important factor that the requirements of users cannot be satisfied. We consider that the uncertainty of the service composition is mainly caused by the following reasons:

(1) QoS information is incomplete. The Web service description emphasizes more on functional information, so the description of QoS attributes is not accurate enough.

(2) There is no unified QoS metric of Web services, and only one way can gain service QoS value from service providers. Due to the fuzziness of the natural language, there may be divergence between the service providers and the users even about the same attributes. For example, there is one attribute called

throughput, one definition is the number of requests for a service successfully executed, and the other is the average rate of successful message size delivery over a communication channel per second.

(3) Internet environment cannot be repeated. When services are used on the Internet, some attributes of QoS of the Web services are influenced by the running state of the Internet, such as security and response time. The real situation of service cannot be described by the QoS obtained from once request.

Because of the existing uncertainties, the composite service could hardly be accomplished in the dynamic execution process. These uncertainties also affect the individual requirements of users. When the QoS deviation occurs, we must replan and compose services; otherwise it would cause a lot of problems. But the replanning will lead to the waste of computing resources, so this paper studies how to ensure that service composition is feasible with the uncertain conditions, thus avoiding the replanning of composited service. For the uncertainty of the QoS properties, this paper combined the uncertain optimization theory, described the uncertainty by using the "interval," modeled a service composition model based on workflow (order, circulation, choice, and parallel), proposed the service selection model under the conditions with uncertain QoS, put forward the corresponding normalization and aggregation function, and ensured that when the change occurred in the part of QoS attributes, the obtained solution can adapt to the most cases and close to the optimal.

The rest of the sections are organized as follows. Section 2 describes related works. Section 3 details the robust optimization method used in service selection model based on the uncertain QoS. Section 4 presents the experiment and analysis result. Finally, Section 5 draws some conclusions.

## 2. Related Works

In Zeng et al.'s paper [1] QoS was introduced into Web service at the first time; the QoS of Web service was used to distinguish the Web services with similar functions. From then on, service selection based on QoS plays a more and more important role in the Web service composition area. QoS-based service selection methods are mainly divided into two directions: semantic QoS [6] and QoS value. According to the different QoS requirements, the Web service selection strategy can be divided into two kinds, namely, the local optimal strategy and the global optimal strategy. Local optimal strategy selects component services for a single service function, respectively, compares the candidate atomic service set, sorts QoS values for each candidate atomic service, and then selects the best one. The main algorithm is the Greedy Algorithm [2]. Global optimal strategy aims to make the global QoS of the composite services to satisfy the given constraints or achieve the desired goals; the main methods are Genetic Algorithm [3], PSO [4], and Ant Colony Algorithm [5].

In recent years, the researchers have also addressed some studies for service selection based on uncertain QoS. Zeng et al. [1] proposed QoS-based service selection; at the same time they proposed that uncertain response time was introduced into the service selection model. They assumed the response time of service subjected to normal distribution and introduced the variance of the normal distribution into the constraints. Alférez and Pelechano [7] leveraged models at runtime to guide the dynamic evolution of context-aware Web service compositions to deal with unexpected events in the open world. In order to manage uncertainty, a model that abstracts the Web service composition self-evolves to preserve requirements. Benouaret et al. [8] represented each QoS attribute of a Web service using a possibility distribution and introduced two skyline extensions on uncertain QoS called pos-dominant skyline and nec-dominant skyline. Wu et al. [9] proposed an "on-demand" strategy for QoS-aware service composition to replace the traditional "best-effort" strategy. Mostafa and Zhang [10] presented two approaches based on the Multiobjective Reinforcement Learning (MORL) to address multiobjective service composition in uncertain and dynamic environments. Wagner et al. [11] developed a new QoS model that helps to predict the resulting QoS of a workflow by considering service failures during the initial selection phase. For each service, a set of possible backup services and the corresponding QoS are computed beforehand.

In recent years, it has been recognized that dealing with uncertain data is a major challenge in optimization. As can be seen from the above literatures, Web service QoS models with uncertainty dispose uncertain data in two ways; one would assume that the data is subject to one kind of distributions to compute the uncertainty of QoS and the other solves the problem by using intelligent algorithms. But the problem of the former method is that the assumption may be not reasonable. In the real world, we could not get the exact distribution of the data. And for the latter method, different intelligent algorithms have different problem, like premature convergence for Genetic Algorithm. And for all intelligent algorithms, they always solve different solutions in different times, even if they use the same data. We need an approach to select service considering all possible situations, but it does not take whether the uncertain data is subject to any distributions into account.

The researches of the traditional optimization problems are mostly about deterministic problems. When solving uncertain optimization problems, researchers mainly focus on the following ways: stochastic programming, fuzzy programming, interval programming, and robust optimization. When solving the uncertain problems, the stochastic programming and fuzzy programming need precise distribution of the parameters, and the precise distribution is difficult to obtain in the real world. When solving these two kinds of programming problems, the commonly used method is the Monte Carlo method, but this method shows slow convergence and low efficiency. Facing more complex constraints, they are only to use approximate method. Interval programming and robust optimization are all converting the uncertain optimization to determine problem for solving. The difference is that the robust optimization considers the worst case; the result may be conservative. And the interval programming eventually provides decision-makers with a range of optimal value; they will have a better space to select [12]. When converting the interval programming, upper and lower

bounds of uncertainty goals and constraints will become a nested optimization model, so the problem is always a multiobjective optimization problem and difficult to solve.

In the 1973s, Soyster [13] proposed an approach to address data uncertainty named robust optimization. Robust optimization means finding a solution which can deal with all possible realizations of the uncertain data. The first robust optimization method proposed by Soyster considers the maximum uncertainty. That means only considering the worst case, the uncertain programming is converted into the corresponding deterministic model to obtain robust optimal solution. The model only considers the worst case and leads to the very conservative solutions. Conservatism in robust optimization means that the solution ensures a high robustness level but is "far away" from optimality. In other words, a robust solution is not optimal for all realizations of the uncertain data but performs well even for the worst case. Later researches focus on how to keep the optimality and robustness at the same time. At the end of the last century, Ben-Tal and Nemirovski [14] gave an approach that converted the linear programming to the robust counterpart problem with the ellipsoid uncertain set, but the converted problem is a Second-Order Cone Programming (SOCP) and difficult to calculate. In 2003, Bertsimas and Sim [15, 16] proposed a new robust approach which obtained the solution considering optimality and robustness and analyzed the probability bounds of constraint violations. Existing researches based on the two approaches used the robust optimization in the field of supply chain planning, network flow, and logistics programming.

In this paper, we propose using robust optimization method to select service. The difference between the existing study and this method is that this method considers all possible situations during the service as invoking. And it is solved by mathematical methods, which means that if the data is constant, the results are constant.

## 3. Service Selection Method Based on Robust Optimization

*3.1. QoS of the Candidate Services.* Since Zeng et al. [1] combined Web service with QoS together, there have been a large number of QoS attribute definitions which describe nonfunctional properties of Web services. In many literatures, QoS attributes are using the determinate data to make it clear, but some QoS attributes should be described as uncertain interval data, such as response time. This paper uses the following three kinds of common QoS attributes as a basis for service selection. Defining that response time and throughput are described by the interval data and the reliability is real data. We consider that the value of QoS is not subject to any distribution.

*Definition 1* (response time). Response time measures the expected delay in seconds from a request sent to the results which are received [1]. It is computed using the expression as follows:

$$Q_{\text{rt}} = T_{\text{tran}} + T_{\text{proc}}, \tag{1}$$

where $T_{\text{tran}}$ is the transmission time of request and response in the network. $T_{\text{proc}}$ is the time of processing the request.

In the past literatures, the response time is often using the minimum value or average value of the response time in a period of time, as a certain value to calculate. But the nature of the response time should be uncertain. Due to the different network status, the same service will show different response times. Using traditional QoS models always masks the uncertainty of response time, ignoring the unstable factors. Without considering the uncertainty of response time, the selected service would take a bad response time. For example, the minimum and average response time of service A, respectively, are 0.1 s and 0.2 s and of service B are 0.15 s and 0.22 s. However, the intervals of response time of services A and B, respectively, are [0.1, 1.0] and [0.15, 0.5]. From the two intervals, it can be seen that the response time of service B is more stable and robust. Using the traditional service selection method, service A is better than service B. When we use the two services, the response time of service A may be 0.9 s and that of service B may be 0.3. The purpose of this paper is to choose a more robust service. We use interval $[Q_{\text{rt}}, Q_{\text{rt}}+\Delta Q_{\text{rt}}]$ to indicate the uncertainty of response time.

*Definition 2* (throughput). For the throughput, one definition is the number of requests for a service successfully executed, and the other is the average rate of successful message size delivery over a communication channel per second. In this paper, we define it as in the second definition. It is also uncertain. So we describe the throughput like response time with $[Q_{\text{tp}}, Q_{\text{tp}} + \Delta Q_{\text{tp}}]$.

*Definition 3* (reliability). Reliability of service is the probability of a request that is correctly responded to within the maximum number of total requests. It is computed using the expression as follows:

$$Q_{\text{re}} = \frac{N_{\text{re}}}{n}, \tag{2}$$

where $N_{\text{re}}$ means the number of successful requests in $n$ invocations and $n$ is the total number of service executions. In this paper, the reliability is derived from the service attribute dataset and is a certain value.

*3.2. Service Composition Based on Workflow.* Most of single Web service is usually providing only a single way to implement a single function, but the composite services can take advantage of many small particle size distributions of the atomic services on the network and combine them to be able to implement complex functions, loosely coupled composite services.

As shown in Figure 1, business logic of each kind of complex application is fixed, which has been divided into service classes by domain experts, and then the service class is advertised. After that, service providers could publish their Web service into registry and point out which certain service class it belongs to. When users' requirement is coming (meaning that user selects a certain business logic), then the proper services should be selected according to users' QoS. In order to solve the problem, a global service selection model based

Table 1: Aggregate formulas of QoS.

| | Response time | Throughput | Reliability |
|---|---|---|---|
| Sequence | $\left[\sum_{i=1}^{n} Q_{\mathrm{rt}i}, \sum_{i=1}^{n} (Q_{\mathrm{rt}i} + \Delta Q_{\mathrm{rt}i})\right]$ | $\left[\min_i Q_{\mathrm{tp}i}, \min_i (Q_{\mathrm{tp}i} + \Delta Q_{\mathrm{tp}i})\right]$ | $\prod_{i=1}^{n} Q_{\mathrm{re}i}$ |
| Selection | $\left[\sum_{i=1}^{n} c_i Q_{\mathrm{rt}i}, \sum_{i=1}^{n} c_i (Q_{\mathrm{rt}i} + \Delta Q_{\mathrm{rt}i})\right]$ | $\left[\min_i c_i Q_{\mathrm{tp}i}, \min_i c_i (Q_{\mathrm{tp}i} + \Delta Q_{\mathrm{tp}i})\right]$ | $\sum_{i=1}^{n} c_i Q_{\mathrm{re}i}$ |
| Loops | $[kQ_{\mathrm{rt}}, k(Q_{\mathrm{rt}} + \Delta Q_{\mathrm{rt}})]$ | $[Q_{\mathrm{tp}}, Q_{\mathrm{tp}} + \Delta Q_{\mathrm{tp}}]$ | $Q_{\mathrm{re}}^{k}$ |
| Parallel | $\left[\max_i Q_{\mathrm{rt}i}, \max_i (Q_{\mathrm{rt}i} + \Delta Q_{\mathrm{rt}i})\right]$ | $\left[\min_i Q_{\mathrm{tp}i}, \min_i (Q_{\mathrm{tp}i} + \Delta Q_{\mathrm{tp}i})\right]$ | $\prod_{i=1}^{n} Q_{\mathrm{re}i}$ |

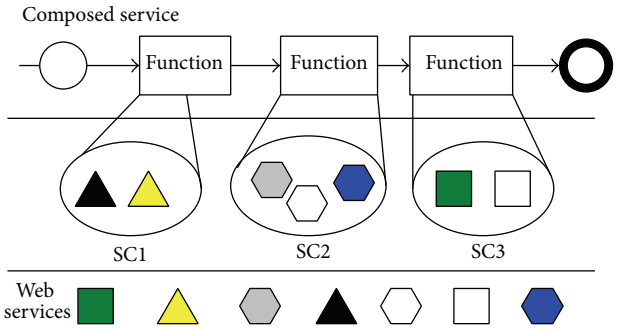



Figure 1: QoS-based Web services composition problem.

on uncertain QoS was proposed, including the corresponding normalization and aggregation functions, and then a robust optimization model is adopted to transform the model.

Techniques of composite services are various. Mainly methods include modeling based on workflow, modeling based on Petri nets, and semantic modeling. This paper focuses on how to select the atomic service to accomplish the function, regardless of the composite modeling techniques; how to choose the specific combination service workflow model is also the key problem.

Workflow model has four basic atomic structures, namely, sequence, selection, loop, and parallel. The flow of composite services can be combined or nested by the basic structures. If we can find a corresponding calculation method of each QoS for each basic structure, we can give the global QoS of composite services. In the traditional QoS-driven service selection algorithms when aggregating the QoS of the services, the algorithms only consider the determined QoS, so that the aggregate formula is simple. But there is uncertainty in the QoS, and we described it by interval numbers, so the traditional function cannot be used for calculating global QoS. In order to solve this problem, we combined the interval number algorithm with the characteristics of response time and throughput and proposed the aggregation formulas of three QoS properties corresponding to each basic structure in Table 1.

*3.3. Normalization.* Different QoS attributes always have different dimensions. Some value of QoS attributes is the bigger the better, such as throughput and reliability, and this kind of attributes is called benefit type attribute; some is the smaller the better, such as response time, called cost type attribute. In this case, integrating global QoS of composite service could result in additional cost. In order to easily calculate and compare, data in the decision matrix needs to be normalized into the same dimension. The service selection algorithm based on real number generally used 0-1 normalization method; the attributes' values are converted to interval [0, 1] by linear calculations. However, in this paper, response time and throughput are interval data and cannot use the method of real number to normalize. The researches about normalization methods of interval data are not too many and mainly focused on two ways. One is to convert the interval data into a real number. The other is to use some kinds of normalization functions to directly convert the interval number; the result is still interval number. While the former method will lose the uncertain information which is what this paper focused on, we use the latter method.

Common methods convert interval numbers into standardized interval number including the method based on interval number operation method, vector-based normalization method, method based on a linear scale transformation, and method based on the range transformation [17]. These three former methods combine the normalization method of real number with interval number operation method. However, the data obtained by these methods always is out of [0, 1], and these methods are based on nonlinear transformation. There has been a huge impact on the decision results. So we use the method based on the range transformation to normalize response time and throughput. Because the value of the reliability is taken in the range of [0, 1], it does not need to be normalized.

There are $n$ services realizing one function. Assuming that the QoS attribute of each service takes values in the interval $[Q_i, Q_i + \Delta Q_i]$, the QoS after normalization is written as $[Q_i', Q_i' + \Delta Q_i']$. In order to easily calculate, we need to normalize the attributes data to be cost type data. The normalization formula is established as follows:

Cost type:

$$Q_i' = \frac{Q_i - \min_i Q_i}{\max_i (Q_i + \Delta Q_i) - \min_i Q_i},$$
$$Q_i' + \Delta Q_i' = \frac{(Q_i + \Delta Q_i) - \min_i Q_i}{\max_i (Q_i + \Delta Q_i) - \min_i Q_i}. \quad (3)$$

Benefit type:

$$Q_i' = \frac{\max_i (Q_i + \Delta Q_i) - Q_i}{\max_i (Q_i + \Delta Q_i) - \min_i Q_i},$$
$$Q_i' + \Delta Q_i' = \frac{\max_i (Q_i + \Delta Q_i) - (Q_i + \Delta Q_i)}{\max_i (Q_i + \Delta Q_i) - \min_i Q_i}. \quad (4)$$

*3.4. Service Selection Model Based on Uncertain QoS.* At present, many scholars transform the service selection problem as mathematical problems from different ways and designed the corresponding selection model. The transformed mathematical problems include linear or integer programming problem, multiple choice knapsack problem, multiple attribute decision-making problems, single objective combinatorial optimization problems with QoS constraints, and multiobjective combinatorial optimization problem with QoS constrained. But these problems always do not consider the uncertainty of QoS, so we propose a service selection model considering the uncertainty.

According to the global service selection strategy, we propose a service selection model based on uncertain QoS. Assuming that there are $p$ tasks in the workflow, each task has $q$ candidate atomic services, so that the number of all possible composite services is $q^p$. We propose a method based on robust optimization (RO), which can select one service with high robustness level from $q^p$ composite services. In order to implement our approach, firstly we need to formulate a linear programming model.

There are three inputs in linear programming problem: variables, objective function, and constraint conditions, in which the objective function and constraints must be linear. Linear programming is to maximize or minimize the value of objective function by adjusting the value of the variables. The output of linear programming is maximum (or minimum) value of objective function. We assume that the variable is $x_{ij}$, which means whether the service $s_{ij}$ is selected to execute plan or not. If the service $s_{ij}$ is selected, then the value of the variable $x_{ij}$ is 1; otherwise it is 0. The objective function of the model is the following formula:

$$\min Q_{\text{rt}} \cdot \omega_1 + Q_{\text{tp}} \cdot \omega_2 + Q_{\text{re}} \cdot \omega_3, \quad (5)$$

where we transformed three QoS attributes into cost type attributes and $Q_{\text{rt}}$, $Q_{\text{tp}}$, and $Q_{\text{re}}$ are global QoS. $\omega_i$ is the weight of the QoS, and $\sum_{i=1}^{n} \omega_i = 1$, $\omega_i \in [0, 1]$.

We used $Q_{\text{rt}}(x_{ij})$ to express the response time of service $s_{ij}$ for task $t_i$, so we have the following constraint:

$$Q_{\text{rt}} = \sum_{i=1}^{n} \sum_{j=1}^{m_i} Q_{\text{rt}}\left(x_{ij}\right) \cdot x_{ij}. \quad (6)$$

Assume that there are $n$ workflow tasks and there is a set of $m_i$ services that can be provided to the user for each task, but there is only one service selected in each set, so $x_{ij}$ ($x_{ij} = 0, 1$) indicates whether the service is selected. $x_{ij}$ has the following constraints:

$$\sum_{j=1}^{m_i} x_{ij} = 1. \quad (7)$$

For example, there are 100 candidate Web services that can execute the task $t_i$; however, only one of them will be selected to the task, so it is $\sum_{j=1}^{100} x_{ij} = 1$.

We used $Q_{\text{tp}}(x_{ij})$ to express the throughput of service $s_{ij}$. From Table 1, we can see that the aggregate formula of throughput is looking for the minimum throughput in the process as the global throughput. But before the calculation of QoS value, we need to normalize it; the throughput will be converted from the benefit type to the cost type, so the global throughput is the maximum value. The constraint to the throughput is as follows:

$$Q_{\text{tp}} = \max Q_{\text{tp}}\left(x_{ij}\right) \cdot x_{ij}. \quad (8)$$

$Q_{\text{re}}(x_{ij})$ means the reliability of service $s_{ij}$. Like throughput, the reliability also needs to be converted, so the global reliability is as follows:

$$Q_{\text{re}} = 1 - \prod_{i=1}^{n} \prod_{j=1}^{m_i} Q_{\text{re}}\left(x_{ij}\right) \cdot x_{ij}. \quad (9)$$

However, the constraint of attribute can be introduced into the linear programming problem which is not limited to the defined one in this section. If the aggregate function could be given, it also can be added to the model.

We can propose the following model:

$$\begin{aligned} \min \quad & \left\{ \omega_1 \sum_{i=1}^{n} \sum_{j=1}^{m_i} Q_{\text{rt}}\left(x_{ij}\right) \cdot x_{ij} + \omega_2 \cdot \max Q_{\text{tp}}\left(x_{ij}\right) \cdot x_{ij} + \omega_3 \left( 1 - \prod_{i=1}^{n} \prod_{j=1}^{m_i} Q_{\text{re}}\left(x_{ij}\right) \cdot x_{ij} \right) \right\} \\ \text{s.t.} \quad & \sum_{i=1}^{3} \omega_i = 1 \\ & \sum_{i=1}^{n} x_i = 1 \\ & x_{ij} \in \{0, 1\}, \end{aligned} \quad (10)$$

where there exists uncertainty in two QoS attributes: response time and throughput. For each entry $Q_{rt}(x_{ij})$ is independent and bounded random variable and takes values in $\lfloor Q_{rt}(x_{ij}), Q_{rt}(x_{ij}) + \Delta Q_{rt}(x_{ij}) \rfloor$. $\Delta Q_{rt}(x_{ij})$ means the deviation of the coefficient $Q_{rt}(x_{ij})$. It is the same as throughput, in that, for each entry, $Q_{tp}(x_{ij})$ is independent and bounded random variable and takes values in $\lfloor Q_{tp}(x_{ij}), Q_{tp}(x_{ij}) + \Delta Q_{tp}(x_{ij}) \rfloor$ and $\Delta Q_{tp}(x_{ij})$ means the deviation of the coefficient $Q_{tp}(x_{ij})$.

Obviously, model (10) cannot be solved with the traditional optimization method or intelligent algorithm. In order to be able to consider the uncertainty of QoS in the solution, we used the robust optimization model proposed by Bertsimas to deal with the uncertainty. Due to the uncertainty that only appears in the objective function, we introduce a parameter $\Gamma$, which takes value in the interval $[0, k]$, where $k$ is the number of selected services in the model. For one flow, there will be only one composite service selected, so the range of $\Gamma$ should be $[0, 1]$. So we propose the following service selection model:

$$\begin{aligned}
\min \quad & \left\{ \omega_1 \sum_{i=1}^{n} \sum_{j=1}^{m_i} Q_{rt}\left(x_{ij}\right) \cdot x_{ij} + \omega_2 \cdot \max Q_{tp}\left(x_{ij}\right) \cdot x_{ij} + \omega_3 \left( 1 - \prod_{i=1}^{n} \prod_{j=1}^{m_i} Q_{re}\left(x_{ij}\right) \cdot x_{ij} \right) + \max_{\{S|S\subseteq A, |S|\le \Gamma\}} \left\{ \omega_1 \sum_{(i,j)\in S} \Delta Q_{rt}\left(x_{ij}\right) \cdot x_{ij} + \omega_2 \max \Delta Q_{tp}\left(x_{ij}\right) \cdot x_{ij} \right\} \right\} \\
\text{s.t.} \quad & \sum_{i=1}^{n} \omega_i = 1, \quad \omega_i \in [0,1] \\
& \sum_{j=1}^{m_i} x_{ij} = 1 \\
& x_{ij} \in \{0,1\}.
\end{aligned} \tag{11}$$

$\Gamma$ in objective function (11) is to adjust the robustness of the solution against the level of conservatism of the solution. $\Gamma = 0$: the model will completely ignore all uncertain factors, so that the solution is the same as the solution obtained by solving with minimum values $Q_{rt}(x_{ij})$ and $Q_{tp}(x_{ij})$; this solution is called the optimal solution. $\Gamma = 1$: the model will consider all uncertainties, but the solution is too conservative; this solution is called robust solution. If the value of $\Gamma$ is more close to 1, we say the corresponding solution has a higher level of robustness. So robust solution has the highest level of robustness. In general, a higher value of $\Gamma$ increases the level of robustness at the expense of higher nominal cost.

## 4. Experiment Analyses

In order to verify the effectiveness of the proposed method, we developed Web service testing and selection management system in the early study. In the literature, the researchers usually solved deterministic model. The value of QoS they used is from one test result or some random numbers produced in a certain range. This paper is to study and calculate uncertain data, once test result or random numbers cannot describe uncertainty of the service used in the real world. We use the response time and throughput datasets from Zheng et al.'s paper. In the literature [18], about 1.9 million pieces of data were provided by 339 users calling the 5825 atomic services. In this dataset, we found that when users invoked the service, the request may succeed or not, and the values of response time and throughput are always varied.

In this experiment, the workflow is a simple sequence flow with four tasks; each task would invoke in order. Each task has 10 candidate services, so the number of all possible composite services is $10^4$. Set three QoS attributes that have the same weight and calculate the uncertainty QoS interval of composite service through the front of the aggregation function. Table 2 is the QoS value of 40 candidate services.

In Table 2, the data comes from Zheng et al.'s dataset. We find out the maximum and minimum value as the upper and lower bounds of the response time and throughput intervals. In the dataset, −1 means the request is failed, so we calculate the reliability by failure number and total number.

Since $\Gamma > 1$, the obtained composite service is the same as when $\Gamma = 1$, so in the experiment, $\Gamma$ takes value in the interval $[0, 1]$. If $\Gamma$ is an integer, it can only take 0 or 1; in both cases, obtained composite services are, respectively, the optimal solution and the robust solution mentioned in Section 3.2. To validate the method, $\Gamma$ will be taking decimal to analysis.

When solving (11), we have two ways. The first way is solving the robust counterpart of (11), and the second is using one algorithm. The solving algorithm is as follows:

(1) For $l = 1, \ldots, n+1$, solve the $n+1$ nominal problems

$$G^l = \Gamma \cdot d_l + \min_{x\in X} \left( c' \cdot x + \sum_{j=1}^{l} \left( d_j - d_l \right) \cdot x_j \right) \tag{12}$$

and let $l^*$ be an optimal solution of the corresponding problem.

(2) Let $l^* = \arg\min_{l=1,\ldots,n+1} G^l$.

(3) $Z^* = G^{l^*}$, $\mathbf{x}^* = \mathbf{x}^{l^*}$.

Although there are a lot of methods and tools for solving integer programming, it would spend too much time, and the algorithm can quickly find a solution in few seconds. So this experiment uses the above algorithm to solve it.

The experiment is divided into two parts. The first part is to adjust the value of $\Gamma$ and observe parameter effect on

Table 2: The QoS value of 40 candidate services.

| Number | $Q_{rt}$ | $Q_{tp}$ | $Q_{re}$ |
|---|---|---|---|
| | Task 1 | | |
| 1 | [0.432, 16.523] | [0.02, 4.629] | 97.94% |
| 2 | [0.078, 4.29] | [0.932, 51.282] | 98.23% |
| 3 | [0.084, 13.835] | [0.433, 71.428] | 98.53% |
| 4 | [0.061, 3.317] | [0.602, 32.786] | 98.53% |
| 5 | [0.071, 3.103] | [0.644, 28.169] | 98.53% |
| 6 | [0.015, 6.582] | [0.455, 200.0] | 99.71% |
| 7 | [0.154, 9.626] | [0.207, 12.987] | 99.71% |
| 8 | [0.22, 9.552] | [11.201, 486.363] | 99.71% |
| 9 | [0.115, 3.726] | [0.536, 17.391] | 98.53% |
| 10 | [0.169, 13.237] | [0.226, 17.751] | 98.23% |
| | Task 2 | | |
| 11 | [0.301, 16.648] | [0.12, 6.644] | 97.94% |
| 12 | [0.444, 15.867] | [0.126, 4.504] | 94.69% |
| 13 | [1.541, 16.807] | [3.093, 33.744] | 98.82% |
| 14 | [0.455, 9.651] | [1.347, 28.571] | 97.94% |
| 15 | [0.191, 19.652] | [0.101, 10.471] | 74.63% |
| 16 | [0.116, 15.877] | [0.06, 17.241] | 98.23% |
| 17 | [0.466, 11.27] | [0.798, 19.313] | 98.23% |
| 18 | [0.502, 11.187] | [0.804, 17.928] | 97.94% |
| 19 | [0.284, 10.87] | [0.275, 10.563] | 100.00% |
| 20 | [0.164, 2.149] | [1.395, 18.292] | 100.00% |
| | Task 3 | | |
| 21 | [0.164, 3.594] | [0.834, 18.292] | 100.00% |
| 22 | [0.325, 9.772] | [0.306, 9.23] | 100.00% |
| 23 | [0.162, 3.512] | [0.854, 18.518] | 100.00% |
| 24 | [0.162, 3.816] | [0.786, 18.518] | 100.00% |
| 25 | [0.345, 7.115] | [0.421, 8.695] | 100.00% |
| 26 | [0.17, 2.066] | [1.452, 17.647] | 100.00% |
| 27 | [0.169, 2.173] | [1.38, 17.751] | 100.00% |
| 28 | [0.17, 2.349] | [1.277, 17.647] | 100.00% |
| 29 | [0.379, 14.363] | [0.208, 7.915] | 100.00% |
| 30 | [0.194, 4.32] | [0.694, 15.463] | 100.00% |
| | Task 4 | | |
| 31 | [0.195, 14.418] | [0.208, 15.384] | 100.00% |
| 32 | [0.217, 13.192] | [0.227, 13.824] | 98.82% |
| 33 | [0.2, 3.272] | [0.916, 15.0] | 98.82% |
| 34 | [0.202, 3.61] | [0.831, 14.851] | 98.82% |
| 35 | [0.383, 19.796] | [0.146, 7.832] | 99.41% |
| 36 | [0.216, 5.397] | [0.555, 13.888] | 99.71% |
| 37 | [0.276, 11.531] | [0.433, 18.115] | 98.53% |
| 38 | [0.496, 16.851] | [0.534, 18.145] | 100.00% |
| 39 | [0.422, 15.991] | [2.814, 106.635] | 100.00% |
| 40 | [0.343, 9.15] | [0.546, 14.577] | 100.00% |

the results by solving the robust optimization method. The second part is to analyze the different results solved by using robust optimization (RO), linear programming (LP), and Linear Programming with Uncertain Response Time [1] (LPURT).

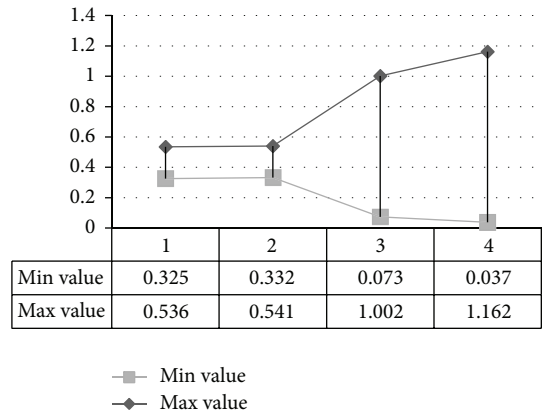



Figure 2: The global QoS range of four composite services.

In order to compare the results of the different methods, this paper will show the global QoS of services which are selected by methods. If the global QoS of the selected composite service is lower, we believe that the optimality of the service is better. If the deviation of service QoS is smaller and the upper QoS is smaller, we believe the robustness of the service is better. In all results of the experiment, the QoS values of all component services are standardized before, except that the total QoS in Table 4 is the result of the standardization and the aggregation.

*4.1. Effect of Adjusting the Value of* $\Gamma$. In the experiment, the value of $\Gamma$ only can be taken in the interval $[0, 1]$, so choose $\Gamma \in \{1, 0.8, 0.6, 0.4, 0.2, 0\}$ to solve the proposed model. The specific results are shown in Table 3.

It can be seen from Table 3 that when $\Gamma$ dropped to 0.4, the atomic services of composite service began to change. In Table 4, we give out the QoS attributes of four composite services.

The global QoS has been normalized and aggregated. When we are calculating the global QoS, first we adjust reliability to cost type attribute (the smaller the better) and then sum the three QoS attributes with weight (weight of three QoS attributes is equal; i.e., $\omega_1 = \omega_2 = \omega_3 = 1/3$).

We named the four combination services from $CS_1$ to $CS_4$. It can be seen from Figure 2 that $\Gamma$ takes value from 1 to 0, and the deviations of global QoS are increasing. If the services are invoked in the real world, we contrast $CS_4$ and $CS_1$; the former may realize a smaller QoS value, but the latter is more stable and more robust.

The experiment proved that we can take a tradeoff between optimality and robustness by adjusting the value of $\Gamma$ with user demand.

*4.2. Contrast RO, LP, and LPURT.* First, respectively, we use three methods for selecting services in the mentioned workflow, obtaining three services called composite services A, B, and C, as shown in Table 5. Among them, when using LP, the response time is the minimum response time of each atomic service, and throughput is the maximum throughput (the best case). As noted above, all of these methods consider

Table 3: Solutions of different values of Γ.

| Values of Γ | 1 | 0.8 | 0.6 | 0.4 | 0.2 | 0 |
|---|---|---|---|---|---|---|
| Result of service selection | [5, 20, 26, 33] | [5, 20, 26, 33] | [5, 20, 26, 33] | [4, 20, 26, 33] | [8, 14, 26, 39] | [8, 13, 24, 39] |

Table 4: The QoS attributes of four composite services.

| Number | Γ | Composite service | Response time | Throughput | Reliability | Global QoS |
|---|---|---|---|---|---|---|
| 1 | 1 | [5, 20, 26, 33] | [0.605, 10.59] | [1.452, 28.169] | 97.30% | [0.325, 0.536] |
| 2 | 0.4 | [4, 20, 26, 33] | [0.595, 10.804] | [1.452, 32.786] | 97.30% | [0.332, 0.541] |
| 3 | 0.2 | [8, 14, 26, 39] | [1.267, 37.26] | [11.201, 486.36] | 97.60% | [0.073, 1.002] |
| 4 | 0 | [8, 13, 24, 39] | [2.345, 46.166] | [11.201, 486.36] | 98.50% | [0.037, 1.162] |

Table 5: The results of three methods.

| Solving method | Robust optimization | Linear programming | Linear programming with uncertain response time |
|---|---|---|---|
| Composite service | A | B | C |
| Atomic services | [5, 20, 26, 33] | [6, 16, 23, 31] | [4, 20, 24, 33] |
| Response time | [0.605, 10.59] | [0.488, 40.389] | [0.587, 12.554] |
| Throughput | [0.644, 15] | [0.06, 15.384] | [0.602, 15] |
| Reliability | 97.32% | 97.91% | 97.32% |

Table 6: The global QoS of three selected composite services.

| Methods | Composite service (CS) | Atomic services | User ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| RO | A | [5, 20, 26, 33] | Response time | 1.51 | 1.517 | 0.956 | 0.92 | 0.989 | 1.361 | 1.582 | 1.001 | 1.469 | 1.004 |
| | | | Throughput | 4.878 | 6.369 | 9.049 | 8.13 | 8.583 | 6.172 | 4.538 | 8.474 | 6.355 | 7.272 |
| LP | B | [6, 16, 23, 31] | Response time | 4.011 | 2.727 | 1.268 | 1.109 | 1.677 | 2.386 | 2.053 | 1.649 | 2.355 | 1.074 |
| | | | Throughput | 1.066 | 1.809 | 5.181 | 4.672 | 2.702 | 1.915 | 2.929 | 2.832 | 2.057 | 5.665 |
| URTLP | C | [4, 20, 24, 33] | Response time | 1.536 | 1.525 | 0.959 | 0.94 | 0.993 | 4.504 | 1.527 | 1.009 | 1.467 | 1.013 |
| | | | Throughput | 4.962 | 6.153 | 9.09 | 8.032 | 8.474 | 0.614 | 4.538 | 8.51 | 6.355 | 7.194 |

determined values. When using LPURT, the response time is assumed to be subject to a normal distribution in [1], while throughput still is the maximum throughput without uncertainty.

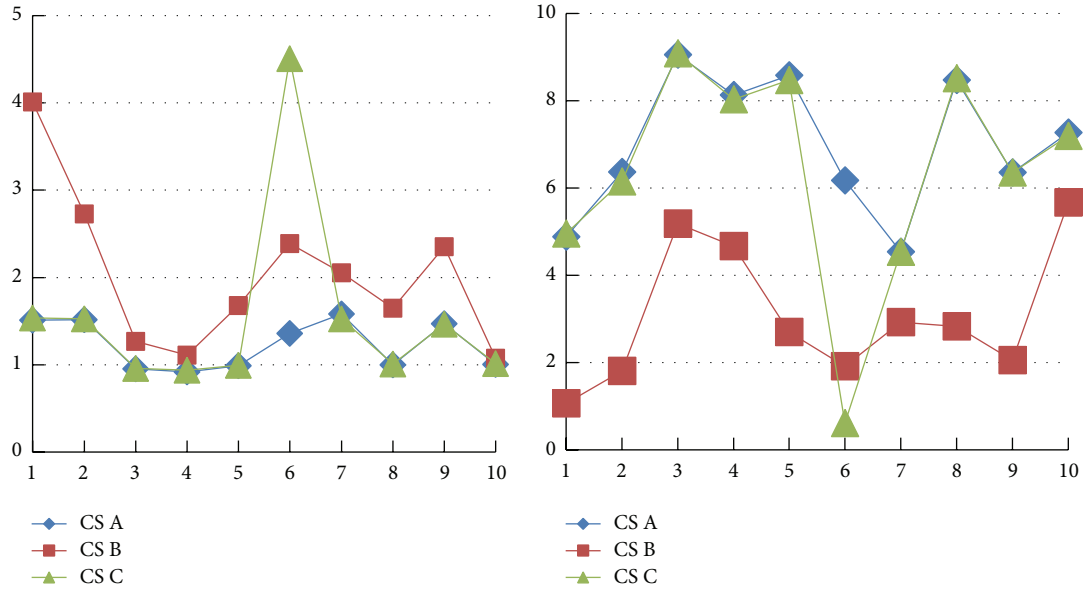
As it is shown in Table 5, there are two different atomic services between composite services A and C. They have the same reliability, but the deviations of response time and throughput of composite service A are smaller. In the best case, composite service B is seen a little better than A and C; however, the deviation is bigger than the other ones.

Then, we select randomly QoS of 40 candidate atomic services invoked by 10 users from dataset and compute global QoS of composite services. Table 6 and Figure 2 show the global QoS of three selected composite services.

We firstly compare composite services A and B. From the data in Figure 3 and Table 6, we can see that response time and throughput of composite service A are obviously better than B. Then compare composite services A and C. From the data, the QoS of composite service A is better than service C, in most cases. From Figure 3, we can see the QoS of the sixth user and response time and throughput of composite service C have an obviously deteriorated case. After checking the QoS of each atomic service, we found that one atomic service of composite service C has a bad case at this user. The reason for this is that LPURT assumed the response time of atomic service subject to normal distribution and introduced the mean and variance to the constraints. While, in the real world, we cannot confirm what the probability distribution of the response time of services is subject to, when we request some services, their quality may be more likely to deteriorate as we have mentioned. Here we could calculate the QoS deviations of the three composite services by 10 users invoked, as shown in Table 7. From Table 7, we can see that the response time and throughput deviations of composite service A are minimal. Influenced by the result of the sixth user, the deviation

Table 7: The QoS fluctuation of combination services.

| | Deviation of response time | Deviation of throughput |
|---|---|---|
| Composite service A | 0.662 | 4.511 |
| Composite service B | 2.937 | 4.599 |
| Composite service C | 3.564 | 8.476 |




(a) Response time (b) Throughput

FIGURE 3: The global QoS of three selected composite services.

of QoS of composite service C is maximum. That means robustness level of service A is higher than B and C.

Therefore, the composite service selected by our approach has optimality and robustness. And we can control the tradeoff between optimality and robustness by adjusting the attribute. The proposed method without considering whether the uncertain QoS is subject to some distributions is more suitable for the real network environment.

## 5. Conclusions

This paper studies the robust optimization used in the Web service selection based on uncertain QoS. On the basis of certain QoS-based Web service selection method, first we considered the uncertainty of QoS, second established uncertain QoS-based service selection model, and then used robust optimization methods to convert and solve the proposed model. In the end, the experiment verifies that the proposed model and method can deal with uncertain data without obvious distribution. And we compared the composite services obtained by our approach and other two methods; the result showed that the service selected by our approach had higher robustness and optimality levels.

Uncertain QoS-based Web service composition has the following advantages: firstly, for the software developer, paying attention to this can improve users' acceptance, enhance product competitiveness, and corporate reputation. Secondly, for users, paying attention to this can improve users' productivity, reduce training and support costs, improve satisfaction and comfort for users' work, and increase construction investment system efficiency and utilization. Thirdly, service selection study under uncertainty will increase the accuracy of the services provided and give full potential of each service and the effect of cooperating with other services. Finally, the composite services with uncertainty focus on providing continuity of services to ensure that when QoS changes in a certain range, the composite service is robust and minimizes the impact of service errors or failure.

Over the past few years, cloud computing has been more and more attractive as a new computing paradigm due to high flexibility for provisioning on-demand computing resources that are used as services through the Internet [19, 20]. It is very necessary that cloud users be vigilant while selecting the service providers present in the cloud [21, 22]. The proposed approach in this paper is valid in cloud service selection.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 61100166) and Xi'an Science and Technology Program (Grant no. CXY1437(8)).

## References

[1] L. Z. Zeng, B. Boualem, D. Marlon, J. Kalagnanam, and Q. Z. Sheng, "Quality driven Web services composition," in *Proceedings of the 12th International Conference on World Wide Web (WWW '03)*, pp. 411–421, Budapest, Hungary, May 2003.

[2] Q. Yu, M. Rege, A. Bouguettaya, B. Medjahed, and M. Ouzzani, "A two-phase framework for quality-aware Web service selection," *Service Oriented Computing and Applications*, vol. 4, no. 2, pp. 63–79, 2010.

[3] C.-W. Zhang, S. Su, and J.-L. Chen, "Genetic algorithm on web services selection supporting QoS," *Chinese Journal of Computer*, vol. 29, no. 7, pp. 1029–1037, 2006 (Chinese).

[4] T. Wen, G.-J. Sheng, Q. Guo, and Y.-Q. Li, "Web service composition based on modified particle swarm optimization," *Chinese Journal of Computers*, vol. 36, no. 5, pp. 1031–1046, 2013 (Chinese).

[5] Y.-M. Xia, B. Cheng, J.-L. Chen, X.-W. Meng, and D. Liu, "Optimizing services composition based on improved ant colony algorithm," *Chinese Journal of Computers*, vol. 35, no. 2, pp. 270–281, 2012 (Chinese).

[6] S.-G. Deng, L.-T. Huang, B. Wu, J.-W. Yin, and G.-X. Li, "QoS optimal automatic composition of semantic Web services," *Chinese Journal of Computers*, vol. 36, no. 5, pp. 1015–1030, 2013 (Chinese).

[7] G. H. Alférez and V. Pelechano, "Facing uncertainty in Web service compositions," in *Proceedings of the IEEE 20th International Conference on Web Services (ICWS '13)*, pp. 219–226, Santa Clara, Calif, USA, July 2013.

[8] K. Benouaret, D. Benslimane, and A. Hadjali, "Selecting skyline web services from uncertain QoS," in *Proceedings of the IEEE 9th International Conference on Services Computing (SCC '12)*, pp. 523–530, IEEE, Honolulu, Hawaii, USA, June 2012.

[9] Q. Wu, Q. Zhu, X. Jian, and F. Ishikawa, "Broker-based SLA-aware composite service provisioning," *Journal of Systems & Software*, vol. 96, no. 4, pp. 194–201, 2014.

[10] A. Mostafa and M. Zhang, "Multi-objective service composition in uncertain environments," *IEEE Transactions on Services Computing*, 2015.

[11] F. Wagner, F. Ishikawa, and S. Honiden, "Robust service compositions with functional and location diversity," *IEEE Transactions on Services Computing*, 2014.

[12] J. K. Zhang, *Optimality and duality for generalized convex uncertain programming [Ph.D. thesis]*, Xidian University, Xi'an, China, 2012, (Chinese).

[13] A. L. Soyster, "Convexprogramming with set-inclusive constraints and applications to inexact linear programming," *Operations Research*, vol. 21, no. 5, pp. 1154–1157, 1973.

[14] A. Ben-Tal and A. Nemirovski, "Robust solutions of uncertain linear programs," *Operations Research Letters*, vol. 25, no. 1, pp. 1–13, 1999.

[15] D. Bertsimas and M. Sim, "Robust discrete optimization and network flows," *Mathematical Programming*, vol. 98, no. 1-3, Ser. B, pp. 49–71, 2003.

[16] D. Bertsimas and M. Sim, "The price of robustness," *Operations Research*, vol. 52, no. 1, pp. 35–53, 2004.

[17] M. L. Hu, C. X. Fan, and K. Q. Shi, "Character analysis of standardization methods of decision matrix with intervals," *Computer Science*, vol. 40, no. 10, pp. 203–207, 2013 (Chinese).

[18] Z. B. Zheng, Y. L. Zhang, and M. R. Lyu, "Distributed QoS evaluation for real-world Web services," in *Proceedings of the IEEE 8th International Conference on Web Services (ICWS '10)*, pp. 83–90, IEEE, Miami, Fla, USA, July 2010.

[19] Y. Zhong, W. Li, W. Guo, L. Gong, and G. Lodewijks, "A method of modeling and service encapsulation on cloud logistics resources," in *Proceedings of the 19th IEEE International Conference on Computer Supported Cooperative Work in Design (CSCWD '15)*, pp. 383–388, IEEE, Calabria, Italy, May 2015.

[20] W. Li, Y. Zhong, X. Wang, and Y. Cao, "Resource virtualization and service selection in cloud logistics," *Journal of Network and Computer Applications*, vol. 36, no. 6, pp. 1696–1704, 2013.

[21] P. Bedi, H. Kaur, and B. Gupta, "Trustworthy service provider selection in cloud computing environment," in *Proceedings of the International Conference on Communication Systems and Network Technologies (CSNT '12)*, pp. 714–719, Rajkot, India, May 2012.

[22] R. Karim, C. Ding, A. Miri, and X. Liu, "End-to-end QoS mapping and aggregation for selecting cloud services," in *Proceedings of the International Conference on Collaboration Technologies and Systems (CTS '14)*, pp. 515–522, IEEE, Minneapolis, Minn, USA, May 2014.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2016, Article ID 1531282, 8 pages
http://dx.doi.org/10.1155/2016/1531282

*Research Article*

# Alternative Trade-Offs in Data Envelopment Analysis: An Application to Hydropower Plants

**Mohsen Mirzaei,[1] Sohrab Kordrostami,[1] Alireza Amirteimoori,[2] and Mehrdad G. Chegini[3]**

[1]*Department of Applied Mathematics, Islamic Azad University, Lahijan, Iran*
[2]*Department of Applied Mathematics, Islamic Azad University, Rasht, Iran*
[3]*Department of Management, Islamic Azad University, Rasht, Iran*

Correspondence should be addressed to Alireza Amirteimoori; aamirteimoori@gmail.com

Received 27 September 2015; Accepted 6 December 2015

Academic Editor: Fazal M. Mahomed

Copyright © 2016 Mohsen Mirzaei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In multidimensional input/output space, the behavior of the firms can be analyzed by using efficient frontier or supporting surfaces of production technology. To this end, mathematicians are interested to use marginal rates of substitutions. The piecewise linear frontier of data envelopment analysis (DEA) technology is not differentiable at the extreme points and marginal rates calculation is valid only for small changes in one or more variables. The existing trade-off analysis methods calculate the maximum changes in a specific throughput when another throughput is changed. We will show that binding efficient supporting surfaces of an efficient point may be used to define different marginal rates of substitutions and in this sense, we get different marginal rates to each frontier point.

## 1. Introduction

Data envelopment analysis (DEA) is a LP-based nonparametric technique for measuring the relative performances of firms that use multiple inputs to produce multiple outputs. The literature on the performance measurement using DEA has grown substantially since Charnes et al. [1] introduced the traditional CCR model based on the maximum radial reduction in inputs. It has been widely used in performance measurement of many business and industry applications. A complete literature and surveys can be found in Cooper et al. [2] and Cook and Seiford [3].

While much DEA-based research has been directed to the application of DEA for performance measurement, fewer works have been made on its properties as a production function model. Production technology in DEA is extrapolated with the observed inputs and outputs and the boundary points of this set construct the efficient frontier or supporting hyperplanes of the technology. Supporting hyperplanes of production technology enable us to analyze the relation between two different throughputs. Mathematicians are very interested to work with efficient frontiers or supporting surfaces of production possibility set to analyze the behavior of firms in the multidimensional input/output space.

The knowledge of trade-offs in a production process is an important subject to the managers. For instance, managers are interested to know the additional amount of a certain input that is required in order to increase a particular output by a small fixed amount.

Huang et al. [4] proposed a general method for calculating the rates of change of outputs to the inputs along efficient surface of DEA production set. Rosen et al. [5] directly studied the problem of marginal rate of substation on efficient frontier and presented a general framework for the calculation of trade-offs between two variables in DEA. Cooper et al. [6] modified the classic additive DEA model and from the optimal slack values of this model they derived the marginal rates and elasticities of substitution.

Krivonozhko et al. [7] have used the supporting hyperplanes and efficient surfaces to calculate marginal rates of substitution in DEA. In the meantime, using supporting hyperplanes of production technology, Førsund and Hjalmarsson [8] have calculated scale elasticity in DEA models. Førsund et al. [9] have also proposed two ways of

obtaining numerical values of scale elasticity by direct-indirect approaches and then they compared the two approaches by real data. Balk et al. [10] generalized the concept of scale elasticity to accommodate changes in any given direction in input-output space. Khoshandam et al. [11] proposed a production in which a group of variables are changed in a given direction and the effect of this change on some throughput is calculated. In their second paper, Khoshandam et al. [11] studied the problem of marginal rates of substitution in the presence of nondiscretionary factors.

The abovementioned approaches calculate a single measure of marginal rates of substitution and these trade-off analysis methods are given in optimistic case in the sense that they calculate the maximum changes in a specific throughput when another throughput is changed by a small quantity. In this sense, the new frontier point may be very far from the original frontier point. However we can equivalently look for a piece of the frontier in which the new frontier point is closing as much as possible to the original point.

We show that binding supporting surfaces of production set in an efficient point can be used to define marginal rates of substitutions and hence we may calculate different marginal rates in efficient points. We show this geometrically by a simple example.

The rest of the paper is organized as follows: marginal rates substitutions are introduced in the next section. Section 3 introduces two marginal rates in optimistic and pessimistic cases. Marginal rates in the presence of undesirable outputs and in the absence of explicit inputs are given in Section 4. A real application is used to illustrate the proposed approach. The paper ends with concluding remarks.

## 2. Marginal Rates of Substitution

In mathematical economics, marginal rate of substitution is employed to calculate the relative marginal utility. Mathematicians are often interested to work with efficient surfaces, supporting hyperplanes and Pareto-efficient surfaces. In multidimensional input-output space, the behavior of a production unit can easily be read using the supporting hyperplanes of the DEA production possibility set without any loss of mathematical rigor. In this section, we briefly introduce the marginal rate of substitution. Consider a general process in which an output vector $y = (y_1, \ldots, y_s) \in \mathbb{R}^s_+$ is produced by consuming the input vector $x = (x_1, \ldots, x_m) \in \mathbb{R}^m_+$. There are a set of $n$DMU$_j$: $j = 1, \ldots, n$, each of which is characterized by a throughput vector $z_j = (x_j, y_j)^t$ in which $x_j = (x_{1j}, \ldots, x_{mj})$ and $y_j = (y_{1j}, \ldots, y_{sj})$ are nonzero and nonnegative vectors. To simplify, the throughput matrix is shown as follows:

$$z = (z_1, \ldots, z_n) = \begin{pmatrix} z_{11} & \cdots & z_{1j} & \cdots & z_{1n} \\ \vdots & & \vdots & & \vdots \\ z_{i1} & \cdots & z_{ij} & \cdots & z_{in} \\ \vdots & & \vdots & & \vdots \\ z_{(m+s),1} & \cdots & z_{(m+s),j} & \cdots & z_{(m+s),n} \end{pmatrix}. \quad (1)$$

Let $F(x, y) = 0$ be the boundary of the production technology and without loss of generality, we assume that

$$\frac{\partial F(x, y)}{\partial x_i} < 0 \quad i = 1, \ldots, m,$$
$$\frac{\partial F(x, y)}{\partial y_r} > 0 \quad r = 1, \ldots, s. \quad (2)$$

Moreover, we assume that the production technology shows free disposability or in other words $F(x, y)$ is assumed to be continuously differentiable. Let $z_o = (x_o, y_o)$ be a frontier point; that is, $F(z_o) = F(x_o, y_o) = 0$.

*Definition 1.* The marginal rate of substitution of the $j$th throughput to the $k$th throughput at the frontier point $z_o$ is defined as follows:

$$\mathrm{MR}^+_{jk}(z_o) = \left(\frac{\partial z_{jo}}{\partial z_{ko}}\right)_{Z_o^+} = \lim_{h \to 0^+} \frac{f_j(z_{1o}, \ldots, z_{ko} + h, \ldots, z_{m+so})}{h},$$
$$\mathrm{MR}^-_{jk}(z_o) = \left(\frac{\partial z_{jo}}{\partial z_{ko}}\right)_{Z_o^-} = \lim_{h \to 0^-} \frac{f_j(z_{1o}, \ldots, z_{ko} + h, \ldots, z_{m+so})}{h}. \quad (3)$$

As we know, in real applications, $F(x, y)$ is not known exactly and hence we instead use the empirical production technologies such as DEA production possibility set. This set is structured axiomatically and the boundary points of this set construct the empirical production function that is a piecewise linear function. Due to the nature of the DEA efficient frontier, marginal rates are not uniquely defined in the extreme efficient units on the frontier and hence, we should calculate the marginal rate of substitution to right and left.

A DEA-based procedure to calculate the marginal rates of substitution of throughput $j$ to throughput $k$ from right at the frontier point $z_o = (x_o, y_o)$ is given by Asmild et al. [12] as follows:

$$\mathrm{MR}^+_{jk}(z_o) = \frac{z^*_{jo} - z_{jo}}{h}, \quad (4)$$

in which $h$ is a small positive number and $z^*_{jo}$ is the optimal solution to the following LP problem:

$$\begin{aligned} \text{Max} \quad & Z^*_{jo} \\ \text{s.t.} \quad & \sum_{t=1}^{n} \lambda_t z_{lt} \geq z_{lo} \quad l \neq j, k \\ & \sum_{t=1}^{n} \lambda_t z_{jt} \geq z^*_{jo} \\ & \sum_{t=1}^{n} \lambda_t z_{kt} \geq z_{ko} + h \end{aligned}$$

Table 1: The data set for simple example.

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| $x_1$ | 0.9 | 0.5 | 1.1 | 0.2 | 2.2 | 2.8 | 3 |
| $x_2$ | 1.63 | 1.36 | 1.55 | 2.15 | 2.04 | 1.40 | 2.04 |
| $y_1$ | 0.65 | 0.35 | 0.65 | 0.55 | 1.2 | 0.8 | 1.3 |

$$\sum_{t=1}^{n} \lambda_t = 1$$
$$\lambda_t \geq 0. \tag{5}$$

Substituting $-h$ instead of $h$ in the foregoing procedure gives the left marginal rates of substitution.

In this procedure, the new frontier point $z'_o = (z_{1o}, \ldots, z_{ko} + h, \ldots, z^*_{jo}, \ldots, z_{m+so})$ is calculated in optimistic case in the sense that $z^*_{jo}$ is maximized. There may be another point on the frontier and we can use this point instead of $z^*_{jo}$. In the next section, two different marginal rates of substitutions are calculated, one in optimistic case and another one in pessimistic case.

## 3. Different Marginal Rates

As we stated in the previous section, the classic trade-off analysis is given in the optimistic case and hence they determine the maximum changes in a specific component of the throughput vector when a specific component is increased (or decreased) by a small quantity.

Generally, we solve the following LP problem to calculate the marginal rates of substitution:

$$\begin{aligned} &\text{Max} \quad z'_{ko} \\ &\text{s.t.} \quad z'_o \in T, \end{aligned} \tag{6}$$

in which $z'_o = (z_{1o}, \ldots, z_{ko} + h, \ldots, z_{m+so})$ is the new input/output vector after changes. The maximization objective function in the LP problem (6) is given in optimistic case and hence, it finds the maximum value to a specific throughput when another throughput is changed by a small quantity. In model (6), we look for a piece of the frontier in which the specified component is maximized. We can equivalently look for a piece of the frontier in which the component is minimized. Needless to say in both cases the new point is a frontier point and this is what we want.

We show this by the following simple example. We use a simple example with seven DMUs with two inputs $x_1$ and $x_2$ and one output $y_1$. The data are summarized in Table 1.

The BCC model of Banker et al. [13] has been used and we saw that all DMUs are efficient. The production technology in variable returns to scale case is depicted in Figure 1.

Consider $\text{DMU}_\text{A}$ with $z_1$: $(-x_{11}, -x_{21}, y_1) = (-0.9, -1.63, 0.65)$. Suppose that we change $y_1$ to $y_1 + h$ and the response of this change on $x_1$ and $x_2$ is needed. As the figure shows, in $\text{DMU}_\text{A}$, four efficient facets, ABC, ABD, ADE, and ACE,

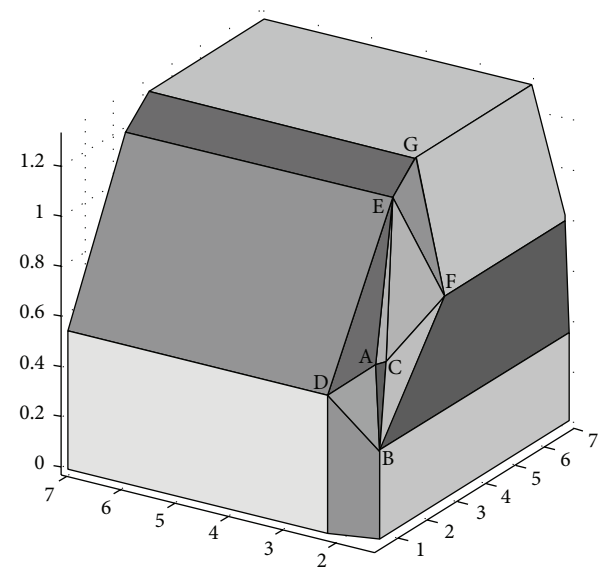



Figure 1: The production technology.

are binding and the optimistic approach finds a new point on facet ACE with $(-x_{11}, -x_{21}, y^*_1) = (-1.1364, -1.7045, 0.75)$ and $\text{MR}^{(+)}_{13} = 2.364$ and $\text{MR}^{(+)}_{23} = 0.745$. $\text{MR}^{(+)}_{13}$ is the value of change in first input that results when the single output is increased by 0.1 units. Note that $(-x_{11}, -x_{21}, y^*_1) = (-1.1364, -1.7045, 0.75)$ is the result of the following model:

$$\begin{aligned} \text{Min} \quad & z^*_1 + z^*_2 \\ \text{s.t.} \quad & -0.9\lambda_1 - 0.5\lambda_2 - 1.1\lambda_3 - 0.2\lambda_4 - 2.2\lambda_5 \\ & \quad - 2.8\lambda_6 - 3\lambda_7 \geq -z^*_1 \\ & -1.63\lambda_1 - 1.36\lambda_2 - 1.55\lambda_3 - 2.15\lambda_4 - 2.04\lambda_5 \\ & \quad - 1.4\lambda_6 - 2.04\lambda_7 \geq -z^*_2 \\ & 0.65\lambda_1 - 0.35\lambda_2 - 0.65\lambda_3 - 0.55\lambda_4 - 1.2\lambda_5 \\ & \quad - 0.8\lambda_6 - 1.3\lambda_7 \geq 0.65 + 0.1 \\ & \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 = 1 \\ & \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7 \geq 0. \end{aligned} \tag{7}$$

However, all points on the facets ABC, ABD, and ADE can provide us a frontier point to calculate a new marginal rate. If we change the direction of movement, a new point on another frontier is obtained. For example, point $(-x_{11}, -x_{21}, y^*_1) = (-0.8154, -2.1162, 0.75)$ is a frontier point on the facet ADE with $\text{MR}^{(-)}_{13} = -0.8462$ and $\text{MR}^{(-)}_{23} = -4.8615$. Let us compare these two rates: In $\text{DMU}_\text{A}$, $\text{MR}^{(+)}_{13}$ and $\text{MR}^{(+)}_{23}$ mean that if we increase $y_1$ from 0.65 to 0.75, we will remain on the efficient frontier if the two inputs are, respectively, changed to 1.1364 and 1.7045, while $\text{MR}^{(-)}_{13}$ and $\text{MR}^{(-)}_{23}$ mean that, with $h = +0.1$, we will remain on the efficient frontier if the two inputs are, respectively, changed to 0.8154 and 2.1162. It is interesting; there is a substantial gap between the two new points on two different efficient facets. Note that

$(-x_{11}, -x_{21}, y_1^*) = (-0.8154, -2.1162, 0.75)$ is the optimal solution of the following LP problem:

$$\begin{aligned}
\text{Min} \quad & -z_1^* + z_2^* \\
\text{s.t.} \quad & -0.9\lambda_1 - 0.5\lambda_2 - 1.1\lambda_3 - 0.2\lambda_4 - 2.2\lambda_5 \\
& \quad - 2.8\lambda_6 - 3\lambda_7 \geq -z_1^* \\
& -1.63\lambda_1 - 1.36\lambda_2 - 1.55\lambda_3 - 2.15\lambda_4 - 2.04\lambda_5 \\
& \quad - 1.4\lambda_6 - 2.04\lambda_7 \geq -z_2^* \\
& 0.65\lambda_1 - 0.35\lambda_2 - 0.65\lambda_3 - 0.55\lambda_4 - 1.2\lambda_5 \\
& \quad - 0.8\lambda_6 - 1.3\lambda_7 \geq 0.65 + 0.1 \\
& \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 = 1 \\
& \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7 \geq 0.
\end{aligned} \tag{8}$$

As a result, we have calculated two marginal rates: $(\text{MR}_{13}^{(+)}, \text{MR}_{23}^{(+)})$ and $(\text{MR}_{13}^{(-)}, \text{MR}_{23}^{(-)})$.

We can summarize the foregoing procedure as follows.

To calculate the two marginal rates of substitutions of throughputs in group $M$ to throughputs in $N$, at the frontier point $z_o = (-x_o, y_o)$, we can use the following three-step procedure:

(i) Choose a small number $h$ for throughput $k$.

(ii) Solve the following two LP problems ($d_l^{(1)}$ and $d_l^{(2)}$ are user-defined coefficients):

$$\begin{aligned}
z_o^{(+)} = \text{Max} \quad & \sum_{l \in M} d_l^{(1)} z_{lo}^* \\
\text{s.t.} \quad & \sum_{t=1}^{n} \lambda_t z_{tl} \geq z_{lo}^*, \quad l = 1, 2, \ldots, m+s, \\
& \sum_{t=1}^{n} \lambda_t = 1, \\
& z_{lo}^* = z_{lo}, \quad l \notin M,\ l \notin N, \\
& z_{ko}^* = z_{ko} + h, \quad l \in N, \\
& \lambda_t, z_{lo}^* \geq 0, \quad \forall t, l,
\end{aligned} \tag{9}$$

$$\begin{aligned}
z_o^{(-)} = \text{Max} \quad & \sum_{l \in M} d_l^{(2)} z_{lo}^* \\
\text{s.t.} \quad & \sum_{t=1}^{n} \lambda_t z_{tl} \geq z_{lo}^*, \quad l = 1, 2, \ldots, m+s, \\
& \sum_{t=1}^{n} \lambda_t = 1, \\
& z_{lo}^* = z_{lo}, \quad l \notin M,\ l \notin N, \\
& z_{ko}^* = z_{ko} + h, \quad l \in N, \\
& \lambda_t, z_{lo}^* \geq 0, \quad \forall t, l.
\end{aligned} \tag{10}$$

(iii) Calculate the optimistic and pessimistic marginal rates of substitutions from right as follows:

$$\begin{aligned}
\text{MR}_{jk}^{(+)}(z_o) &= \frac{z_{jo}^{(+)} - z_{jo}}{h}, \\
\text{MR}_{jk}^{(-)}(z_o) &= \frac{z_{jo}^{(-)} - z_{jo}}{h} \\
& \qquad \forall j \in M,\ k \in N.
\end{aligned} \tag{11}$$

Note that both marginal rates of substitutions from left can be calculated by replacing $-h$ instead of $h$. It is easy to show that both projection points obtained from models (9) and (10) are frontier points. Moreover, different weight vectors $d$ lead to different marginal rates on different efficient facets. The weight vector $d$ is a user-defined vector that determines the direction of moving to efficient facet.

## 4. Marginal Rates in the Presence of Bad Outputs

Now suppose we face a production system without explicit inputs and the system has just desirable and undesirable outputs (indeed, in such a system, inputs consumption is not important to the decision makers). Such a system may have considerable applications in real life when the process has focused on outputs production. In what follows, we reformulate the procedure to production systems without explicit inputs in the presence of both, desirable and undesirable outputs.

Suppose we have $n\text{DMU}_j : j = 1, \ldots, n$, with the good output vector $y_j = (y_{1j}, \ldots, y_{sj}) \geq 0$ and bad output vector $z_j = (z_{1j}, \ldots, z_{pj}) \geq 0$. Masoumzadeh et al. [14] have used the weak disposability assumption of Shephard [15] to propose the following linear production technology to handle undesirable outputs in production system without explicit inputs:

$$\begin{aligned}
T_{\text{WI}} = \Bigg\{ & (y, z) : \sum_{j=1}^{n} \lambda_j y_j \geq y, \sum_{j=1}^{n} \lambda_j z_j = z, \sum_{j=1}^{n} \lambda_j = \theta,\ 0 \\
& \leq \theta \leq 1,\ \lambda_j \geq 0,\ \forall j \Bigg\}.
\end{aligned} \tag{12}$$

If we want to evaluate firm "$o$" in terms of the abatement potential in bad outputs, we can use the following linear programming problem:

$$\begin{aligned}
\text{Min} \quad & \rho \\
\text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, 2, \ldots, s, \\
& \sum_{j=1}^{n} \lambda_j z_{pj} \geq \rho z_{po}, \quad p = 1, 2, \ldots, P, \\
& \sum_{j=1}^{n} \lambda_j = \theta, \\
& 0 \leq \theta \leq 1,\ \lambda_j \geq 0, \quad j = 1, 2, \ldots, n.
\end{aligned} \tag{13}$$

Now, we reformulate the foregoing three-step procedure to calculate the marginal rates of substitution of the undesirable outputs in $N$ to the desirable outputs in $M$ from right:

(i) Choose a small number $h$.

(ii) Solve the following two LP problems:

$$\begin{aligned}
\text{Max} \quad & \sum_{p\in M} d_p^{(1)} z_p^{(+)} \\
\text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, 2, \ldots, s, \ r \notin N, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} + h, \quad r \in N, \\
& \sum_{j=1}^{n} \lambda_j z_{pj} \geq z_{po}, \quad p = 1, 2, \ldots, P, \ p \notin M, \\
& \sum_{j=1}^{n} \lambda_j z_{pj} \geq z_p^{(+)}, \quad p \in M, \\
& \sum_{j=1}^{n} \lambda_j = \theta, \\
& 0 \leq \theta \leq 1, \ \lambda_j, z_p^{(+)} \geq 0, \quad \forall j, p,
\end{aligned} \tag{14}$$

$$\begin{aligned}
\text{Max} \quad & \sum_{p\in M} d_p^{(2)} z_p^{(-)} \\
\text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro}, \quad r = 1, 2, \ldots, s, \ r \notin N, \\
& \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{ro} + h, \quad r \in N, \\
& \sum_{j=1}^{n} \lambda_j z_{pj} \geq z_{po}, \quad p = 1, 2, \ldots, P, \ p \notin M, \\
& \sum_{j=1}^{n} \lambda_j z_{pj} \geq z_p^{(-)}, \quad p \in M, \\
& \sum_{j=1}^{n} \lambda_j = \theta, \\
& 0 \leq \theta \leq 1, \ \lambda_j, z_p^{(-)} \geq 0, \quad \forall j, p.
\end{aligned} \tag{15}$$

(iii) Calculate the optimistic and pessimistic marginal rates of substitutions from right as follows:

$$\begin{aligned}
\text{MR}_{pk}^{(+)}(y_o, z_o) &= \frac{z_p^{(+)} - z_{po}}{h}, \\
\text{MR}_{pk}^{(-)}(y_o, z_o) &= \frac{z_p^{(-)} - z_{po}}{h}, \\
& \forall p \in M, \ k \in N.
\end{aligned} \tag{16}$$

Again, both marginal rates of substitutions from left can be calculated by replacing $-h$ instead of $h$.

## 5. Real World Application

To illustrate the marginal rates of substitution in the real cases, consider the following example that uses data describing the quality of the services provided by the largest hydropower plants operating in Brazil. The data has been taken from Zanella et al. [16]. These plants are under the performance monitoring of the Brazilian Independent System Operator (ISO) with capacity above 100-megawatt (MW) electricity. The data consist of 81 hydropower plants in Brazil. As Zanella et al. [16] stated, around 90 percent of the electrical energy of Brazil is supplied by hydropower plants and these 81 plants cover 82% of the Brazilian total installed capacity. In this sense, this selected sample is a good representation of the power supply in Brazil.

The data set consist of three different indicators as outputs that have been provided by Brazilian Independent System Operator. Of these three indicators, two indicators are considered as undesirable outputs and one is the single desirable output.

Now, we introduce these three indicators: The proportion of time that the plant is available or in operation is considered as the single desirable output (availability factor ($y_1$)). The first undesirable output is the average time to repair ($z_1$) which accounts for the number of hours unavailable to operate and under forced maintenance per forced outages. Finally, the failure rate ($z_2$) is considered as the second undesirable output. This indicator shows the number of failures occurring per hour of operation. Model (13) has been used in this data set and we saw that ten plants are efficient (plants #4, #6, #7, #13, #17, #20, #38, #39, #62, and #74). The data set and results are listed in Table 2. Columns five and ten show the relative efficiencies obtained from our proposed model (13).

Suppose that $M = \{1, 2\}$ and $N = \{1\}$ and we want to calculate the response of changes in the undesirable outputs $z_1$ and $z_2$ to the change in the desirable output $y$. It is important to note that both $h = 0.5$ and $h = -0.5$ yield feasible points for all efficient plants. This means that increasing or decreasing 0.5 units in the desirable output leads to a point inside of the production technology. The results for two different values of $h$ are given in Tables 3 and 4. The first two columns of Tables 3 and 4 show the new values of $z_1$ and $z_2$. The marginal rates of substitutions are given in the last two columns.

Now, we analyze the results for three sample plants, plants 7, 13, and 17. As the last two columns of Table 3 show, of ten plants, nine plants confront reduction in the bad outputs if we increase the single desirable output by 0.5 units.

First consider plant 13: If we increase the availability factor from its current value 95.64 to 96.14 ($h = 0.5$), with $d = (1, 1)$, the average time to repair ($z_1$) and the failure rate ($z_2$) are, respectively, decreased to 0.1293 and 0.2568 with $\text{MR}_{11}^{(+)} = -1.9414$ and $\text{MR}_{12}^{(+)} = -0.0028$. In plant 17, if we increase 0.5 units in the availability factor, both undesirable factors $z_1$ and $z_2$ are increased to 0.1307 and 0.2613, respectively. The marginal rates in this case are $\text{MR}_{11}^{(+)} = 0.0013$ and

Table 2: The data set and results.

| Plant $J$ | $y_1$ | $z_1$ | $z_2$ | Efficiency | Plant $J$ | $y_1$ | $z_1$ | $z_2$ | Efficiency |
|---|---|---|---|---|---|---|---|---|---|
| U1 | 93.62 | 165.56 | 4.25 | 0.0699 | U41 | 92.8 | 6.38 | 7.55 | 0.0331 |
| U2 | 56.97 | 24.71 | 3.09 | 0.0509 | U42 | 92.22 | 4.88 | 5.01 | 0.0496 |
| U3 | 90.28 | 15.11 | 1.3 | 0.1947 | U43 | 97.9 | 6.39 | 10.82 | 0.347 |
| U4 | 98.46 | 9.94 | 2.35 | 1 | U44 | 92.95 | 31.83 | 4.74 | 0.0539 |
| U5 | 93.45 | 15.26 | 2.83 | 0.0902 | U45 | 73.52 | 12.77 | 3.32 | 0.0601 |
| U6 | 98.6 | 14.92 | 1.58 | 1 | U46 | 96.19 | 24.23 | 2.86 | 0.0933 |
| U7 | 86.98 | 314.69 | 1.51 | 1 | U47 | 95.21 | 5.76 | 5.78 | 0.0444 |
| U8 | 89.23 | 273.47 | 1.86 | 0.318 | U48 | 92.34 | 8.15 | 0.89 | 0.2879 |
| U9 | 94.39 | 3.72 | 2.44 | 0.1044 | U49 | 65.99 | 79.01 | 3.07 | 0.0641 |
| U10 | 93.66 | 38.9 | 1.5 | 0.1865 | U50 | 85.54 | 1.15 | 12.19 | 0.122 |
| U11 | 96.74 | 0.72 | 1.24 | 0.3203 | U51 | 96.48 | 5.13 | 6.55 | 0.0397 |
| U12 | 95.16 | 136.68 | 2.66 | 0.1208 | U52 | 73.89 | 6.17 | 3.7 | 0.0539 |
| U13 | 95.64 | 1.1 | 0.26 | 1 | U53 | 95.72 | 0.58 | 2.07 | 0.2295 |
| U14 | 97.18 | 20.31 | 0.57 | 0.9174 | U54 | 97.61 | 24.21 | 3.61 | 0.2453 |
| U15 | 90.85 | 5.09 | 15.16 | 0.0245 | U55 | 94.79 | 3.03 | 14.11 | 0.0446 |
| U16 | 94.29 | 5.36 | 2.44 | 0.1044 | U56 | 93.1 | 21.82 | 9.86 | 0.0255 |
| U17 | 96.66 | 0.13 | 0.26 | 1 | U57 | 78.82 | 74.25 | 5.86 | 0.0379 |
| U18 | 79.25 | 574.77 | 4.51 | 0.0973 | U58 | 97.83 | 11.87 | 6.31 | 0.293 |
| U19 | 92.44 | 19.31 | 3.15 | 0.0804 | U59 | 92.49 | 92.49 | 8 | 0.0324 |
| U20 | 97.69 | 2.79 | 1.19 | 1 | U60 | 91.65 | 3.87 | 1.93 | 0.1283 |
| U21 | 95.04 | 13.59 | 2.78 | 0.0932 | U61 | 91.87 | 14.14 | 3.44 | 0.0726 |
| U22 | 84.07 | 26.03 | 14.72 | 0.0154 | U62 | 98.17 | 0.33 | 10.11 | 1 |
| U23 | 92.95 | 7.15 | 10.82 | 0.0231 | U63 | 96.96 | 5.47 | 4.08 | 0.146 |
| U24 | 75.23 | 43.27 | 9.83 | 0.0208 | U64 | 93.38 | 6.99 | 1.53 | 0.1662 |
| U25 | 97.56 | 6.94 | 5.44 | 0.2658 | U65 | 93.87 | 2.85 | 0.39 | 0.6628 |
| U26 | 93.08 | 0.97 | 15.58 | 0.183 | U66 | 88.28 | 4.83 | 10.46 | 0.0247 |
| U27 | 94.58 | 5.56 | 1.81 | 0.1416 | U67 | 78.45 | 24.45 | 4.12 | 0.0521 |
| U28 | 96.79 | 5.57 | 6.86 | 0.0645 | U68 | 95.57 | 50.45 | 4.94 | 0.0539 |
| U29 | 95.77 | 4.7 | 1.07 | 0.2436 | U69 | 93.98 | 38.87 | 21.63 | 0.0117 |
| U30 | 94.42 | 17.88 | 0.9 | 0.3049 | U70 | 98.2 | 6.57 | 4.12 | 0.8721 |
| U31 | 90.96 | 30.29 | 6.45 | 0.0384 | U71 | 96.98 | 34.69 | 8.86 | 0.0546 |
| U32 | 96.87 | 2.33 | 3.38 | 0.1792 | U72 | 96.74 | 2.22 | 13.19 | 0.0639 |
| U33 | 90.88 | 1.83 | 7.89 | 0.0702 | U73 | 93.97 | 6.61 | 21.4 | 0.0196 |
| U34 | 91.04 | 143.22 | 2.9 | 0.1049 | U74 | 97.64 | 0.32 | 2.58 | 1 |
| U35 | 97.01 | 3.86 | 8.1 | 0.1156 | U75 | 84.68 | 8.28 | 3.68 | 0.0622 |
| U36 | 85.86 | 14.29 | 19.88 | 0.0116 | U76 | 72.69 | 116.59 | 10.11 | 0.0202 |
| U37 | 96.21 | 0.93 | 1.37 | 0.189 | U77 | 95.86 | 11.94 | 2.6 | 0.1005 |
| U38 | 98.4 | 20.24 | 0.75 | 1 | U78 | 94.62 | 11.25 | 2.36 | 0.1092 |
| U39 | 98.45 | 7.05 | 5.15 | 1 | U79 | 94.81 | 23.17 | 2.06 | 0.1289 |
| U40 | 73.91 | 4.21 | 1.49 | 0.1343 | U80 | 97.61 | 1.01 | 10.01 | 0.3034 |
| | | | | | U81 | 80.03 | 3.23 | 7.07 | 0.0335 |

Table 3: Results with $h = 0.5$ and $d = (1, 1)$.

| | $z_1^{(+)}$ | $z_2^{(+)}$ | $\mathrm{MR}_{11}^{(+)}$ | $\mathrm{MR}_{12}^{(+)}$ |
|---|---|---|---|---|
| U4 | 0.13313 | 0.2662 | −19.6138 | −4.1676 |
| U6 | 0.1333 | 0.2666 | −29.5734 | −2.6269 |
| U7 | 0.1177 | 0.2353 | −629.145 | −2.5494 |
| U13 | 0.1293 | 0.2586 | −1.9414 | −0.0028 |
| U17 | 0.1307 | 0.2613 | 0.0013 | 0.0027 |
| U20 | 0.1321 | 0.26411 | −5.3159 | −1.8518 |
| U38 | 0.1330 | 0.2660 | −40.2140 | −0.9680 |
| U39 | 0.1330 | 0.2662 | −13.8338 | −9.7677 |
| U62 | 0.1327 | 0.2654 | −0.3946 | −19.6892 |
| U74 | 0.1320 | 0.2634 | −0.3760 | −4.6320 |

Table 4: Results with $h = -0.5$ and $d = (1, 1)$.

| | $z_1^{(+)}$ | $z_2^{(+)}$ | $\mathrm{MR}_{11}^{(+)}$ | $\mathrm{MR}_{12}^{(+)}$ |
|---|---|---|---|---|
| U4 | 0.1318 | 0.2635 | 19.6165 | 4.1730 |
| U6 | 0.1319 | 0.2639 | 29.576 | 2.6323 |
| U7 | 0.1163 | 0.2326 | 629.15 | 2.5548 |
| U13 | 0.1270 | 0.2559 | 1.9441 | 0.0082 |
| U17 | 0.1293 | 0.2587 | 0.0013 | 0.0027 |
| U20 | 0.1307 | 0.2614 | 5.3186 | 1.8572 |
| U38 | 0.1317 | 0.2633 | 40.2167 | 0.9733 |
| U39 | 0.1317 | 0.2635 | 13.8365 | 9.7731 |
| U62 | 0.1314 | 0.2627 | 0.3973 | 19.6946 |
| U74 | 0.1306 | 0.2613 | 0.3787 | 4.6374 |

$\mathrm{MR}_{12}^{(+)} = 0.0027$. In plant 7, we confront a different case; 0.5-unit increment in the availability factor leads to a decrement of 1.2747 in $z_1$ and a huge decrement of 314.5723 in the failure rate. Now, consider $h = -0.5$ that the results are given in Table 4.

Again, for plant 13, if we decrease the availability factor from its current value 95.64 to 95.14 ($h = -0.5$), with $d = (1, 1)$, the average time to repair and the failure rate are, respectively, decreased to 0.1270 and 0.2559 with $\mathrm{MR}_{11}^{(+)} = -1.9441$ and $\mathrm{MR}_{12}^{(+)} = 0.0082$.

An interesting point in this example is that we have run the procedure for different nonnegative $d$ and all directions lead to same results.

## 6. Conclusions

Marginal rates of substitution are one of the most frequently studied subjects in the content of nonparametric efficiency analysis. The problem of calculating marginal rates of substitution has been studied from different point of views. The existing trade-off analysis methods calculate the maximum changes in a specific throughput when another throughput is changed. In this case, a single measure of marginal rate is calculated. The current paper has shown how supporting surfaces of production technology in nonparametric techniques can be used to analyze the behavior of the firms. In this paper, we have shown that binding efficient supporting surfaces can be used to define different marginal rates of substitutions and in this sense, instead of getting a single measure of marginal rate of substitution as was done earlier in the literature, we can get different marginal rates of substitution to each frontier point. Numerical examples are used to illustrate the applicability of the proposed approach.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

Financial support by Rasht Branch, Islamic Azad University Grant no. 4.5830, is gratefully acknowledged.

## References

[1] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.

[2] W. W. Cooper, L. M. Seiford, and J. Zhu, *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers, Norwell, Mass, USA, 2004.

[3] W. D. Cook and L. M. Seiford, "Data Envelopment Analysis (DEA)—thirty years on," *European Journal of Operational Research*, vol. 192, no. 1, pp. 1–17, 2009.

[4] Z. Huang, S. X. Li, and J. J. Rousseau, "Determining rates of change in data envelopment analysis," *Journal of the Operational Research Society*, vol. 48, no. 6, pp. 591–593, 1997.

[5] D. Rosen, C. Schaffnit, and J. C. Paradi, "Marginal rates and two dimensional level curve in DEA," *Journal of Productivity Analysis*, vol. 9, no. 3, pp. 205–232, 1998.

[6] W. W. Cooper, K. S. Park, and J. T. Pastor Ciurana, "Marginal rates and elasticities of substitution with additive models in DEA," *Journal of Productivity Analysis*, vol. 13, no. 2, pp. 105–123, 2000.

[7] V. E. Krivonozhko, O. B. Utkin, A. V. Volodin, I. A. Sablin, and M. Patrin, "Constructions of economic functions and calculations of marginal rates in DEA using parametric optimization methods," *Journal of the Operational Research Society*, vol. 55, no. 10, pp. 1049–1058, 2004.

[8] F. R. Førsund and L. Hjalmarsson, "Are all scales optimal in DEA? Theory and empirical evidence," *Journal of Productivity Analysis*, vol. 21, no. 1, pp. 25–48, 2004.

[9] F. R. Førsund, L. Hjalmarsson, V. E. Krivonozhko, and O. B. Utkin, "Calculation of scale elasticities in DEA models: direct and indirect approaches," *Journal of Productivity Analysis*, vol. 28, no. 1-2, pp. 45–56, 2007.

[10] B. M. Balk, R. Färe, and G. Karagiannis, "On directional scale elasticities," *Journal of Productivity Analysis*, vol. 43, no. 1, pp. 99–104, 2015.

[11] L. Khoshandam, R. K. Matin, and A. Amirteimoori, "Marginal rates of substitution in DEA with undesirable outputs: a directional approach," *Measurement*, vol. 68, pp. 49–57, 2015.

[12] M. Asmild, J. C. Paradi, and D. N. Reese, "Theoretical perspectives of trade-off analysis using DEA," *Omega*, vol. 34, no. 4, pp. 337–343, 2006.

[13] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in data envelopment analysis," *Management Science*, vol. 30, no. 9, pp. 1078–1092, 1984.

[14] A. Masoumzadeh, M. Toloo, and A. Amirteimoori, "Performance assessment in production systems without explicit inputs: an application to basketball players," *IMA Journal of Management Mathematics*, 2014.

[15] R. W. Shephard, *Theory of Cost and Production Functions*, Princeton University Press, Princeton, NJ, USA, 1970.

[16] A. Zanella, A. S. Camanho, and T. G. Dias, "Undesirable outputs and weighting schemes in composite indicators based on data envelopment analysis," *European Journal of Operational Research*, vol. 245, no. 2, pp. 517–530, 2015.