

ISSN: 1940-9893 (Print), 1940-9907(Online) Vol.1, No.3, December 2008

# **Journal of Service Science and Management**

Chief Editor : **Samuel Mendlinger**



# Journal Editorial Board

ISSN: 1940-9893 (Print), 1940-9907(Online)

<http://www.scirp.org/journal/jssm>

---

## Editor-in-Chief

**Prof. Samuel Mendlinger**

Boston University, USA

## Editorial Board (According to Alphabet)

<b>Dr. Sónia Bentes</b>	Institute of Accounting and Management of Lisbon, Portugal
<b>Dr. Chia-Fen Chi</b>	National Taiwan University of Science & Technology, Taiwan, China
<b>Dr. Tsan-Ming Choi</b>	Hong Kong Polytechnic University, Hong Kong, China
<b>Dr. Yong Deng</b>	Shanghai Jiao Tong University, China
<b>Dr. Jerry Fjermestad</b>	New Jersey Institute of Technology, USA
<b>Dr. Leung Kit-nam Francis</b>	City University of Hong Kong, Hong Kong, China
<b>Dr. Javier Sánchez García</b>	Universitat Jaume I, Spain
<b>Dr. Dai-Ping Hu</b>	Shanghai Jiao Tong University, China
<b>Dr. Patrick C. H. Ma</b>	Hong Kong Polytechnic University, Hong Kong, China
<b>Dr. Mehron Sepehri</b>	Sharif University of Technology, Iran
<b>Dr. Chun-Shan Shen</b>	Chinese Academy of Sciences, China
<b>Dr. Kan-Liang Wang</b>	Xi'an Jiao Tong University, China
<b>Dr. Shu-Liang Wang</b>	Wuhan University, China
<b>Dr. Xu-Song Xu</b>	Wuhan University, China
<b>Dr. Feng-Qi You</b>	Carnegie Mellon University, USA

## Editorial Assistants

<b>Xiao-Qian Qi</b>	Wuhan University, Wuhan, Hubei, China
<b>Dong-Peng Yang</b>	George Washington University

---

## Guest Reviewers (According to Alphabet)

Nezih Altay	Metin Kozak	Kisang Ryu
Sónia R. Bentes	Bert Kramer	John J Shi
Paolo Boccardelli	Jongsu Lee	Florence Mihaela Singer
Truong Ton Hien Duc	Zhen-Shan Lin	Lixin Tang
Volkan S. Ediger	Chi-Lun Liu	Dave Ulrich
Xavier Font	J. J. M. Mendes	Yu Wei
R. Thomas George	Tomoe Moore	Danqin Yang
W. P. Hogan	Antonio Nicolò	Jie Wu
Taehoon Hong	Danny Oppenheimer	Xiang Zhu
Chin-Lung Hsu	Nitin Pangarkar	
Kai Huang	Constantin Popescu	
Timo Hyvönen	Bocquet Rachel	

# CONTENTS

**Vol. 1 No. 3**

**December 2008**

## **Improving Engineering Data Management with Semantic Web Techniques**

K. Wang.....199

## **Intertemporal Pricing and Allotment of Sea-cargo Capacity under Reference Effect**

X. Z. Bu, L. Xu & L. Su.....206

## **Innovation in the Financial Sector: Persistence and Schumpeterian Hypotheses**

R. Napoli.....215

## **A Way to Improve Knowledge Sharing: from the Perspective of Knowledge Potential**

L. L. Zhang, X. Y. Zheng, J. Li, G. L. Nie, G. Q. Huo & Y. Shi.....226

## **Combining Keyword Search Advertisement and Site-Targeted Advertisement in Search Engine Advertising**

L. Chen.....233

## **Market Segmentation for Mobile TV Content on Public Transportation by**

## **Integrating Innovation Adoption Model and Lifestyle Theory**

C. C. Tao.....244

## **Holonic Production System to Obtain Flexibility for Customer Satisfaction**

G. Dominici.....251

## **The Evaluation of Enterprises' Sustainable Superiority Degree of Technical Innovation Based on DEA Method**

B. Jiang, L. Yang & J. Fang.....255

## **Examining Human Resource Competencies and Their Relationship to the Success Factors of HR Profession**

C. S. Long.....259

## **Analysis of University Science Research Capability Elements and Evaluation Based on BP Neural Network**

W. W. Liu & C. S. Shi.....266

# **Journal of Service Science and Management (JSSM)**

## **SUBSCRIPTIONS**

The *Journal of Service Science and Management* (Online at Scientific Research Publishing, **www.SciRP.org**) is published quarterly by Scientific Research Publishing, Inc. 5005 Paseo Segovia, Irvine, CA 92603-3334, USA.

E-mail: **jssm@scirp.org**

**Subscription rates:** Volume 1 2008

Print: \$50 per copy.

Electronic: freely available on **www.SciRP.org**.

To subscribe, please contact Journals Subscriptions Department, E-mail: **jssm@scirp.org**

**Sample copies:** If you are interested in subscribing, you may obtain a free sample copy by contacting Scientific Research Publishing, Inc at the above address.

## **SERVICES**

### **Advertisements**

Advertisement Sales Department, E-mail: **jssm@scirp.org**

### **Reprints (minimum quantity 100 copies)**

Reprints Co-ordinator, Scientific Research Publishing, Inc. 5005 Paseo Segovia, Irvine, CA 92603-3334, USA.

E-mail: **jssm@scirp.org**

## **COPYRIGHT**

Copyright© 2008 Scientific Research Publishing, Inc.

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as described below, without the permission in writing of the Publisher.

Copying of articles is not permitted except for personal and internal use, to the extent permitted by national copyright law, or under the terms of a license issued by the national Reproduction Rights Organization.

Requests for permission for other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works or for resale, and other enquiries should be addressed to the Publisher.

Statements and opinions expressed in the articles and communications are those of the individual contributors and not the statements and opinion of Scientific Research Publishing, Inc. We assume no responsibility or liability for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained herein. We expressly disclaim any implied warranties of merchantability or fitness for a particular purpose. If expert assistance is required, the services of a competent professional person should be sought.

## **PRODUCTION INFORMATION**

For manuscripts that have been accepted for publication, please contact:

**E-mail: jssm@scirp.org**

# Improving Engineering Data Management with Semantic Web Techniques

Kai Wang

Mathematics Department, Guizhou University  
Email: mathsfan@hotmail.com

Received August 27<sup>th</sup>, 2008; revised December 15<sup>th</sup>, 2008; accepted December 22<sup>nd</sup>, 2008.

## ABSTRACT

*Throughout the IT communities, it has been acknowledged that ontology plays a key role in representation and reuse of knowledge. This paper discusses some issues of ontology construction for engineering data management, such as knowledge discovery, knowledge representation and semantic services. All discussions are followed by a running example of engineering data management. Based on the user needs with five phases of engineering design, the issue of knowledge discovery will be presented. The built knowledge base contains basic knowledge models and vocabularies for knowledge integration. With the semantics of semistructured data, the hierarchical relations of concepts in engineering data have been extracted for reuse in future engineering design based on some clustering techniques of data mining. Semantic services for engineering design will be provided with the ontology-based schema.*

**Keywords:** engineering data management, semistructured data, ontology, semantic web

## 1. Introduction

Nowadays, throughout the IT communities, it has been acknowledged that ontology plays a key role in representation and reuse of knowledge. However, there is no so far general methodology for a domain ontology construction [1]. For reuse and representation of knowledge in engineering design, one may consider to use ontology as an unified knowledge model for knowledge representation and vocabularies.

In this paper, a methodology of ontology-based schema for engineering data management (EDM) will be discussed. This paper presents an ontology-based methodology aiming at reuse and representation of knowledge in engineering design. The built knowledge base will consist of basic knowledge models and vocabularies. The ontology-based schema provides formally defined semantics for capturing and reusing of design knowledge. With the machine-interpretable, flexible data structures that can be modified at run time, ontologies will provide a general way to knowledge representation and reuse in the EDM systems. For instance, semantic searching services will be provided among the heterogeneous engineering databases.

The rest of the paper is organized as follows. In Section 2, we present a brief overview of the state of the art for engineering data management. In Section 3, the user needs of an engineering data management are presented. Section 4 discusses the knowledge representation in engineering data management. In Section 5, some facets of ontology-based services will be discussed. In Section 6, the conclusion and future works will be discussed.

## 2. Related Work

### 2.1. Background

For ontology construction of engineering design, we need at first to represent a domain ontology in engineering design. With the integrated knowledgebase, semantic services will be provided with the ability to detect and eliminate inconsistencies in heterogeneous sources [2]. In literature, there are many standards of a construction project, including the phases in construction [2]. In general, these phases can be divided into three parts: planning; programming; design. In this paper, based on a five phases of design, we present a domain ontology of engineering design.

EDM (Engineering Document Management) systems support many facets for maintaining engineering ranging data from schema design to documentation stage [3]. In industry practice, EDM systems are typically embedded in PDM (Product Data Management), PLM (Product Lifecycle Management), and CAE (Computer Aided Engineering). PDM systems provide handling detailed product information, ranging from design to production stage. PLM systems integrate information on CAM systems and CAD systems, and the information about ERP (Enterprise Resource Planning) processes. However, traditional PDM/PLM systems suffer from the inability of reusing of the creative design knowledge, particular in the conceptual design stage. CAE (Computer Aided Engineering) systems also suffer the same deficiencies as PDM/PLM

systems, lacking of the well-structured product representation, especially in the early design phases. Like PDM systems, CAE systems fail to support more complex development activities. Thus, reusing of product knowledge has been hindered by the lack of capabilities to knowledge representation. Furthermore, the capabilities for direct process support are also hindered.

In [4,5], the knowledge representation (KR) formalisms of Quillian's was proposed as the semantic memory models. As a knowledge representation method, ontology represents the relevant domain entities and their relationships by means of classes and relations. Ontologies provide a unified knowledge model and vocabulary for knowledge integration in many specific domains. In general, ontologies provide machine- interpretable, flexible data structures that can be changed and adapted at run time [3,6,7]. In other ways, as XML provides an unified data exchange schema, the Web ontology language OWL [8,9,10] is a language for expressing sophisticated class definitions and properties. In [8,9], Tim Berners-Lee proposed for the formalization of the resources of web (no only the web page, but also all the data and resources on the web). The architecture of semantic web consists of uniform resource identifier at the lowest layer, XML and XML schema at upper layer and XML name space, constructing the fundamental syntax of semantic web.

## 2.2. Document Management System

Engineering Document Management systems(EDM) are important tools for maintaining engineering data [3]. EDM Systems such as Windream [11] or Documentum [12] are widely used in industrial practice for the storage, maintenance, and distribution of documents. A step further, Product Data Management (PDM) systems provide extended facilities for the handling of detailed product information, ranging from design to production stage. Product Lifecycle Management (PLM) systems is the successor of PDM. These systems are widely applied in the fields of industry manufacturing design. However, PLM/PDM systems lack essential capabilities to reflect the function, behavior, and structure for each design phase. Thus, the management and reuse of design knowledge are less suited, particular in the conceptual design stage.

In [13,14,15,16,17,18], the semistructured model was proposed for the data model and algebra of XML.

As our running example of document management system, the semistructured data model will be used for the query and navigating the engineering data. In practice, the running example offers typical services such as version management, change management, notification (if some changes have been committed), and simple navigation and retrieval functionality, shown as Figure 1. In our running example of EDM, the database is modeled in terms of labeled, directed graph, a semistructured data model. For improving efficiency, An indexing structure

has also been provided, where the indexing data with its address are organized in double-level with hierarchy. For example, when a date data is indexed, the year and month data at the first level are stored, meanwhile, the date data is stored at the second level. At running time, the year and month data will be searched at first level. The second searching will be completed within the first matching results. The implementation aspects include some techniques in C++ programming, such as the multiple- user interface, P2P and components implementation in distributed environment. As a consequence, irregular engineering data can be accommodated in this model. In the sequel, the running example offers traditional services, such like query and navigating the data for five stages (introduced in the earlier section) in distributed multiple-user environment, meanwhile, a data object is a collection of attributes without blank-value.

## 3. User Needs

So far, functional modeling research within engineering design theory studies the flows-the materials, energies, and signals on which functions operate [19,20,21]. For functional modeling, domain knowledge conceptualization is critical to ontology construction.

A general knowledge systems methodology of ontology engineering for creating engineering ontology can be found in literature [22], including four typical steps as follows:

- Step 1-Determine the ontologies for engineering data
- Step 2-Enumerate important terms and rules in the ontology
- Step 3-Define the relationship within the ontology, including hierarchy and the properties of classes
- Step 4-Create instances for engineering data

Analysis of the user needs of engineering data management is critical to ontology construction for engineering data management. As an ontology-based methodology, we

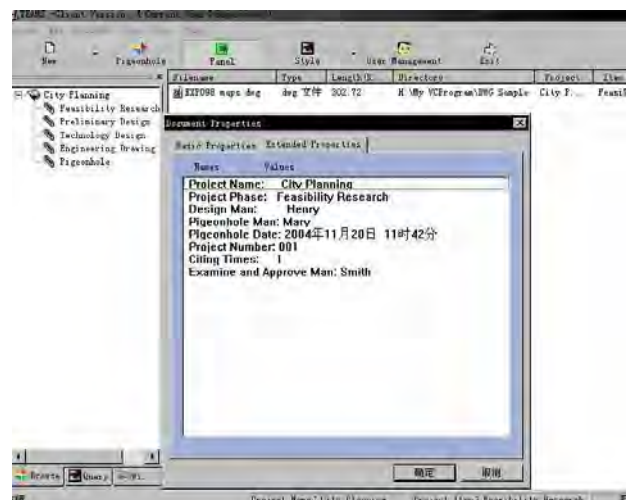


Figure 1. A running example of EDM



should establish our controlled vocabulary (terms) at first. In the running example of EDM system, there could be several kinds of ontologies for different purposes, e.g., design ontology, documentation (pigeonhole) ontology, citing ontology, etc. As machine-interpretable, flexible with data structures ontologies can be modified at run time.

As the proposed paradigm to knowledge integration for engineering design, we at first investigate the domain semantics for engineering design.

There are different phases in the domain of engineering data management [2]. In general, these phases are divided into three parts: planning; programming; design. In this paper, without loss of generality, we consider five phases in the engineering design, e.g., (1) Feasibility Analysis, (2) Preliminary Design, (3) Conceptual Design, (4) Schematic Design, and (5) Documentation, shown in Figure 3. In each phase, the vocabularies are defined. At each stage, one need approving services once specific works completed.

- Step 1–Feasibility Study Phase

In this phase, some methods and techniques are applied to examine the technical feasibility with the cost evaluation in a project, which proceeds to Conceptual Design phase once Feasibility Study has been approved.

- Step 2–Conceptual Design Phase

The goal of the conceptual design phase is to identify very general types of solution. Based on the market requirements and the state-of-the-art of the relevant technology, the degree of innovation, and the scope for innovation in a design project within a product are determined. The project proceeds to Schematic Design phase once Conceptual Design has been approved.

- Step 3–Schematic Design Phase

The goal of the conceptual design phase is to identify the schematic solution of the project. Schematic Design establishes a general scope, a conceptual design, scale and relationships among the components of the project. The primary objective is achieved with a clearly definition, feasible concept. A series of rough plans, known as schematics, should be prepared. In practice, models may be prepared to help visualizing the project. The project proceeds to Design Development phase once Schematic Design has been approved.

- Step 4–Design Development Phase

This phase is to develop more detailed design with the other aspects of the proposed design. The detailed design results will served as the achievements of engineering development. The project proceeds to Documentation Management phase when Schematic Design has been approved.

- Step 5–Documentation Phase

Once the Design Development phase had been approved, the detailed works, as well as the documents in

previous stages should be documented into the library for future citation and/or reuse.

One could illustrate five phases in the engineering design as shown in Figure 2, meanwhile, the engineering document process can be defined as the procedure shown in Figure 3.

As mentioned in the literature, the user needs of a domain ontology are served as the skeleton of a domain ontology. Based on the user needs of a domain ontology, the syntactic elements of an ontology, such as properties, vocabularies and their relationships, are obtained. Abstract concepts can be formed with some clustering techniques, such as formal concept analysis (FCA). Rules with respect to properties and vocabularies can also be formed primitively. In the sequel, knowledge about refined concepts and their relationship is learned statically or dynamically with earlier mentioned models or an iterated process.

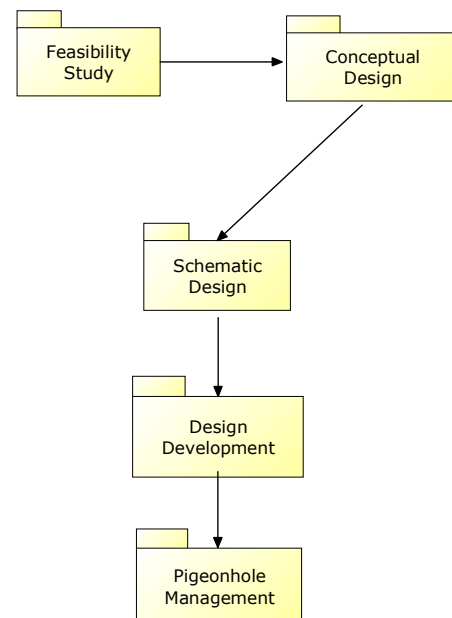


Figure 2. Overview of engineering design process

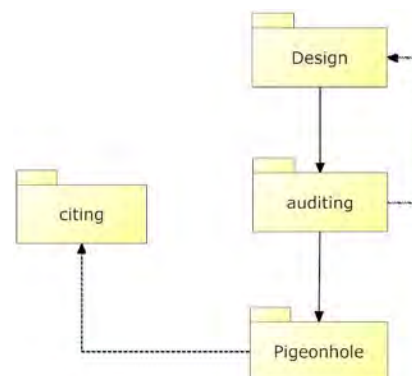


Figure 3. Overview of documentation process

#### 4. Knowledge Discovery in EDM

In literature, knowledge discovery is the process of finding novel, interesting, and useful patterns in data. Domain conceptualization aims at identifying and defining concepts, specifying relationships among the concepts. Capturing domain concepts and their relationships is a bottleneck for ontology engineer [1]. One could capture concepts and their relations automatically by some clustering techniques, such as formal concept analysis (FCA). Then knowledge discovery and semantic services for engineering design will be provided.

For knowledge reuse and semantic services in engineering data management, knowledge discovery is crucial for a knowledge management system. However, in industrial practice, people are reluctant to share knowledge by making their knowledge explicit. One of the reasons for this is that it requires extra effort and they seldom hardly benefit from the explicit knowledge. Meta information about products is often chosen inconsistently and, similarly, work processes and decision making procedures are captured by different users in different ways. Consequently, knowledge acquisition is severely limited and could be found by some query. A solution to these problems is to automatic knowledge acquisition, called automatic knowledge discovery. The current ontology learning systems extracts relevant domain terms and relationships from a corpus of text. As an unsupervised learning method, formal concept analysis is a good tool to extract concepts and implied rules. With formal concept analysis (FCA), one can extract formal concepts and their relationships from a boolean context. The basic relation of the concepts is the hierarchical relation, which is the core component of all ontologies.

Formal concept analysis (FCA) is a clustering technique which concerns about the hierarchical grouped structure of objects, which is a good start point for concept construction [23]. FCA reflects the philosophical understanding of a concept as a unit of thought consisting in two parts: the extent containing all the entities belonging to the concept and the intent which is the collection of all the attributes shared by the entities, based upon Galois connections [2,23,24]. A boolean context  $C$  is defined as a triple  $(G, M, I)$ , where  $G$  is the set of objects or entities and  $M$  the set of properties or attributes. The formal concepts set  $F$  can be derived from this context as follows. For  $A \subseteq G$  and  $B \subseteq M$ , let us define the Galois connections:

$$\begin{aligned} A' &= \{m \in M \mid (\forall g \in A) gIm\}, \\ B' &= \{g \in G \mid (\forall m \in B) gIm\}. \end{aligned}$$

Thus  $A'$  is the set of the attributes that are common to all the entities in  $A$  and  $B'$  is the set of the entities possessing their attributes in  $B$ . These Galois connections

verify all the usual properties of the duality in particular the following relation  $A \subseteq A''$  stands for any subset of the context (consequently  $A' = A'''$ ). Concept is then a pair  $(A, B)$  such that  $A \subseteq G, B \subseteq M, B = A'$ , and  $A = B'$ .  $A$  is the extent of the concept and  $B$  the intent. For concepts  $(A_1, B_1)$  and  $(A_2, B_2)$ , the hierarchical subconcept superconcept relation is formalized by  $(A_1, B_1) \leq (A_2, B_2) \Leftrightarrow A_1 \subseteq A_2 (\Leftrightarrow B_2 \subseteq B_1)$ . The set of all the concepts of the context  $(G, M, I)$  together with this order relation is a complete lattice which is called the concept lattice. Therefore for every set of concepts there exists a unique largest subconcept, the infimum and a unique smallest superconcept, the supremum. Supremum and infimum are respectively given by  $\bigvee_{j \in J} (A_j, B_j) = ((\bigcup_{j \in J} A_j)'', \bigcap_{j \in J} B_j)$  and  $\bigwedge_{j \in J} (A_j, B_j) = (\bigcap_{j \in J} A_j, \bigcup_{j \in J} B_j)''$ .

In semistructured data model, such as Lorel and XML, the data model is a multi-labeled graph, where the data objects are nodes and the attributes, labels [2]. In our running example, a data object is a collection of attributes without blank-value. Thus, it is very easy to identify whether a data object has an attribute. By conceptual scaling, an engineering document can be transformed into a boolean context.

With standard FCA algorithms, the hierarchical grouped structure of engineering data objects, called concept lattice, can be formed. Then the collection of implications (association rules) will be discovered. This will provide useful and applicable knowledge integration for reusing the engineering data.

#### 5. Knowledge Representation

In philosophy, ontology is a theory about the nature of existence. A formal definition of ontology used in this paper bases on [25]: "An ontology is a tuple  $\Omega = (C, R, \psi)$ , whereas  $C$  is a set of concepts,  $R$  a set of relationship names,  $R^H := \{is\_a\} \in R$  the inheritance relation on  $C$ ,  $R^A \in R$  the attribute relation on  $C$  and  $\psi: R \rightarrow \wp(C \times C)$  a function". In short, an ontology is the conceptualization of concepts and their relations.

There are many kinds of representation of ontology. For instance, a typical ontology may include the conceptual entity of the domain, called concept, attributes describing a concept, called property, relationship between concepts, called relation, relation via logic expression, called axiom. Ontology represents the interrelationships between entities or resources, while XML and RDF deal with metadata. An ontology is typically defined as a specification of conceptions, where there are definitions of entities and their relationship with each other, based on semantic nets used in artificial intelligence.



One can discover a domain knowledge through supervised or unsupervised methods. As mentioned earlier, FCA is an unsupervised method to forming abstract concepts and rules. In issues of ontology learning, one can construct an ontology for engineering design automatically. However, the resulted ontologies are often unsatisfiable without human intervention. In this paper, we present an example of knowledge representation for our running example of EDM, shown in Figure 4. In the engineering knowledge domain, there are at least two kinds of entities, persons and documents. The former includes the attributes as name, email and can be divided into deviser, document managers and auditors, etc; the latter includes various attributes of documents, such as visiting time, documentation time, file location, data size, data type, etc. The relationships include hierarchical attributes, etc.

Then we attempt to represent the knowledge in our running example of EDM as axioms of *DLs*. *DLs* are the historical descendants of attempts to formalize semantic networks, which can be addressed by utilizing a subset of first order logic [23]. A *DL*-based system can support modeling of rules for engineering data in a hierarchical fashion. A typical *DLs* knowledge bases consists of a TBox  $\mathcal{CT}$ , an ABox  $\Sigma$  (a set of assertions), and a set of rules  $\tilde{R}$ . The set of rules  $\tilde{R}$  is of the form  $C \Rightarrow D$ , where  $C, D$  are concepts. Given an individual  $a$  (a documented document or person), a new assertion  $D(a)$  will be added into the ABox  $\Sigma$  if  $C(a)$  is already believed to hold. The lower expressivity of description logics, e.g., *AL $\mathcal{EN}$* , limits the ability to define detailed semantics for a domain. However, in return they make several positive tradeoffs, including desirable computability and tractability results [23]. Within the repository application, this loss of expressiveness is acceptable so long as enough design knowledge can be captured to enable necessary operations, such as matching devices against adequately sophisticated searches and comparing mechanisms.

We denote by the set of atomic concepts in  $\Sigma$ ,  $Symbols_T$  [23]. Now let us demonstrate some axiomatic representation with respect to the model shown in Figure 4.

In graph-based presentation of ontology, the atomic concept is often presented as a class. In the ontology of engineering design below, the ABox  $\Sigma$  consists of the assertions with respect to the documented documents and persons, and

$$Symbols_T = \{Person, Document\_Manager, Designer, Auditor, Document\} \quad (1)$$

Besides the inheritance relations, the properties (slots) of concepts could also be described as roles. Thus, the set of atomic roles contains

$$\{hasName, hasEmail, desgin, audit, document, designBy, auditBy, \quad (2)$$

$$documentBy, cite, hasLocation, \quad (3)$$

$$hasType, hasSize, hasDocumentationtime\} \quad (4)$$

$T$  consists of the following axioms

$$Person \sqsubseteq \exists hasName. \top \sqcap \exists hasEmail. \top \quad (5)$$

$$Designer \equiv Person \sqcap \exists desgin. Document \quad (6)$$

$$Auditor \equiv Person \sqcap \exists audit. Document \quad (7)$$

$$Document\_Manager \equiv Person \sqcap \exists document. Document \quad (8)$$

$$Document \sqsubseteq \exists isCited. \top \sqcap \exists hasLocation. \top \sqcap \exists hasType. \top \sqcap \exists hasSize. \top \quad (9)$$

$$\sqcap \exists hasDocumentationtime. \top \quad (10)$$

We provide some rules for the ontology above, where the rules set  $\tilde{R}$  contains

$$\exists desgin. Document \Rightarrow Designer$$

$$\exists audit. Document \Rightarrow Auditor \quad (11)$$

$$\exists document. Document \Rightarrow Document\_Manager \quad (12)$$

$$\exists cite. Document \Rightarrow \exists isCited. Document \quad (13)$$

## 6. Semantic Search Service

After the preliminary ontology construction for engineering data, further steps, e.g., integrating all concepts and relations in distributed environment and evaluating the on-

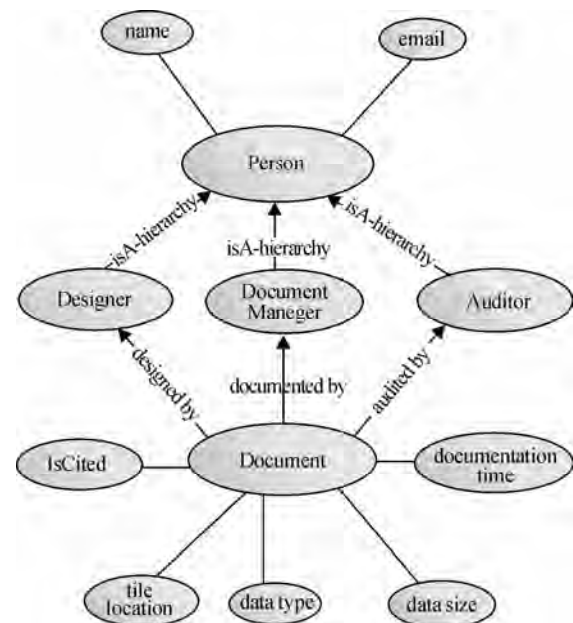


Figure 4. An ontology example of engineering design

tology, should be investigated. As mentioned earlier, the ontology-based semantic services should include constructing domain concepts and their relationships. Furthermore, a sophisticated search and comparing mechanism, which are advanced semantics services, should also be provided. One can demonstrate some possible *DL*-based semantics services as follows:

- 1) Asserting new information about an existing term (individual)
- 2) Recognizing that the updated term is an instance of a class (concept name)
- 3) Firing a rule on the term that is associated with the class
- 4) Propagating information from the updated term to related terms

One can offer inference service to reason with *DLs*-based knowledge. For example, given an individual  $a$  (a documented document or a person), one can find the most specific concepts for  $a$  e.g., the most proper description of  $a$  within the domain. One can also provide other kinds of intelligent services, such knowledge discovery, semantics searching, etc [19].

In *DLs*, given an ontology  $\Sigma$ , a query  $\phi$  could be written in the form of  $C(a)?$ , where  $a$  is an individual and  $C$ , a concept description. It could also be expressed as an evaluation of the consequence relation  $\Sigma \models \phi$ , with the only two answers to that query: either 'yes', or 'no'. For example, in the ontology above, the problem whether the *Designer* are *Auditor* are same for the document *Doc* can be expressed as the entailment problem:

$$\Sigma \models \exists \text{designBy} . (\text{Designer} \sqcap \text{Auditor})(\text{Doc}) \quad (14)$$

## 7. Conclusions and Future Works

As mentioned above, we discuss some issues of ontology-based schema for engineering knowledge systems. Based on the user needs of engineering data management systems, we discuss the issues of knowledge discovery method with formal concept analysis. The knowledge representation of engineering data management systems has also been discussed. Furthermore, we have investigated the domain semantics. The proposed works could be served as major steps for constructing an engineering data knowledge management system.

However, many problems are still open. For example, it is still challenging to meet the needs of providing far more semantics services for creative design processes in engineering data management. In all, creative engineering data management is attractive but challenging topic. As knowledge discovery is currently the major bottleneck of knowledge-based system, one needs more effective and efficient methods to extract relevant domain terms

automatically. Appropriate concepts can meet the requirements of engineering data knowledge management system and describe the domain knowledge properly.

## REFERENCES

- [1] L. Ling, Y. Hu, X. Wang, and C. Li, "An ontology-based method for knowledge integration in a collaborative design environment," *The International Journal of Advanced Manufacturing Technology*, Volume 34: pp. 843–856, 2007.
- [2] S. D. Miller, "A control-theoretic aid to managing the construction phase in incremental software development (extended abstract)," In *Proceedings of the 30th Annual international Computer Software and Applications Conference (Compsac'06)*–Volume 02.
- [3] M. Uschold, M. King, S. Moralee, *et al.*, "The enterprise ontology," *Knowl Eng Rev* 13(1): pp. 31–89, 1998.
- [4] A. M. Collins and M. R. Quillian, "Retrieval time from semantic memory," *Journal of verbal learning and verbal behavior* 8 (2): pp. 240–248, 1969.
- [5] A. M. Collins and M. R. Quillian, "Does category size affect categorization time?" *Journal of verbal learning and verbal behavior* 9 (4): pp. 432–438, 1970.
- [6] Description Logics. Baader, Franz, Horrocks, Ian, Sattler, and Ulrike, Volume, *Handbook on Ontologies in Information Systems of International Handbooks on Information Systems*, chapter I: Ontology Representation and Reasoning, pp. 3–31. Steffen Staab and Rudi Studer, Eds., Springer, 2003.
- [7] Description Logics. Baader, Franz, Horrocks, Ian, Sattler, and Ulrike, Volume, *Handbook on Ontologies in Information Systems of International Handbooks on Information Systems*, chapter I: Ontology Representation and Reasoning, pp. 3–31. Steffen Staab and Rudi Studer, Eds., Springer, 2003.
- [8] T. Berners-Lee, "Semantic Web-XML2000,".
- [9] T. Berners-Lee, J. Handler, and O. Lassila, "The semantic web," *Scientific American*, Vol. 184, pp. 34–43, 2001.
- [10] P. F. Patel-Schneider, P. Hayes, and I. Horrocks, "OWL web ontology language semantics and abstract syntax, W3C Recommendation 10 February 2004", <http://www.w3.org/TR/owl-semantics/>, 2004.
- [11] H. G. Windream (2006), *Managing documents by windream*, Available at <http://www.windream.com/>. Accessed October 2006.
- [12] EMC2, (2006). *Documentum*. Available at <http://software.emc.com/products/product-family/documentum-family.htm>. Accessed September 2006
- [13] S. Abiteboul and D. Quass, "The lorel query language for semistructured data," *International Journal on Digital Libraries*, pp. 68–88, 1997.
- [14] D. Beech, A. Malhotra, and M. Rys, "A formal data model and algebra for XML, W3C. XML Query working group note, September 1999.
- [15] M. Fernandez, J. Simeon, D. Suciu, and P. Wadler, "A data model and algebra for XML query," 1999.

- <http://www.cs.belllabs.com/wadler/topics/xml.html#algebra>.
- [16] B. F. Cooper, N. Sample, M. J. Franklin, G. R. Hjaltason, and M. Shadmon, "Fast index for semistructured data," *Proceedings of the 27th VLDB Conference*, Roma, Italy, 2001.
  - [17] L. Cardelli, "Describing semistructured data," *SIGMOD Record*, Vol. 30, No. 4, December 2001.
  - [18] Y. Papakonstantinou, "Enhancing semistructured data mediators with document type definitions," *ACM SIGMOD ICDE*, pp. 136–145, 1999.
  - [19] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and Peter Patel-Schneider, editors. *The description logic handbook: Theory, implementation, and applications*, Cambridge University Press, 2003.
  - [20] J. Koppena and W. C. Regli, "Functional modeling of engineering designs for the semantic web," *IEEE Data Engineering Bulletin*, 26(4), pp. 55–61, 2003.
  - [21] Y. Kitamura and R. Mizoguchi, "Ontology-based description of functional design knowledge and its use in a functional way server," *Expert Systems with Applications*, 24(2), pp. 153–166, 2003.
  - [22] Angele, Staab, R. Studer, and D. Wenke, "OntoEdit: Collaborative ontology engineering for the semantic web," *International Semantic Web Conference 2002 (ISWC 2002)*, Sardinia, Italia, 2003.
  - [23] B. Ganter and R. Wille, "Formal concept analysis: mathematical foundations," Springer, Heidelberg, 1999.
  - [24] A. Hotho and G. Stumme, "Conceptual clustering of text clusters," In *Proceedings of FGML Workshop*, pp. 37–45, Special Interest Group of German Informatics Society (FGML-Fachgruppe Maschinelles Lernen der GI e.V.), 2002.
  - [25] Faatz, Andreas, Steinmetz, and Ralf, "Ontology enrichment evaluation," *Lecture Notes in Computer Science*. Springer-Verlag. Vol. 3257. pp. 497–498, 2004.

# Intertemporal Pricing and Allotment of Sea-cargo Capacity under Reference Effect<sup>\*</sup>

Xiangzhi Bu<sup>1</sup>, Lei Xu<sup>2</sup> & Li Su<sup>3</sup>

<sup>1</sup>School of Business, Shantou University, Shantou, China, <sup>2</sup>School of Business, Shantou University, Shantou, China, <sup>3</sup>English Language Center, Shantou University, Shantou, China  
Email: xzpu@stu.edu.cn, lxx3@stu.edu.cn, lisu@stu.edu.cn

Received August 16<sup>th</sup>, 2008; revised December 8<sup>th</sup>, 2008; accepted December 15<sup>th</sup>, 2008.

## ABSTRACT

*Reference effect influences the shippers' purchase behavior to some extent, and it makes shippers' behavior more mysterious. Furthermore, it leads to complex to coordinate procurement and pricing of forwarder and carrier. Firstly, we introduce the reference effect into the sea-cargo service supply chain composed with carriers and forwarders, and build the game-theoretical model to decide the optimal contract price and order quantity. Considering the situation where the market demand is affected by current price and reference price, we adopt dynamic programming and classical variation method to analyze the dynamic equilibrium of the game under reference effect with symmetric information or not, respectively. Finally, this paper analyzes the reference effect to price, procurement and profit of carrier with digital analysis.*

**Keywords:** reference effect, anchoring effect, sea-cargo service chain, dynamic pricing, revenue management

## 1. Introduction

The marine environment has two types of markets generally, the contract market and the spot one in reality [1,2]. In order to lower the risk of capacity allocation and price, carriers often choose to sale advance in contract market that is to sell most capacity to the forwarders and other big shippers. And then, in the second stage, the remaining space is sold by carrier to direct-ship shippers on ad-hoc or free-sale basis. As to the forwarders, in the first stage the forwarders purchase large amount of cargo capacity, which qualifies them for a volume discount. The discount may depend upon the size of allotment as well as the actual volume tendered by the forwarder [3]. In the second stage, the forwarder drums up downstream shippers, forwarders can offer more services and often better prices to downstream shippers in comparison to the carriers' standard tariff. In this paper, we just focus on coordination between the carrier and the forwarder in the contract market.

In the downstream freight market, due to the repeated purchase behavior, shippers will develop reference price, through observing prices of other channels and forwarder's past prices. Comparing current price and reference price, thereafter, shippers make their capacity purchasing decision. This theory is similar to the Prospect Theory of Kahneman and Tversky [4]. During purchasing, they buy less and then turn to other carriers or forwarders

if the current price is above reference price, or they do more on site. Throughout the paper, we just analyze shippers' behavior as a whole, not the individual shipper's purchasing behavior. In this way, the downstream shippers' purchasing decision is not only based on the cost-plus price [1], but the price expectations. In such complex situations, the carrier and the forwarder must take into account the shippers' price reference effect before they decide the contract price, procurement level and the resale price. And carrier and forwarder could share information and communication to quantify the reference effect to make optimal ordering and pricing decision.

In the paper, we adopt the common supply chain contract-wholesale contract. In a channel context, we observe in some cases that channel transactions are "governed by simple contracts defined only by a per unit wholesale price" [5]. As noted some time ago by [6], incentive contracts in the real world frequently take simpler forms than what theory often predicts. This can also happen because firms have little to lose using a simpler contract [7]. We build a von-Stackelberg model of the carrier and the forwarder where the carrier is the game leader, and the forwarder is the game follower. In which the carrier firstly announces the contract prices according to the ordering quantity of forwarder, and then the forwarder decides the contract quantity. Here, both the carrier and the forwarder act strategically, and the shippers are only sensitive in price. The mathematical representation of the interaction between carrier and forwarder is based on a stylized

<sup>\*</sup> The research is supported in part by the National Science Foundation of China under Grant NO. NSFC 70701022, the Natural Science Foundation of Guangdong Province under Grant NO. 07301452 and the China Postdoctoral Science Foundation under Grant NO. 20060390847.

model in which certain real-life details have been either omitted or simplified to maintain tractability.

The remainder of this paper is organized as follows. In Section 2, we begin with a review of the relevant literatures. In Section 3 we present the problem and also give the definitions of symbols and parameters. We propose a formal price reference model in Section 4. In Section 5, we firstly give an analysis on the decision making process when price reference effect is common knowledge in part 5.1; when price reference effect is asymmetrical information to the carrier, see part 5.2. We then continue in Section 6 with a presentation of our numerical studies and report on various insights that can be drawn from digital simulation. Section 7 concludes our paper.

## 2. Literature Review

Extensive literature has evolved around the coordination between carrier and forwarder, yet relatively attention has been directed to the problem of intertemporal pricing in multi-stages. [1,2] considered a situation where contract market and spot one coexist; asset provider and intermediary can make coordination through option contract, and the overbooking is also considered. Gupta [8] analyzed a flexible contract where carrier has priority to change contract parameters later to adjust contract quantity. Then, Amaruchkul, Cooper and Gupta [9] analyzed capacity contracts under asymmetric information, and proposed optimal contract. However, these researches mostly focus on a single-stage problem. In real world, carrier and forwarder coordinate in a multi stages, more.

As to the dynamic pricing of finite inventory, the relevant literature has mostly ignored the reference effect. Gallego and van Ryzin [10] and Bitran and Mondschein [11] analyze a continuous and discrete time pricing schemes, in environment where the distribution of consumer valuation does not change over time. Lazer [12] and Aviv and Pazgal [13] discuss dynamic pricing of fashion goods where demand is uncertain. Next, Aviv and Pazgal [14] analyze optimal pricing of seasonal products where the consumer is forwarder-looking (strategic) and distinguish two classes of pricing strategies. For further information on the above body of research, we refer the reader to Popescu and Wu [15] and Bu Xiang-zhi *et al.* [16].

There are a lot of papers study on the effects of reference price to market demand. Helson's [17] Adaptation Level Theory pointed that people measures the current result with his past activities as standards; Consumer's price perception depends on the difference between the interior reference price and exterior reference price. Sorger [18] studies local stability of joint dynamic pricing and advertising policies under reference price effects. Fibich [19] investigates the dynamic pricing problem with linear reference effects. The research on dynamic pricing under reference price effects is also in Popescu *et al.* [15], during which, both consumers of risk neutral and

risk preference are concluded. Kapalle *et al.* [20] bring quality into the reference effect and analyzed the condition under which reference effect of price and quality are both existed. Oliver and Winer [20], Zeithaml [21] consider the reference effect to be the key factor which affects market demand. But all the papers above is in the area of marketing, it doesn't refer to the coordination between retailers and manufacturers or channel distributor's problems of production quantity and inventory in supply chain system.

## 3. Problem Description

The pricing and ordering of repeat-purchase capacity are what we studied in this paper. With the repeated purchase behavior, shippers will develop a reference price  $p_r$  through observing price of other channels and forwarder's past prices. Then, comparing the current price and the reference price, shippers would like to purchase less and turn to other channels to procure if the current price is higher than the reference price; otherwise shippers would purchase more on site [4]. This price reference effect makes the demand much more complex.

Because of the double impact of current price and price reference, the forwarder faces a more complex cargo market. In that case, the price reference effect has to be considered during pricing decision, based on which, forwarder makes the resale prices and quantity at each period to pursue the maximum profit. And furthermore, the market price also affects the reference price. Thus, both current price and reference price change interactively. Due to the double impact of current price and reference price, it becomes difficult that the coordination between carrier and forwarder, and how the forwarder makes the resale prices as well.

In this paper, a finite-period model is involved in a steady economic circumstance, that is to say the contract prices difference and resale prices difference are not attributable to economical conditions change, such as cost changes and competition, but strategically pricing. Moreover, cargo that shippers provide lies in the same type, avoiding the prices difference from cargo type difference. And the ordering tactics is not referred specifically. We assume that the carrier sells capacity on the same cargo market in  $N$  periods, and the contract between the carrier and the forwarder is the classical discount price contract used widely in real economic activities [23]. It means carrier announces a contract price  $k(t)$  according to forwarder's books level. We choose the linear rebate through all the paper [24]. And, the unit cost is constant as  $c$  [25].

The decision process is as following:

Firstly, the carrier announces its discount price contract according to forwarder's future booking decision. Then, the forwarder decides its resale price  $p(t)$  in period  $t$ , based on the discount price  $k(t)$  and the ship-

pers' reference price to achieve the dynamic long run maximum profit. And, procure quantity  $k(t)$  in each single period is determined based on a take-it-or-leave-it basis.

During this paper, we mainly study the problems described as follows: how forwarder determines the resale prices and procurement quantity with the double impact of current price and reference price, and how carrier makes sea-cargo capacity contract under reference effect with symmetric information or not. The mathematical representation of the interaction between the carrier and the forwarder is based on a stylized model in which certain real-life details have been omitted or simplified to maintain tractability. Moreover, we assume rational and risk-neutral decision makers.

The basic parameters of the model are as follows:

$\gamma$  the memory effect of shippers to past price

$c$  the constant unit cost

$k(t)$  the discount price made by the carrier in period  $t$

$d(t)$  the procurement (demand) in period  $t$

$p(t)$  the sales price made by the forwarder in period  $t$

$p_r(t)$  the reference price developed by shippers in period  $t$

$\pi_b(t)$  the forwarder's revenue in period  $t$

$\pi_s(t)$  the carrier's revenue in period  $t$

All the parameters above satisfy  $p(t) \geq k(t) \geq c$ , while  $k(t) = c$  is the situation that the decision is made by the carrier and the forwarder jointly. It is a classical pricing problem of marketing on vertical integration of the carrier and the forwarder [15,16], we don't study on it more in the following.

#### 4. Reference Effect Model

Recently, many empirical studies had tested the price reference effect and set up the reference effect equations, of which important empirical studies were based on prospect theory. Prospect theory considers that the difference between price and reference price affects market demand, and the effect is asymmetrical. Kahneman and Tversky [4] estimated that when the loss and the gain are almost equal, we often emphasize the loss more, and even two times as much as the gain. As to the same variation, consumers often pay attention to that when price is above reference price [26]. Kalyanaram and Winer [27] tested the strong effect of prospect theory under reference effect. The derivation of mathematic equation is following:

Assumption: at period  $t$ , the market demand is  $d(t)$ , as the function of price  $p(t)$  and reference price  $p_r(t)$ .

For simplicity, we adopt the linear demand function [25], as follows:

$$d(t) = a - bp(t) + m\{p_r(t) - p(t)\} \quad (1)$$

where,  $a, b > 0$  and are all positive constants,

$$m = \begin{cases} \theta_1, & p_r(t) > p(t) \\ \theta_2, & p(t) > p_r(t) \end{cases}, \text{ and } \theta_2 > \theta_1 > 0.$$

The penalty for each unit price above reference price is  $\theta_2$ , and the gain for each unit price below reference price is  $\theta_1$ .

Following the past literatures on the formation of reference prices, we model it as a weighted average of past prices [28], that is:

$$p_r(t+1) = \gamma(t)p(t) + (1-\gamma(t))p_r(t) \quad (2)$$

Here, we assume shippers are homogeneous and have the same memory effect  $\gamma$ . Where,  $\gamma$  is shipper's memory effect to historic price, and the bigger  $\gamma$  is, the stronger the memory effect is. In particular,  $\gamma=1$  is ideal state, reference price is equal to the initial reference price  $p_r(t) = p_0(t)$ .

### 5. Coordination Between Carrier and Forwarder

#### 5.1. Decision under Symmetric Information

Under this situation, carrier and forwarder are attached to different enterprises, and make their decision respectively. Carrier decides the contract prices, and forwarder set resale price and order quantity. The Von-Stackelberg game process between carrier and forwarder is:

At period  $t$ , the decision of carrier is:

$$\pi_s(t) = [k(t) - c] \times d(t) \quad (3)$$

Based on that, we adopt dynamic programming method to find the general property of carrier's objective function. Assume  $V_t(p(t), p_r(t))$  is total profit function from period 0 to period  $t$ .

So, the Bellman equation is [29]

$$\begin{aligned} V_t^s(p(t), p_r(t)) &= (k(t) - c)d(t) + V_{t+1}^s(p(t+1), p_r(t+1)) \\ s.t. \quad p_r(t+1) &= \gamma(t)p_r(t) + (1-\gamma(t))p(t) \end{aligned} \quad (4)$$

At period  $t$ , the decision of forwarder is:

$$\pi_b(t) = [p(t) - k(t)] \times d(t) \quad (5)$$

Then, we adopt dynamic programming method to find the general property of carrier's objective function. Assuming  $V_t(p(t), p_r(t))$  is total profit function from period  $N-1$  to period  $t$ .



So, the Bellman equation is [29]

$$\begin{aligned} V_t^s(p(t), p_r(t)) &= (k(t) - c)d(t) + V_{t+1}^s(p(t+1), p_r(t+1)) \\ \text{s.t. } p_r(t+1) &= \gamma(t)p_r(t) + (1 - \gamma(t))p(t) \end{aligned} \quad (6)$$

The forwarder's decision and state are separable, so the long run objective function can be transformed into:

$$\begin{aligned} \pi &= \sum_{t=0}^{N-1} \pi(t) = \sum_{t=0}^{N-1} [(p(t) - k(t))d(t)] \\ \text{s.t. } p_r(t+1) &= \gamma(t)p_r(t) + (1 - \gamma(t))p(t) \\ p_r(0) &= p_{r0} \end{aligned} \quad (7)$$

Then, the objective function of the forwarder is transformed into optimal control problem about single state variable  $p_r(t)$  and single decision variable  $p(t)$ , which we use maximum principle [30] to solve the discrete-time optimal control problem. And based on the analysis of the iterative multi-stage problem, we decompose the multi-stage dynamic coordination problem into single-stage coordination problems with classical variation method, and based on that analyze iterative multi-stage problem.

**Theorem 1:** single-stage market price  $p(t)$  for forwarder is pathwise increasing in the contract price  $k(t)$  and the reference price  $p_r(t)$ , decreasing in reference effect  $\gamma(t)$ .

*Proof:*

Constructing Hamilton Function:

$$\begin{aligned} H(p_r(t), \lambda(t+1), p(t), t) &= [p(t) - k(t)][a - bp(t) + m(p_r(t) - p(t))] \\ &\quad + \lambda(t+1)[\gamma(t)p_r(t) + (1 - \gamma(t))p(t)] \end{aligned} \quad (8)$$

$$\frac{\partial H(p_r(t), \lambda(t+1), p(t), t)}{\partial p(t)} = 0 \quad (p(t) > k(t) > c),$$

then

$$\begin{aligned} [a - 2(b+m)p(t) + mp_r(t) + (b+m)k(t)] \\ + \lambda(t+1)(1 - \gamma(t)) = 0 \end{aligned}$$

Solving the canonical equations:

$$p_r(t+1) = \gamma(t)p_r(t) + (1 - \gamma(t))p(t), \quad p_r(0) = p_{r0}$$

$$\begin{aligned} \lambda(t) &= \frac{\partial H(p_r(t), \lambda(t+1), p(t), t)}{\partial p_r(t)}, \\ &= (p(t) - k(t))m + \lambda(t+1)\gamma(t) \end{aligned}$$

$$\lambda(N) = 0$$

We obtain,

$$p(t) = \frac{a + (b+m)k(t) + mp_r(t) + \lambda(t+1)(1 - \gamma(t))}{2(b+m)}$$

By the first order conditions,  $\frac{\partial p(t)}{\partial k(t)} = \frac{1}{2} > 0$ ,

$$\frac{\partial p(t)}{\partial p_r(t)} = \frac{m}{2(b+m)} > 0, \quad \frac{\partial p(t)}{\partial \gamma(t)} = -\lambda(t+1) < 0.$$

End.

**Theorem 2:** single-stage profit function  $\pi^s(t)$  for the carrier is concave in the contract price  $k(t)$ .

*Proof:*

Solving the first condition of  $\pi^s(t)$  to  $k(t)$ , we obtain

$$\begin{aligned} \frac{\partial \pi^s(t)}{\partial k(t)} &= a + mp_r(t) + (b+m)c \\ &\quad - \lambda(t+1)(1 - \gamma(t)) - 2(b+m)k(t) \end{aligned}$$

The second condition of  $\pi^s(t)$  to  $k(t)$  is  $\frac{\partial^2 \pi^s(t)}{\partial [k(t)]^2} = -2(b+m)$ .

Because of  $b, m > 0$ , we get

$$\frac{\partial^2 \pi^s(t)}{\partial [k(t)]^2} = -2(b+m) < 0.$$

Then Theorem 2 is proved.

According to Theorem 2, then we can get value of  $k(t)$  when  $\pi^s(t)$  get the maximum.

$$\text{That is } k(t) = \frac{a + (b+m)c + mp_r(t) - \lambda(t+1)(1 - \gamma(t))}{2(b+m)}.$$

Substituting  $k(t)$  into  $p(t)$ , we get

$$p(t) = \frac{3a + (b+m)c + 3mp_r(t) + \lambda(t+1)(1 - \gamma(t))}{4(b+m)}.$$

In particular, when  $\gamma(t) = 0$ , there is no reference effect, the market demand function is  $d(t) = a - bp(t)$ , it is definite, through solving Von Stackelberg model, we obtain  $k(t) = \frac{a+bc}{2b}$ ,  $p(t) = \frac{3a+bc}{4b}$ ,  $d(t) = \frac{1}{4}(a-bc)$ .

When  $\gamma(t) = 1$ , it is complete price reference effect. The demand function is transformed into  $d(t) = a - mp_r(t) - (b+m)p(t)$ , and  $p_r(t) = p_r(0)$ . It is also definite, through solving Von Stackelberg model, we obtain

$$k(t) = \frac{a + p_r(0) + (b+m)c}{2(b+m)},$$

$$p(t) = \frac{3a + 3mp_r(0) + (b+m)c}{4(b+m)},$$

$$d(t) = \frac{1}{4}(a + mp_r(0) - (b+m)c).$$

As to the specified coordination problem between carrier and forwarder, we will give a further analysis in the Section 6.1.

## 5.2. Decision under Asymmetric Information

This section analyzes a two-level supply chain composed with single carrier and single forwarder. Where, the carrier as the game leader, and the forwarder as the follower in the Von Stackelberg model, the forwarder faces a multi-stage market demand influenced by reference price. Due to near market, the forwarder owns much more market information than the carrier. That is to say, when the reference effect is asymmetric information, the forwarder knows about the exact state. However, because of the asymmetric information about reference effect, the carrier knows little. Under this condition, constant discount contract can't maximize the profit of the carrier any longer.

Assume  $\gamma(t)$  is reference effect, and  $\gamma(t)$  satisfies  $0 \leq \gamma(t) \leq 1$ . Here, the probability of  $\gamma(t) = \gamma_i(t)$ , ( $i = 1, 2, \dots, n$ ) is  $p(\theta = \theta_i) = p_i$ , and it is common knowledge (except reference effect, cost of the carrier and both profit functions are all common knowledge). Under this condition, the decision process of the forwarder is: firstly, the forwarder determines the retail price according to the reference effect, then inquires the discount price, and determines his ordering quantity.

At period  $t$ , the decision of the forwarder is:  $\pi_b(t) = [p(t) - k(t)] \times d(t)$

Then, we search for the general property of forwarder's objective function using dynamic programming method. Assume  $V_t(p(t), p_r(t))$  is the total profits function from period 0 to period  $t$ . So, the Bellman equation is [29]:

$$\begin{aligned} V_t^b(p(t), p_r(t)) &= (p(t) - k(t))d(t) + V_{t+1}^b \\ &\quad (t+1, p(t+1), p_r(t+1)) \\ \text{s.t. } p_r(t+1) &= \gamma(t)p_r(t) + (1 - \gamma(t))p(t) \end{aligned} \quad (9)$$

According to the optimization principle of dynamic programming method, sub strategy of an optimal strategy is optimal. According to the analysis of Section 4.1, at period  $t$ , the optimal pricing strategy of forwarder is

$$p(t) = \frac{3a + (b+m)c + 3mp_r(t) + \lambda(t+1)(1 - \gamma(t))}{4(b+m)}.$$

The decision process of the carrier is: before forwarder makes the order quantity, the carrier never knows the exact reference effect, only knows about the probability  $p(\theta = \theta_i) = p_i$ . According to Harsanyi transformation principle, we assume that there exist  $n$  kinds of forwarders; every forwarder has one kind of reference effect. Thus, the carrier can't maximize his profit by setting only one discount price. Then, the carrier announces his discount contract in early of each period: set  $d_i(t) \leq d_j(t), i \leq j$ , corresponding discount price is  $k_i(t) \geq k_j(t)$ , and  $c < k_j(t) \leq k_i(t)$ . The carrier maximizes his profit through setting up discount contract. And, the order quantity of the forwarder is continuous, and only the order quantity is equal to or bigger than  $d_i(t)$ , the discount price is equal to  $k_i(t)$ . According to Bayesian and Revelation principle, at third stage the forwarder chooses optimal order quantity and discount price according to his type, for example  $\{q_i(t), k_i(t)\}$ , and this signal releases the forwarder's type.

**Theorem 3:** the combination of discount price and order quantity is  $\{[q_i(t), q_{i+1}(t)], k_i(t)\}$ .

*Proof:*

According to the result of Theorem 1, the market retailing price  $\gamma_i(t)$  is corresponding to  $p_i(t) = \frac{a + (b+m)k_i(t) + mp_r(t) + \lambda(t+1)(1 - \gamma_i(t))}{2(b+m)}$ . And according to Theorem 2, the discount price of the carrier is  $k_i(t) = \frac{a + (b+m)c + mp_r(t) - \lambda(t+1)(1 - \gamma(t))}{2(b+m)}$ .

Based on the formula of  $p_i(t)$  and  $k_i(t)$ , we can obtain the order quantity  $q_i(t) = \frac{1}{4} \times [a + mp_r(t) - (b+m)c - \lambda(t+1)(1 - \gamma_i(t))]$  at period  $t$  when the reference effect is  $\gamma_i(t)$ ; And correspondingly, when the reference effect is  $\gamma_{i+1}(t)$ , the order quantity is  $q_{i+1}(t) = \frac{1}{4} \times [a + mp_r(t) - (b+m)c - \lambda(t+1)(1 - \gamma_{i+1}(t))]$ . So when the forwarder's type is  $\gamma_i(t)$ , the carrier set up the combination of discount price and order quantity is  $\{[q_i(t), q_{i+1}(t)], k_i(t)\}$ .

### 5.2.1. Mechanism Design

Because the reference effect is not common knowledge to the carrier, we can take the carrier's decision as a classical mechanism design problem. And based on the revelation principle [31], the carrier can make  $n$  kinds of combination  $\{[q_i(t), q_{i+1}(t)], k_i(t)\}$ . Each combination matches along with one type of forwarder. Assuming that  $Q_i(t)$  ( $q_i(t) \leq Q_i(t) \leq q_{i+1}(t)$ ) and  $k_i(t)$  is a combina-

tion of discount price and order quantity to type  $i$  forwarder, so the expected profit of the carrier is (10):

$$\begin{aligned}
 & V_t^s(k(t), p(t)) = p_i(k_i(t) - c)d_i(t) \\
 & + V_{t+1}^s(k(t+1), p(t+1)) \\
 \text{s.t.} \quad & \max \pi_i(Q_i(t), k_i(t)) \geq 0, i = 1, 2, \dots, n, \dots (IR_i) \\
 & \pi_i(Q_i(t), k_i(t)) \geq \pi_i(Q_j(t), k_j(t)), i \\
 & = 1, 2, \dots, n, j \neq i, \dots (IC_{i,j}) \\
 & Q_i(t) \geq 0, k_i(t) \geq 0, i = 1, 2, \dots, n
 \end{aligned} \tag{10}$$

Where, the first item is Individual Rational constraint ( $IR_i$ ) which ensures the forwarder to accept the discount contract; the second item is Incentive Compatibility constraint ( $IC_{i,j}$ ) which shows type  $i$  forwarder chooses the combination {discount price, order quantity} corresponding his type.

### 5.2.2. Analysis of Mechanism Design

According to the maximum principle, the forwarder chooses the optimal order quantity, and in the third stage, the forwarder releases his type to the carrier through signal about the forwarder's type.

The segmentation of order quantity that the carrier makes is corresponding to the forwarder's type, that is:

$$q_i(t) = \frac{1}{4}[a + mp_r(t) - (b+m)c - \lambda(t+1)(1 - \gamma_i(t))].$$

**Theorem 4:** the optimal order quantity of the forwarder is the initial quantity  $q_i(t)$  in the interval  $[q_i(t), q_{i+1}(t)]$ . That is:

$$q_i(t) = \frac{1}{4} \times [a + mp_r(t) - (b+m)c - \lambda(t+1)(1 - \gamma_i(t))]$$

*Proof:*

according to Theorem 1 and the principle of Mechanism Design in section 4.2.1, Theorem 4 can be proved simply.

Based on Theorem 4, we transform formula (10) into (11):

$$\begin{aligned}
 & V_t^s(k(t), p(t)) = p_i(k_i(t) - c)d_i(t) + V_{t+1}^s(k(t+1), p(t+1)) \\
 \text{s.t.} \quad & \max \pi_i(q_i(t), k_i(t)) \geq 0, i = 1, 2, \dots, n, \dots (IR_i) \\
 & \pi_i(q_i(t), k_i(t)) \geq \pi_i(q_j(t), k_j(t)), i \\
 & = 1, 2, \dots, n, j \neq i, \dots (IC_{i,j}) \\
 & q_i(t) \geq 0, k_i(t) \geq 0, i = 1, 2, \dots, n
 \end{aligned} \tag{11}$$

**Theorem 5:** 1) the sufficient and necessary condition to solve problem 5.2.2 is: for the given combination  $\{q_i(t), k_i(t)\}$ , problem 5.2.2 must satisfy Individual Rational Constraint  $IR_i$ ; the feasible solution

to  $IC_{i,i-1}, i = 2, 3, \dots, n$  and  $IC_{i,i+1}, i = 1, 2, \dots, n-1$  is equal.

2) The sufficient condition to solve problem 5.2.2 is: for the given condition  $\{q_i(t), k_i(t)\}$ , there must exist feasible solution of 5.2.2 satisfies  $q_i(t) = Q_i(k_i(t))$ , and satisfies Individual Rational Constraint  $IR_i$ , Incentive Compatibility Constraint  $IC_{i,i-1}, i = 2, 3, \dots, n$  and  $IC_{i,i+1}, i = 1, 2, \dots, n-1$ .

As to the proof of Theorem 5, [31,32] have given complete proof.

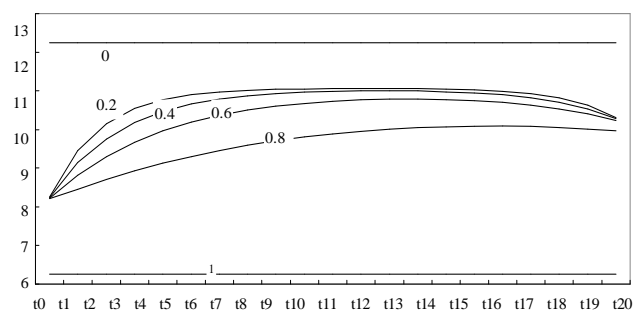
## 6. Case Simulation and Discussion

### 6.1. Digital Simulation under Symmetric Information

The above sections give the property of single-period profit function. However, as to multi-period problem we can't give the pricing trajectory of both the carrier and the forwarder directly. Based on the analysis of Section 5.1, we discuss the optimal pricing trajectory of multi-period discount contract, and the pricing decision of the carrier through case simulation. We take the following parameters for example:  $\gamma(t) = 0$ ,  $\gamma(t) = 0.2$ ,  $\gamma(t) = 0.4$ ,  $\gamma(t) = 0.6$ ,  $\gamma(t) = 0.8$ ,  $\gamma(t) = 1$ ,  $a = 30$ ,  $b = 2$ ,  $m = 3(\theta_1 = 4, \theta_2 = 2)$ ,  $c = 4$ ,  $p_r(0) = 5$ . Because shipper goods, in particular the brand commodities have different procurement frequency. We choose 22 periods as rolling horizon windows. The trajectory of discount price, retailing price and profit under reference effect is following:

(Notes: the Lagrange multipliers are different to different time windows, and then the results of data analysis are different.)

Figure 1 gives the forwarder's optimal pricing trajectory about six kinds of reference effect. When the reference effect  $\gamma(t) = 1$ , the forwarder chooses the lowest price and holds on for the whole time windows. When the reference effect  $\gamma(t) = 0$ , the forwarder makes the highest price and also holds on. Under other reference



**Figure 1.** The forwarder's resale pricing trajectory curve

effects, the initial prices are equal, and from the second period, the forwarder increases the retailing price, the prices converge to the same at last. In some sense, the function of the retailing pricing can be taken as concave of period  $t$ . The forwarder raises the shippers' reference price through increasing retailing price in the early periods, and then lows retailing price to lifting the demand quantity dramatically, in particular in the last period. So that the forwarder could rake in exorbitant profits. Figure 3 also gives the same interpretations about the order quantity, the order quantity of the last period reaches to the maximum level. And the same Lagrange multiplier  $\lambda(22)=0$  leads to the convergence of the last resale price. At the same period the resale price is decreasing in reference effect  $\gamma$ , it is in accord with *Theorem 1*. As Figure 2 and Figure 3 show, When the reference effect  $\gamma(t)=1$ , the carrier gets lowest profits.

As shown in Figure 2, when the reference effect is  $\gamma(t)=1$ , the optimal pricing strategy of the carrier is constant in the whole time windows, and in this situation the carrier takes the lowest price. In combination with Figure 3, we can obtain that the carrier gets the lowest profit. When the reference effect is  $\gamma(t)=0$ , the optimal pricing strategy of the carrier is also constant in the whole time windows where the carrier takes the highest price and the highest profit. Under other situations, the discount prices and corresponding profits of the carrier are between

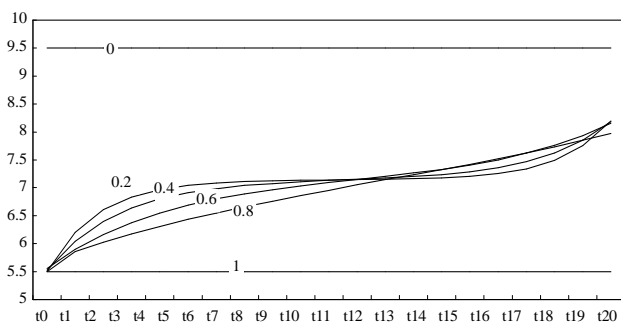


Figure 2. The carrier's contract pricing trajectory curve

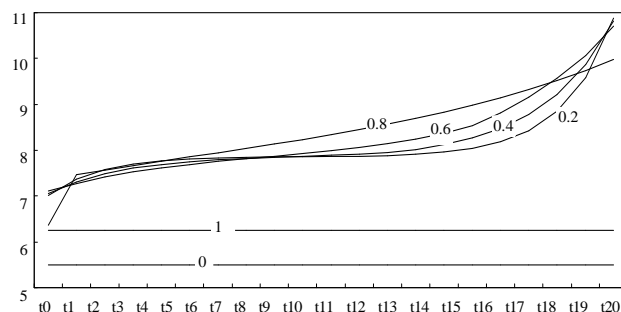


Figure 3. The curve of the forwarder's procurement level

Notes: from Figure 6 to Figure 10, the length of curves represents the rolling time windows. Long curve shows long decision period.

upper boundary value of  $\gamma=1$  and lower boundary value of  $\gamma=0$ .

According to analysis of dynamic von-Stackelberg model, the carrier gets to know that the forwarder will lower the retailing price and by raising the demand quantity dramatically to "seize the last opportunity". And furthermore, the forwarder has made much preparation. Thus, as a believable signal, the carrier will make high price, and only this discount price is not higher than the retailing price, the forwarder will accept the contract (without ordering cost and purchasing cost).

## 6.2. Digital Simulation under Asymmetric Information

According to the principle of mechanism design, we take the following parameters as examples:  $\gamma(t)=0$ ,  $\gamma(t)=0.2$ ,  $\gamma(t)=0.4$ ,  $\gamma(t)=0.6$ ,  $\gamma(t)=0.8$ ,  $\gamma(t)=1$ . And based on these situations, we discuss the combinations of the discount price and order quantity, and tested the profits of the carrier under different reference effect. Other conditions are the same as section 4.1.1.

Based on the result of Figure 4, when the reference effect is asymmetric information, the decision of the forwarder is not affected; the decision process is the same as the former. Thus, we don't repeat it here.

According to the principle of mechanism design, the carrier can maximize his profits through making multi-combination of discount prices and order quantities, and

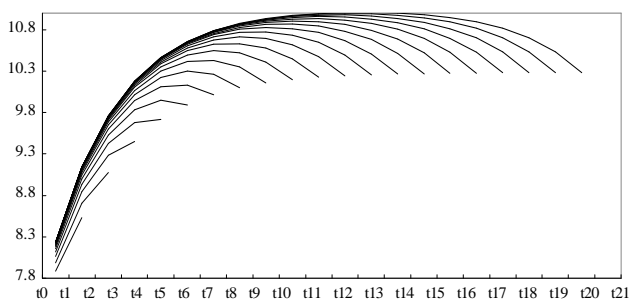


Figure 4. Forwarder's resale prices in different rolling time windows

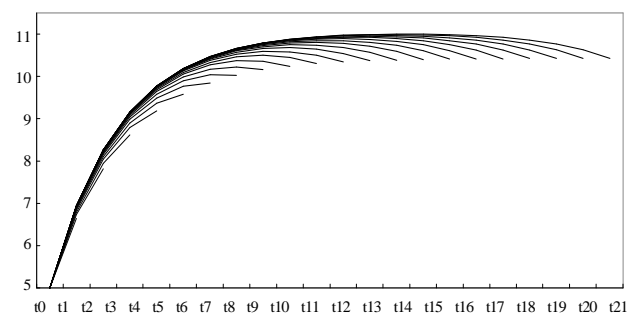


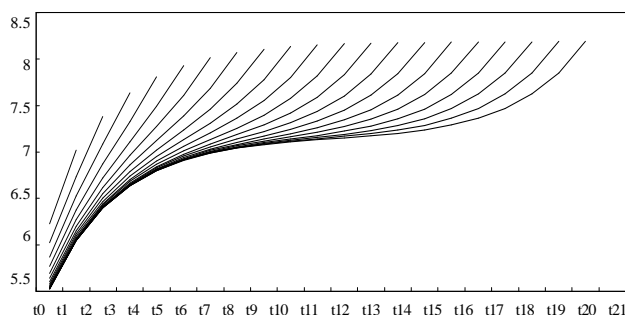
Figure 5. Reference prices in different rolling time windows

with the help of the mechanism, the carrier will know the forwarder's type. That is when the forwarder chooses the relevant combination; he releases real reference effect unconsciously. Thus, the carrier can realize the same profit as the situation where the reference effect is symmetrical information.

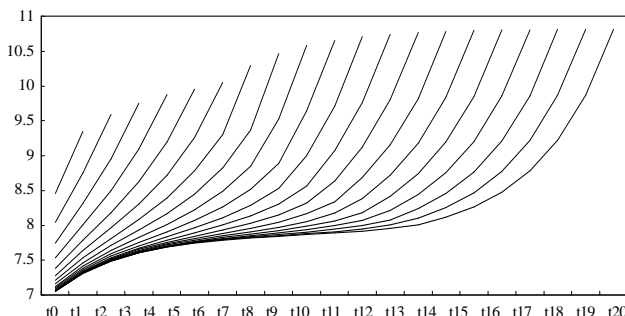
### 6.3. Different Rolling Time Windows Analysis

For the maximum profit, the forwarder decides the optimal resale prices and the optimal procurement quantity; Moreover, he must take into account the price reference effect of its downward market. Due to *Theorem 1*, single-stage resale price  $p(t)$  is pathwise increasing in the reference price  $p_r(t)$ , and decreasing in reference effect  $\gamma(t)$ . However, it's only a single-period optimal decision. As a multi-period decision problem, we should show some managerial suggestions about a long-run pricing decision, in additionally, under price reference effect.

As the simulation result of Figure 4, it shows that the multi-period resale prices are concave in  $t$ . It is different from the result of Popescu and Wu [20], which proved that optimal pricing policies induce a perception of monotonic prices. In our paper, we analyze the resale pricing policies, but also the optimal procurement decision. Theoretically, so long as the forwarder knows about his decision period or about approximate one, he can make the optimal resale prices to maximize his long-run profit.



**Figure 6. Carrier's contract prices in different rolling time windows**



**Figure 7. Forwarder's procurement level in different rolling time windows**

The reference price  $p_r(t)$  is exponential smoothing in resale price  $p(t)$ . And the exponent represents the reference effect of the shippers. We take  $\gamma(t)=0.4$  for example. Based on Figure 4 and Figure 5, it shows the long-run relation between resale price and reference price. As Winer [28] tested,

$$p_r(t+1) = \gamma(t)p(t) + (1-\gamma(t))p_r(t)$$

Based on the decision of the forwarder, the carrier decides the contract prices to maximize his profit. As Figure 6 and Figure 7 show, the carrier's optimal contract prices curve and the forwarder's procurement curve are the kinked S-shaped.

## 7. Conclusions and Future Study

We propose a stylized model of dynamic pricing for carrier that sells a finite quantity of units of capacity, to price reference shippers, using dynamic programming method and classical variation method. In the sea-cargo service supply chain, the carrier applies discount contract under the market demand influenced by current price and reference price. Based on that, the paper considers two situations: the reference effect is symmetric information or not. Under both of situations, how the carrier makes his discount pricing decision and how the forwarder makes his resale price and procurement level are the critical problems in the paper.

For the first case, we identify dynamic von-Stackelberg equilibrium for the game between carrier (pricing strategy) and forwarder (ordering strategy). And we find single-stage market price  $p(t)$  for forwarder is pathwise increasing in the contract price  $k(t)$  and the reference price  $p_r(t)$ , decreasing in reference effect  $\gamma(t)$ , and single-stage profit function  $\pi^s(t)$  for the carrier is concave in the contract price  $k(t)$ . Moreover, an empirical study presents the discount pricing is between upper boundary  $\gamma=1$  and lower boundary  $\gamma=0$ , and the bigger the reference effect is, the lower the discount price is.

For the second case, the carrier designs relevant mechanism to make forwarder to show his type about the market demand, and based on that, to maximize his profit through Bayesian game.

Reference effect interfaces with the drivers of the benefits of price segmentation. Especially, the iterative reference effect leads to a complex where carrier and forwarder make decisions. Carrier and forwarder must take into account both the past price and the past reference price. In the paper, we assume the shippers strategically purchase, and the shippers know about all the resale prices and their own reference price.

There are several directions for future investigations within this line of research. For example, we can consider a situation where the spot market and contract one co-exist at the same time. Facing uncertain demand and

price of spot market, how carrier should allocate the capacity between spot and contract market, respectively. And in the contract market, what price carrier should announce. We don't allow players to overbook in the paper. However in practice, overbooking and other varieties of rules should be used to prioritize loads when demand exceeds total supply. And in the paper, we assume shippers have homogeneous memory effect  $\gamma$ . So a detailed investigation of reference effect with both contractual and spot market capacity purchases is another direction worthy of future study.

## 8. Acknowledgment

The authors would like to thank the referees for their helpful suggestions. The authors are grateful to Associate Professor ZHAO Quan-wu of Shantou University for his helpful comments on an earlier draft of the manuscript, and YANG Min, BEI Xiaofen of Shantou University for their valuable discussions. The research is supported in part by the National Science Foundation of China under Grant NO. 70701022, the Natural Science Foundation of Guangdong Province under Grant NO. 07301452, and the China Postdoctoral Science Foundation under Grant NO. 20060390847.

## REFERENCES

- [1] H. Rolf, "Capacity option for revenue management: Theory and applications in the air cargo industry," P. H. D. thesis, Springer, 2006.
- [2] S. Spinler, "Capacity reservation for capital-intensive technology: A Options Approach," Springer, Heidelberg, 2003.
- [3] D. Sigworth, "Contracting: Making the rate case," *Traffic World*, 268(17): pp. 32–33, 2004.
- [4] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, 47(2), pp. 263–291, 1979.
- [5] M. A. Lariviere and E. L. Porteus, "Selling to the newsvendor: An analysis of price-only contracts," *Manufacturing Service Operation, Management*, 3 (4): pp. 293–305, 2001.
- [6] B. Holmstrom and P. Milgrom, "Aggregation and linearity in the provision of intertemporal incentives," *Econometrica*, 55(2), pp. 303–328, 1987.
- [7] J. S. Raju, V. Srinivasan, "Quota-based compensation plans for multiterritory heterogeneous salesforces," *Management Science*, 42(10), pp. 1454–1463, 1996.
- [8] D. Gupta, "Carrier-Forwarder Contracts in the Air Cargo Industry," Department of Mechanical Engineering, University of Minnesota, Working Paper, 2006.
- [9] K. Amaruchkul, W. L. Cooper, and D. Gupta, "Air-Cargo capacity contracts under asymmetric information," Working Paper, 2007.
- [10] G. Gallego and G. van Ryzin, "Optimal dynamic pricing of inventories with stochastic demand over finite horizons," *Management Science*, 40 (8), pp. 999–1020, 1994.
- [11] R. G. Bitran and S. V. Mondschein, "Periodic pricing of seasonal products in retailing," *Management Science*, 43 (1), 1997, pp. 64–79.
- [12] E. P. Lazer, "Retail pricing and clearance sales," *The America Economic Review* 76(1) pp. 14–32, 1986.
- [13] Y. Aviv and A. Pazgal, "A partially observed markov decision process for dynamic pricing," *Management Science*, 51 (9), pp. 1400–1416, 2005.
- [14] Y. Aviv and A. Pazgal, "Optimal pricing of seasonal products in the presence of forwarder-looking customers," Working Paper, 2005.
- [15] I. Popescu and Y. Z. Wu, "Dynamic pricing strategies with reference effects," *Operations Research*, 55(3), pp. 413–429, 2007.
- [16] X. Z. Bu and X. lei, "Integration operation for production and pricing of dynamic supply chain under reference effect," working paper, pp. 10, 2007.
- [17] M. Helson, "Adaptation level theory," Harper and Row, New York, 1964.
- [18] G. Sorger, "Reference price formation and optimal marketing strategies," *Optimal Control Theory and Economic Analysis*, 3, pp. 97–120, 1988.
- [19] G. Fibich, A. Gavious, and O. Lowengrart, "Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effect," *Operations Research*, 51(5), pp. 721–734, 2003.
- [20] K. P. Kopalle, A. G. Rao, and J. L. Assuncao, "Asymmetric reference price effects and dynamic pricing policies," *Marketing Science*, 15(1), pp. 60–85, 1994.
- [21] Oliver, L. Richard, and S. Russell Winer, "A framework for the formation and structure of consumer expectations: Review and propositions," *Journal of Economic Psychology*, 469–499, December 1987.
- [22] V. A. Zeithaml, "Consumer Perceptions of price, quality and value: a means-end model and synthesis of evidence," *Journal of marketing*, 52(7), pp. 2–22, 1988.
- [23] A. Burnetas, S. M. Gilbert, and C. Smith, "Quantity discounts in single period supply contracts with asymmetric demand information," Working paper, 2005.
- [24] T. Taylor, "Supply chain coordination under channel rebates with sales effect," *Management Science*, 48(8), pp. 992–1007, 2002.
- [25] H. Varian, "Microeconomic Analysis," New York: Norton, 1992.
- [26] G. Kalyanaram and R. S. Winer, "Empirical generalizations from reference price and asymmetric price response research," *Marketing Science*, 14(3), pp. 161–169, 1995.
- [27] D. Putler, "Incorporating reference price effects into a theory of shipper choice," *Marketing Science* (summer), pp. 287–309, 1992.
- [28] Winer and S. Russell, "A reference price model of brand choice for frequently purchased products," *Journal of Shipper Research* 13 (September), pp. 250–256, 1986.
- [29] N. L. Stokey, R. E. Lucas, and E.C. Prescott, "Recursive method in economic dynamics," Harvard University Press, 1989.
- [30] R. F. Hartl and S. P. Sethi, "Optimal control problems with differential inclusions: Sufficiency conditions and an application to a production inventory model," *Optimal Control Applications and Methods* 5 (1984), pp. 289–307.
- [31] D. Fudenberg and J. Tirole, "Game Theory," China people publisher, 10 (in Chinese) 2002.
- [32] W. Y. Zhang, "Game theory and information economics," Shanghai people publisher, 8 (in Chinese), 1996.



# Innovation in the Financial Sector: Persistence and Schumpeterian Hypotheses

## —Econometric Evidence in Germany

Roberto Napoli

Faculty of Economics, University of Trento, Italy

Email: Roberto.napoli@unitn.it

Received August 22<sup>nd</sup>, 2008; revised November 4<sup>th</sup>, 2008; accepted November 21<sup>st</sup>, 2008.

### ABSTRACT

*The paper analyses innovation features in the German financial sector. The first topic is persistence of innovation. Our research question is: Do innovators plan further innovation for the subsequent year? In addition, since the sector is so far poorly researched, very basic questions are investigated in the paper: the relationship between firm size and innovation (both linear and quadratic), as well as the impact of market structure on innovation (i.e. Schumpeterian and neo-Schumpeterian hypotheses). Finally, Suttons argument of R & D sunk costs is investigated as a possible explanation for persistence. Basing on the CIS IV survey, our empirical evidence is consistent with the results of similar researches carried out in different sectors.*

**Keywords:** financial innovation, CIS, schumpeterian hypothesis

### 1. Introduction

Modern evolutionary economics sees the development and diffusion of innovations as a complex and unsteady process. Periods of radical changes that cause shifts in the technological paradigm alternate with phases of incremental innovation of given technologies. In trying to understand the drivers of such phenomena, much attention has traditionally been paid to the manufacturing sector, while only in the last few decades the interest of researchers has been devoted to services. Specifically, the financial sector is gaining centrality in the innovation process, and it has been recently described as crucial in influencing technological trajectories. In a neo-Schumpeterian framework, Perez [1,2] sheds new light on the role of financial intermediaries. She recalls the clear separation between borrower and lender, i.e. between entrepreneur and banker, which can be traced back to Schumpeter [3,4]. However, she argues that the role of financial intermediaries has been formally stated, but substantially not recognized from the neo-Schumpeterian literature, and from Schumpeter himself. Instead, she considers the banker as capable of true innovative commitment, just like the Schumpeterian entrepreneur.

This paper is understood as a first step in the analysis of the innovative dynamics going on in the financial sector.

The author thanks the Center for European Economic Research (ZEW) in Mannheim, the seminar participants at the University of Strasbourg (BETA) 2007 DIMETIC Doctoral Training "Micro Approaches to Innovation and Innovation Networks". The views expressed here are the author's. Any remaining errors are the author's responsibility.

As such, basic research questions are analysed, with the aim of providing some consistent answers which may then serve as basis for future, more detailed research. Questions involved in the analysis are mainly concerned with persistence of innovation, firm size and market structure effects on innovation (Schumpeterian Hypotheses), as well as the neo-Schumpeterian hypothesis of an inverse U shaped relationship between firm size and innovation. Furthermore, Suttons argument of R&D sunk costs is investigated as a possible explanation for persistence. The focus is on the financial sector and the analysis is carried out on a sample of 242 German firms. This sector is worldwide still poorly researched, as stressed by numerous studies, which makes it interesting to analyse very basic questions.

The first section defines innovation and addresses the problematic issue of measuring innovation. In the second section, theoretical issues and the main empirical findings about persistence of innovation are highlighted, and two different approaches based on patent statistics on the one hand, and on the Community Innovation Survey (CIS) on the other hand, are analyzed. Furthermore, Schumpeterian and neo-Schumpeterian hypotheses are briefly described, as well as the controversial empirical results recovered in the literature. This part serves as the theoretical framework for the subsequent analysis. The third part briefly describes some characteristics of the German financial sector. In the fourth section the data used in the model are described, as well as the model developed to

investigate persistence and the different hypotheses highlighted in earlier paragraphs; subsequently, the results of the model are presented. The final part discusses the findings and concludes with some suggestion for possible extensions of the model.

## 2. Defining and Measuring Innovation

In the present work, a firm having introduced a new or improved product or service or a new or improved process during the period covered by the survey, is considered an innovator. This means that we consider as an innovator a firm which reported innovative activities in the last three years, in terms of new products/services/processes introduced into the market. However it may be problematic to identify these innovations. In fact, the intangible nature of services, as well as the close interaction between production and consumption, makes the distinction between product and process innovation unclear. In addition, there is no clear cut between what should be considered true innovation and, on the other hand, what should be viewed as mere product differentiation. Unfortunately, incremental innovation, which is typical for the service sector and is highly interesting when analysing innovation, is difficult to distinguish from mere product customization, which in turn has to be excluded from the analysis. The more, radical innovation in the Schumpeterian sense occurs very rarely and is often little more than a theoretical eventuality. This makes it quite difficult to identify financial innovations in terms of single events. For the purposes of the present work, three definitions are relevant<sup>1</sup>:

- If the innovation involves new or significantly improved characteristics of the service offered to customers, it is a *product innovation*.
- If the innovation involves new or significantly improved methods, equipment and/or skills used to perform the service, it is a *process innovation*.
- If the innovation involves significant improvements in both the characteristics of the service offered and in the methods, equipment and/or skills used to perform the service, it is both a product and a process innovation.

These definitions are reported in the Community Innovation Survey (CIS) questionnaire. Since CIS data are used in the present paper to test the empirical model, we adopt the same definitions of innovation. This seems reasonable, given that respondents to the CIS survey are asked to self-identify as an innovator or as a non innovator basing on the same definition.

Measurement of innovation is a strongly debated issue in the economic literature. There are many different instruments to measure innovation. Input indicators like R&D expenditures belong to the first generation of

innovation indicators. They rely on the assumption of linear relationship between inputs and outputs of innovation, which has been rejected from the literature especially since evolutionarists like e.g. Nelson, Winter, Dosi began to influence heavily the scientific community in the early eighties. Patent statistics are one of the most traditional indicator for firm innovativeness: as an output indicator, they may work properly for manufacturing sectors (however with strong and well known limitations, see e.g. Malerba *et al.* [5] but fail completely in capturing innovation in most services, where patents are not an effective instrument to prevent imitation. Interestingly, Lerner [6] analyzes the dramatic increase in financial patents, observed in the US financial market between 1996 and 2001, and explains it as a consequence of changes in the federal law. However, financial formulas cannot be patented in most countries outside the US, especially in Europe. Furthermore, financial formulas are often developed in Universities. All this factors make patents an unfit tool to measure innovation in the financial sector.

A further group of measuring instruments, composed by those indicators capturing both inputs and outputs of innovation, as well as the process inbetween, overcome the drawbacks of “pure input” and “pure output” indicators, in that they recognize the complexity of the innovative phenomenon, at the cost of being often quite complicated themselves. Finally, a recently established instrument is the Community Innovation Survey (CIS), which has been introduced in Europe in the early nineties. Outcomes of the CIS approach are also highly disputed, due to the fact that self-definition of managers as an innovator is often considered too “soft” a tool to measure innovation<sup>2</sup>.

## 3. Previous Findings

### 3.1. Persistence of Innovation

Schumpeter distinguished between two market situations, known as Schumpeter Mark I and Mark II. The idea of persistence can be found in Schumpeter Mark II, also called “deepening pattern of innovation” in Malerba and Orsenigo [7], as opposed to the “widening pattern” (Mark I). In Schumpeter Mark II a few well established firms with large R&D divisions accumulate knowledge and innovative capabilities, which results in continuous innovation. Similarly, Winter [8] defined two technological regimes: the entrepreneurial regime, characterized by small firms, low entry barriers and high mortality; and the routinized regime, where bigger firms establish solid R&D departments with structured innovative activity. Much of the literature investigating innovation persistence aims at identifying the one or the other innovative pattern in the analyzed sector.

<sup>1</sup> See Oslo Manual (OECD 2005), p. 53

<sup>2</sup> See Tether [9] for an extensive analysis of advantages and drawbacks of CIS analysis

The idea of persistence is embedded in the concept of cumulation, defined as “the fact that existing innovators may contribute to be so also in the future with respect to non innovators” [7]. Malerba and Orsenigo [7] consider cumulativeness, and hence persistence of innovation, as directly linked to appropriability conditions: market power enables effective appropriability of innovation benefits, which in turn imply high cumulativeness conditions and hence ensure persistence of innovative behaviour in large and well established firms. In this perspective, innovation protection mechanisms build up a shield against imitation and allow profits (and rents) to innovations. This view, however, depicts rather extreme situations, which are more common in the manufacturing sector than in services. Specifically, the financial sector shows some features of the “widening pattern of innovation”, in that only 1,8% firms use patents as a protection mechanism<sup>3</sup> and imitation is amongst the biggest worries of managers, making the sector quite turbulent. At the same time it is characterised by high concentration and large firms, which makes it more similar to the sectors characterised by a “deepening pattern of innovation”. Consequently, it would be hard to forecast some specific features of persistence in the financial sector if we follow this classification. In fact, the Schumpeterian argument that firms have an advantage in R&D in the markets in which they have high market shares because market power enables them to capture the returns to innovation, doesn’t seem to hold for the financial sector, according to the widespread agreement that imitation is difficult to avoid and innovation returns difficult to capture. In sum, this view seems to rest on the core idea that innovation protection mechanisms, which can be enforced by large and well established firms, are effective in fostering innovation. However, innovation protection mechanisms is a much disputed theme on which traditional neoclassical views are challenged by the evolutionarist view [10], so that no assumption is made in the present paper as to how appropriability conditions work in the financial sector.

It is worth noting that the choice of the innovation measure may heavily affect outcomes of the analysis. As Gerosky *et al.* [11] point out, an overestimation of persistent innovative behaviour may be expected if R&D expenditures are used to measure innovation, as they occur on a routine basis. On the other hand, using patents as an innovation measure may be problematic too, as the link between patents and innovation outputs is still unclear. Roper *et al.* [12] argue that patent activity and firms’ innovation are only weakly related, whilst Dosi *et al.* [10] point out that the relationship between patents and innovation tends to differ between sectors and depends on industry-specific knowledge basis. Furthermore, patents may be registered on an irregular basis by the Patent Offices, which may not reflect the periodicity of firms’

decision to patent: this would heavily affect outcomes if persistence is to be analyzed [11]. Moreover, an underestimation of innovative activity may occur if patents are used as a measure for innovation. If firms undertake single innovative projects that last longer than one year, then their persistent innovative behaviour may turn into irregular patterns of innovations [11]. In this cases, firms may well be persistent innovators if their stream of innovative activity continues after the first multiple-year project, but in fact a year-by-year survey would misleadingly identify them as non-persistent innovators.

For the purposes of the present work, it seems useful to distinguish two groups of studies about persistence: in the first group patent statistics or R&D expenditures are used as a measure of innovation, while the second group is based on the CIS survey.

### 3.1.1. Patents and R&D As a Measure of Innovation

Common view of the first group of studies is that a small core of persistent innovators exist in most manufacturing sectors. As Cabbagnol and Le Bas [13] point out, big oligopolistic firms are more likely to carry out their innovative activities continuously and for long periods. Studying the British market, Geroski *et al.* [11] find that very few firms are persistently innovative, and that a critical mass of patents at firm level is necessary to pursue continuous innovative activity. Furthermore, even persistent innovators are so for short periods of time. It is noteworthy that Gerosky *et al.*’s results are rather extreme, as they tend to exclude altogether any influence of past innovation activity on the actual innovative behaviour of firms<sup>4</sup>. Le Bas *et al.* [14] as well as Le Bas and Latham [15] find similar results for French firms, suggesting that the size of innovation activity (measured, for instance, by the volume of R&D expenditures) be the main factor fostering persistence. Furthermore, on the background of previous studies (Malerba and Orsenigo [16,17,18], Malerba *et al.* [5]), Cefis and Orsenigo [5] ask if persistence of innovation is determined by the existing technological regime (as defined by Nelson and Winter [20], Dosi [21]) or rather is industry-specific. They also analyze cross-country differences in the degree of innovation persistence and find some degree of persistence both in innovators and in non innovators. Interestingly, non-innovators have a high probability to remain in the same innovative state over time. Furthermore, Cefis and Orsenigo [19] find relevant cross-country differences, while intersectoral differences do not vary substantially across countries, which leads to the conclusion that persistence is up to a certain extent a technology-specific variable. Malerba *et al.* [5]<sup>5</sup> suggest in

<sup>4</sup> “It is very hard to find any evidence at all that innovative activity can be self-sustaining over anything other than very short periods of time, at least for the kind of innovative activity we have focused on here.” (Gerosky *et al.* [11], p. 45).

<sup>5</sup> Malerba *et al.* [5] link innovation persistence to industry heterogeneity, arguing that firms having a competitive advantage in some field tend to enhance their commitment to innovation in the specific field and by this

<sup>3</sup> Mannheim Innovation Panel, ZEW, year 2004.

their patent-based cross-country analysis that a minimum threshold of innovative activity is necessary to become a persistent innovator. Cefis [22] analyzes in a more systematic way the nature of this threshold, and finds that the probability to switch from non-innovator to innovator by introducing one patent is much lower than the probability to increase the number of patents if this is not zero. Furthermore, Cefis [22] suggests that once the threshold is crossed, innovative activities may enjoy economies of scale, hence leading to persistent innovation. Bottazzi *et al.* [23] choose a slightly different approach, however still based on patent statistics. In order to study innovation in the pharmaceutical sector, they analyse the distribution of innovative drugs, both “New Chemical Entities” and patented products, into the US market<sup>6</sup>. Interestingly, they find that the introduction of different innovations in the market cannot be considered as independent events. Spill-over effects, as well as firm-specific learning effects of innovative activity may spread across research projects and influence subsequent innovation, which can be interpreted as a hint to persistence of innovative activity at firm level.

It is worth noting that most of the cited studies, show that innovation (in terms of number of patents) is persistent in a small number of firms only, which are normally characterized by large size and market power, hence showing features similar to the ones described in Schumpeter Mark II. As Malerba and Orsenigo [18] further point out, around this core of big and persistent innovators, a fringe of turbulent, occasional innovators, primarily composed by small firms, enter and exit the market, surviving only for short periods in the innovators group.

### 3.1.2. The CIS-Based Studies

The second group of studies uses the CIS approach to analyse innovative patterns related to persistence. In fact, patent statistics used from the first group tend to underestimate innovative activity, and hence persistence, since they capture only innovation first introduced in the market by the firm. As Duguet and Monjon [24] point out, this means that patent data could measure persistence of innovative leadership rather than persistence of innovation. Duguet and Monjon base instead on the Community Innovation Survey (CIS), where detailed data at firm level is provided and innovation is measured as the percentage of firms that self-identify as innovators. Duguet and Monjon find a high rate of persistence, and that size effects are in fact important in explaining persistence.

Specifically, smaller firms are motivated by dynamic increasing returns in the production of innovations, whereas persistence of innovation in larger firms, as also explained by the patent-statistics approach, originates from continuous R & D investments. Interestingly, Peters [25] involves in her analysis also the service sector, and finds that German manufacturers show higher rates of persistency than services, whereas in both cases true state dependence exists, in the sense that the decision to innovate in one period positively influences the probability to innovate in the subsequent period(s). Peters introduces in her model Suttons view of R&D investments as sunk costs [26]. The fact that R&D costs cannot be recovered, and that they are incurred to implement long term research departments, commit the firm to employ them over time. This may translate into persistent innovation. More recently, Roper and Hewitt-Dundas [12] analyze persistence in Ireland and Northern Ireland using both a quantitative approach and a qualitative case-studies analysis to get deeper insights about innovation patterns in persistently innovative plants. They distinguish between product and process innovation. They find high rates of persistence both in innovators and in non-innovators; moreover, they find a positive relationship between plant size and product as well as process innovation.

The first point which seems worth stressing is the effectiveness of the CIS approach to analyze innovation. Admittedly, patents are an objective measure of innovative activity, while CIS surveys are based on a self-identification as innovator by the respondent. Yet it is not easy to see how else to measure innovation in services, if not using CIS surveys. The second point is the focus of the studies belonging to the second group, which in most cases is on the manufacturing sector. Peters however compares persistence of innovation in the manufacturing and the services sectors, which is only possible using the CIS database. Finally, and most importantly, it is worth noting that whenever CIS analyses are concerned, each observation of the panel covers innovative activities over a 3-year period and data are collected with a four-year interval<sup>7</sup>. This implies that a firm is considered as a persistent innovator if it introduced one or more innovations, say, in the period 1996-1998, and again in the period 2000-2002<sup>8</sup>. However, this seems to provide a too weak definition of persistent innovator. In fact, one should consider the dynamics going on in services and even more in the financial sector, where new products are quickly replaced by newer ones. Service firms introduce regularly new products, which may differ from old ones only through slightly changed characteristics or added

---

way they reproduce initial asymmetries and end up with generating further heterogeneity. In their paper, persistence is not really investigated, but rather used as an explanatory variable to describe firm-level innovative activities across sectors and countries. Malerba *et al.* [5] also investigate implications of persistence and firm-heterogeneity on concentration, market entry and exit, firm size. However, these interesting relationships go beyond the purposes of the present work.

<sup>6</sup> This approach reduces the limitation of the traditional patent-statistics approach, in that it considers also new products introduced in the market without being patented.

---

<sup>7</sup> CIS I (1990-92), CIS II (1994-96), CIS III (1998-2000), CIS IV (2002-04). E.g. the survey of 2001 refers to years 1998-2000, next survey of 2005 refers to 2002-2004.

<sup>8</sup> In Germany instead, where data are collected yearly, a further overlapping problem arises, since e.g. data collected in 2001 refer to the period 1998-2000, and data collected in 2002 refer to the period 1999-2001.

services. In this sense, they appear to be persistent innovators over short periods of time. Interestingly, prevailing in the timing of the launch of new products into the market is the most important strategy of German financial firms to overperform competitors (Napoli [27])<sup>9</sup>. This suggests that financial innovations “expire” very quickly, and firms react by replacing them quickly with new innovations. As a consequence, the analysis based on subsequent waves of three-year periods, may lead to an artificial overestimation of persistence. Instead, the period under analysis should be kept as short as possible to correctly identify persistent innovators.

One way to overcome this problem is proposed by Peters [25], who uses input measures (innovation expenditures), which are available on a yearly basis, rather than output measures. However, this point is problematic too, as it assumes that innovation inputs transform linearly into innovation outputs, thereby denying much of the evolutionary literature dealing with learning effects, human capital contribution, complexity of the whole innovative process etc. In the present work a further solution is proposed. The idea is to keep the time lag as short as possible<sup>10</sup>, so as to capture firms that innovate in the three-year period and plan to innovate immediately thereafter, i.e. in the subsequent year. These firms would be then defined as persistent innovators. Admittedly, this may not suffice to assess persistence in longer periods. However it allows a stronger assessment of persistence of innovation in the short run, which seems interesting given the short life-cycle of innovations in the financial sector. In contrast, a different approach which would identify as a persistent innovator a firm which introduced innovations in the period, say, 1998–2000 and then again in 2002–2004, seems less adequate given the mentioned characteristics of the financial sector, where products are quickly replaced and easily imitated (see e.g. Tufano [28], Roper and Hewitt-Dundas [12]).

There are some counterarguments to the existence of persistence, like e.g. standardization. Once a new technology has been successfully introduced and sufficiently imposed as a standard in the market, some conservative-rather than innovative-forces can be at work in the firm, and make continuous innovation or persistency less likely. In this direction work path dependence, learning processes and network externalities, thus reinforcing standardization<sup>11</sup> and perhaps discouraging further innovation from the innovator itself, which may now be more concerned with establishing a market for its new product rather than developing new ideas.

<sup>9</sup> “Timing advantage” (‘Zeitlicher Vorsprung’) is seen as the most effective way to protect IPR in the German financial sector (MIP, 2005 survey).

<sup>10</sup> Needless to say, this contrasts with the necessity to measure persistence over a longer period.

<sup>11</sup> See e.g. Teece [29].

Furthermore, firms could cannibalise rents of their own innovations by introducing new products, hence having a negative incentive towards persistent innovation (“replacement effect”, see Le Bas, Latham 2004). However, the opposite may hold as well: new products introduced in period *t* may complete or improve the performance of products introduced in period *t* (Gilbert and Newberry [30]).

### 3.2. Schumpeterian and Neo-Schumpeterian Hypotheses

There is a broad literature dealing with the so called Schumpeterian hypotheses<sup>12</sup>, i.e. with the relationship between market structure and innovation on the one hand, and firm size and innovation on the other hand. There is no doubt that the search for consistent findings in this area failed in coming up with general results (see e.g. Symeonidis [31], Teece [29]). Still, some firm-size or market structure effects on innovation may be relevant in subsectors, and failing in capturing them may lead to incomplete explanations. Gellatly and Peters [32] for example, analysing three service subsectors, find higher innovation rates in more concentrated segments (financial services) than in less concentrated ones (communication and technical business services).

Our data suggest that a size effect exists in the German financial sector. The 242 analyzed firms have been divided into 10 subgroups, each with approximately 24 firms. The first group (1–53 employees) shows innovation rates which are lower than the average; the second group of firms (54–600 employees) moves around the mean, while the last group of large firms show the highest innovation rate. This figures suggest some positive relationship between size and innovation rate, which will hence be tested in the model.

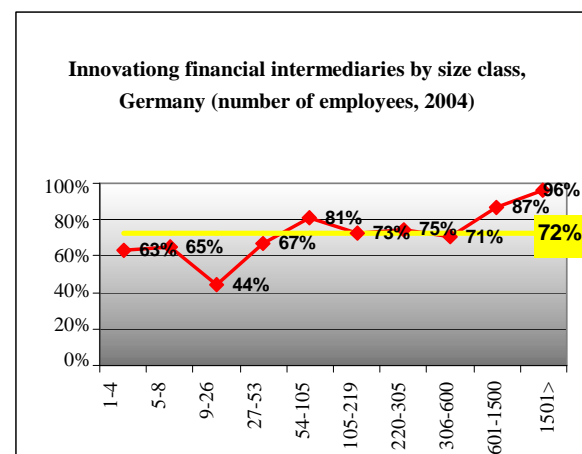


Figure 1. Source: Mannheim innovation panel

<sup>12</sup> Kamien and Schwartz [33] summarize the neo-Schumpeterian hypotheses and the inconclusive empirical work on these arguments. See also Cohen [34] and Cohen and Levin [35] and, more recently, Vaona and Pianta [36] for a literature review on the relationship between size and innovation.

The idea of the firm size being related to the innovative activity can be further expanded. It may well be that a positive relationship, which we expect to find between size and innovation, is quadratic rather than linear. It seems reasonable that the positive effect on innovation of one additional employee expires at a certain firm size level. This may be due to inefficiencies or to organizational problems, which may arise when the firm size grows. This is the so called “Neo-Schumpeterian hypothesis”, which is understood as an extension of the Schumpeterian hypothesis. In order to test it, the relationship between the squared size and innovation activity is analyzed. There are a few examples in the literature, where higher degree relationships have been found between firm size and R&D. Acs and Audretsch [37] and Siddharthan [38] report a quadratic U-shaped relationship, while further studies found also evidence of a cubic relationship between firm size and R&D activities (see Kumar and Aggarwal [39] for more details). The idea of a cubic relationship however, seems too extreme, and some doubts may arise as to how to interpret results. The quadratic relationship instead, seems interesting in terms of management issues: an inverse-U relationship, as argued by neo-Schumpeterians, would mean that expanding the firm size may ensure advantages in terms of innovative activity only up to a certain level, and may turn into an hampering factor if the firm becomes too large. To test this hypothesis in the model discussed later the square of firms size (number of employees) will be used as a regressor. It seems appropriate to keep in the model both measures of the firm size<sup>13</sup>, so as to investigate both the linear and the quadratic relationship of size with the probability to innovate. In fact, the outcome (which we expect) of a positive linear relationship between size and innovation would fall short of a complete explanation about the extent of this relationship (does size effects indefinitely foster innovation or do they expire once a certain level is reached?). In this case, introducing the second degree variable could add useful insights on that. In turn, the squared relationship alone would explain the relationship in a poor way, as the linear relationship cannot be inferred from the quadratic one<sup>14</sup>.

As far as known, no studies have yet analysed persistence of innovative activities in financial firms, while only a few studies have recently tested Schumpeterian hypotheses in the financial sector [40,41]. None of them, however, concentrated on the neo-Schumpeterian hypothesis. More in general, the lack of empirical literature on the determinants of financial innovation has been repeatedly

stressed (see e.g. Frame and White [42], Heffernan *et al.* [40]). This makes the topic even more interesting, since the sector is gaining growing attention. The contribution of the present study to the literature is twofold. First, it is one of the few empirical studies of financial innovation. Second, it identifies some possible factors underlying financial innovation.

The present study is based on CIS data to study persistence mainly for two reasons. The first concerns with the well recognized and already mentioned limitations of the patent statistics, like e.g. underestimation of innovative activity, which can be even more effective in services than in the manufacturing sector. But there is an even stronger argument that makes it impossible to use patent data. In fact, patents are not a widespread mechanism to protect innovations in the financial sector, since less than 2% of German bankers and insurers use them to protect innovation<sup>15</sup>. The neglect of patents as an effective protection mechanism is likely to hold also in neighbour States due to common laws at European level, which e.g. exclude patentability of financial formulas<sup>16</sup>. Furthermore, as Tufano [43] points out, the easily imitated nature of financial innovation does not lend itself to models based on patent statistics.

#### 4. The German Financial Sector

As shown in following figure, the incidence of big firms in the German financial sector is much higher than the incidence of big firms in German services. In fact, 677 out of 2.742 financial intermediaries (or 25%) have more than 250 employees, while the percentage falls to 5% if the whole service sector is considered.

Concentration measures in the German financial sector are calculated basing on revenues stated by firms and reported in the 2005 MIP survey. The CR4 Concentration Ratio (40% for the financial sector) and the CR8 (60%)

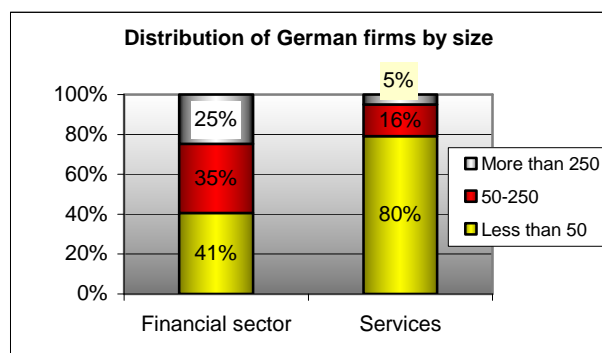


Figure 2. Source: Mannheim innovation panel, 2004

<sup>13</sup> I.e. the logarithm of employees and the squared logarithm of employees.

<sup>14</sup> As an example, if we find a negative quadratic relationship, but don't know anything about the linear relationship, we are not able to understand if size has a positive or negative effect on innovation, as the negative quadratic relationship contains both effects and does not allow, on its own, to understand which one prevails.

<sup>15</sup> Mannheim Innovation Panel, ZEW, year 2004.

<sup>16</sup> Lerner [6] shows a dramatic increase in the number of U.S. financial patent awards due to patentability of financial formulas newly introduced in the U. S. law. However, similar patterns are not likely to show up in Europe.



show a highly concentrated financial market. Concentration is even higher if data are disaggregated by sub-sectors. Furthermore, the Herfindahl-Hirschman Index (HHI) confirms that the banking sub-sector<sup>17</sup> is more concentrated than the insurance sub-sector<sup>18</sup> and than the financial sector as a whole.

## 5. Econometric Analysis

### The Database

The data used for the analysis are firm level data from the Mannheim Innovation Panel (MIP) in the German financial services sector (NACE3 651, 652, 660, 671, 672). The MIP is based on innovation surveys carried out by the Centre for European Economic Research (ZEW) on behalf of the German Federal Ministry of Education and Research. The target population covers all legally independent firms with 5 or more employees and the surveys are drawn as stratified random samples (stratified by firm size, branches of industries and East/West region). The samples are constructed as panels and about 10.000 firms in manufacturing and 12.000 service firms are questioned each year. Participation is voluntary and the response rate varies between 20% and 25%. The survey methodology is detailed in the OSLO-Manual (OECD 2005). The data which are used to test hypotheses stem from the 2005 survey. Following table summarizes the population of German financial firms and the sample used for the estimation of the model.

### Hypotheses

1) The first relationship analysed in the proposed model is the one between innovation activities in the period 2002-2004 and innovative projects for 2005, with the aim of assessing short-run persistence of innovation at firm level. The rationale behind this choice is straightforward:

**Table 1. Concentration in the German financial sector**

	Banks	Insurances	Financial sector
CR4	70%	52%	40%
CR8	80%	70%	60%
HHI	1.739	830	562

Source: Mannheim Innovation Panel, 2004

**Table 2. The German financial sector**

NACE3		Population	Sample
65	Financial intermediation except insurance and pension funding	2.053	117
66	insurance and pension funding except compulsory social security	490	55
67	Activities auxiliary to financial intermediation	199	70
	Totale	2.742	242

<sup>17</sup> NACE3= 651, 652, 671

<sup>18</sup> NACE3= 660, 672

persistence of innovative behaviour requires one firm to be an innovator and to plan new innovation for the subsequent year. As Malerba *et al.* [5] point out, “in the simplest statistical interpretation, the notion of innovative persistence can be defined as the conditional probability that innovators at time  $t$  will innovate a time  $t+1$ ” (p. 804).

We expect to find a high rate of persistence in the financial sector for two reasons. First of all, such a result would be consistent with prior researches using CIS data for other sectors (see previous section). Secondly, persistence at industry level is evident from following figure:

The figure shows persistence at industry level in the financial sector. Between 1994 and 2005 the rate of innovating firms moved around 70%, with some lower values in 2002–2003. However, this figure does not provide insights about firms identity. One possible explanation of this figure is that there might be continuous (or frequent) new entrance in the markets, which increments the innovation rate. Malerba and Orsenigo [18] find an extreme turbulence in innovative activities in the manufacturing sector, and a high turnover of innovative firms, which would exclude high rates of innovative persistence at firm level. On the other hand, there might be a big group of firms that innovate persistently, where turbulence would concern and a smaller group of non innovators which steadily enter and exit the market. The conclusion cannot be drawn from the figure, and firm level analysis is required.

*Hypothesis 1: There is a positive relationship between past innovative activity and innovative (expected) behaviour in year  $t+1$ .*

2) The second relationship tests the Schumpeterian hypothesis of positive correlation between firm size and innovation. In past paragraphs hints of a positive relationship between size and innovation have been highlighted in the German financial sector. However, there are also counterarguments to this Schumpeterian hypothesis. Scherer and Ross [44], e.g., argue that small firms innovate more because too much bureaucracy



**Figure 3. Source: MIP, own calculation**

inhibits innovative activities, and this is more likely to happen in larger firms. Given the hints of a positive relationship between German financial firms and innovation, we expect to provide empirical evidence confirming the Schumpeterian hypothesis.

*Hypothesis 2: Firm size is positively correlated to innovativeness*

3) Furthermore, referring to the mentioned literature, the so-called neo-Schumpeterian hypothesis of an inverse U-shaped relationship between size and innovativeness will be tested, in order to evaluate if size effects of innovation vanish above a certain firm size.

*Hypothesis 3: A negative second degree relationship exists between firm size and innovativeness*

4) The fourth hypothesis can be traced back again to Schumpeter, as it deals with the relationship between market structure and innovativeness. The rationale is that one firm's market power can be measured by the number of competitors who market similar products, i.e. substitutes. In the German CIS questionnaire, firms are asked about the number of direct competitors they face in the market. If this number is low (up to five), then a firm is considered to have high market power. We expect to confirm that these firms are more innovative. This would mean that the German financial sector displays features which are similar to the Schumpeter Mark II scenario described in previous chapters.

*Hypothesis 4: Market power, in terms of small number of competitors, is positively associated to innovation*

5) Finally, empirical evidence is provided to the Sutton hypothesis of R&D sunk costs and innovation. If the amount of investments for innovative activities in year  $t$  is positively associated with innovation both in year  $t$  and  $t+1$ , this may be due to the lock-in effect caused by R&D sunk costs. In other words, expenditures in R&D in year  $t$  commit firms to innovate in year  $t+1$ .

*Hypothesis 5: Innovation in year  $t+1$  is positively influenced by investments for innovative activities in year  $t$ .*

### Model Specification

The above hypotheses are tested with a data set of 242 German firms in the financial sector from the Mannheim Innovation Panel. Data refer to year 2004, with the exception of the dependent variable as explained hereafter. The variables are defined as follows:

$INNO_{i2005}$  (dummy variable): innovative activities planned by firm  $i$  for 2005. Firms have been requested if they planned some innovative activity for subsequent years (2005 and 2006). Since the survey has been carried out in 2005, the answer is to be considered a forecasted value, or some sort of "expected innovation". As such, the planned innovation rate may differ from the true value. However, given the short horizon of the forecast, the "planned value" can be considered as a reliable proxy for the true value (which of course was unknown in 2004). It

seems realistic that planned innovation transforms into effective innovation in the subsequent year. Some support to this belief is provided by the low rate of firms that give up innovative projects before completing them.

$INNO_{i2004}$  (dummy variable): innovative activities carried out by firm  $i$  in 2004.

$EMPL_{i2004}$ : number of employees of firm  $i$ . It will be used in logarithmic form. Also, the squared value will be tested for its influence on the dependent variable.

$OLIG_{i2004}$ : (dummy variable) equals one if firm  $i$  has five or less competitors that market similar products.

$INNOEXPS_{i2004}$ : Expenditures for innovative activities as a proportion of revenues of firm  $i$ .

$EAST_{i2004}$ : (dummy variable) firms headquartered in East Germany.

$EXP_{i2004}$ : Export value (=sales abroad).

In the following, descriptive statistics of the variables introduced in the model are reported. In order to provide further relevant insights about the German financial market, the number of employees is also reported (in non-logarithmic form), as well as the absolute value (i.e. not as a proportion of revenues) of expenditures for innovative activities.

Since the dependent variable is dichotomic, a probit model is used in order to test the influence of independent variables. Summarising the above discussion and hypotheses in a functional form:

### Econometric Results

The following table reports the estimation results of the probit model with all the independent variables including the control variables. Note that marginal effects are reported, as well as the p-value (in parenthesis).

The first important result is a positive and significant relationship between innovation activities carried out in 2002–2004 and plans to innovate in 2005, as predicted by hypothesis 1. This means that firms that innovated in the

	Type	Mean	Std.dev.
$INNO_{i2005}$	dummy	0.756198	0.430264
$INNO_{i2004}$	dummy	0.723141	0.4483740
$EMPL_{i2004}$ (ln)	cont.	4.547479	2.220977
$EMPL_{i2004}$ (ln, squared)	cont.	26.136261	20.625551
$OLIG_{i2004}$	dummy	0.474790	0.500416
$INNOEXPS_{i2004}$	cont.	0.038873	0.112905
$EAST_{i2004}$	dummy	0.165290	0.372211
$EXP_{i2004}$	cont.	143.0131	1388.594
Employees	(cont., not in the model)	742.9417	2520.252
Innov. expenditures	(cont., not in the model)	6.095196	33.76934

**Table 3. Marginal effects**

$Y = INNO_{i2005}$	Marginal effects (p-value)	
$INNO_{i2004}$	0.203	(0.000)**
$EMPL_{i2004}$ (ln)	0.042	(0.010)*
$EMPL_{i2004}$ (ln, squared)	-0.003	(0.044)*
$OLIG_{i2004}$	-0.021	(0.191)
$INNOEXPS_{i2004}$	0.635	(0.032)*
$EAST_{i2004}$	0.011	(0.532)
$EXP_{i2004}$	0.001	(0.135)
Constant	-3.347	(0.000)**
Observations	138	
* significant at 5%; ** significant at 1%		
Pseudo $R^2 = 0.3937$		

(See appendix for correlations)

period 2002-2004 are likely to innovate also in 2005. Hence, according to our estimation and to our short-run definition of persistence, German financial firms display a persistent innovative behaviour over the analyzed period.

Furthermore, expenditures for innovative activities incurred in 2004 positively and significantly influence the probability to innovate in the subsequent year, suggesting some lock-in effect of R&D investments, as argued by John Sutton [26].

Finally, the number of direct competitors in the market does not seem to have an impact on the probability to innovate. The Schumpeterian hypothesis of market structure influencing innovation could not be confirmed. Admittedly, this may be due to the fact that OLIG is a bad proxy for market power. However, different results may have been obtained focusing only on product innovation. In fact, OLIG directly refers to product competitors, and may be a better proxy for market power as far as only product innovation is concerned. Instead, we analysed product and process innovations together, hence results may be biased.

## 6. Discussion and Concluding Remarks

How can we explain persistence, given the financial sector's characteristics so far described (e.g. low protection mechanisms, high rates of imitation, high concentration) and the results of the model that suggests persistence, albeit limited to the short run? According to Tufano [28], who analyses the first mover advantages in the financial sector, innovators gain know how advantages and new knowledge while developing an innovation. This

knowledge capital can be further improved and applied to develop further innovations, hence leading to persistence. Similarly, Merton [45] uses the metaphor of "financial innovation spiral" meaning that one innovation begets the next. Both these ideas are consistent with our findings that the most firms showing innovative behaviour in 2002–2004 have already planned innovations for upcoming years. In fact, not only they plan innovations for year 2005 as shown in the model. They also have plans for year 2006 (results not shown in the model), which suggests that persistence may hold also beyond our limited 2-years horizon. Consistently with this explanation, Tufano [43] provides examples of financial innovations built upon recent new products and aimed at improving their performances or better accomplish their functions. In all these cases, low appropriability conditions, along with ease of imitation, seem to play a major role in committing firms to innovate continuously, in order to offset competitors' gains from imitation. By this way a reinforcing loop may be at work, resulting in persistence of innovation at firm level. In the same direction may work past investments in innovative activities, as shown in the model. The commitment to innovation can be further reinforced by past R&D expenditures, which have been found to influence future innovation.

For what concerns firm size and innovation, our results are straightforward: while the linear relationship shows that firm size is important in determining innovation, the negative quadratic relationship suggests that this is only true up to a certain level. Large firm size can be detrimental to innovation: one simple explanation suggests that organisational diseconomies may be at work. Therefore, medium-sized firms are responsible for the bulk of the R&D activity. However, an estimation of the point of inflection could provide useful insights to understand to what extent large firm size negatively affects innovation.

## Limits of the Model and Further Research

The first concern is about the insights which can be drawn from the present model about persistence. Given the lack of data allowing to test for persistence in the services sector (patent statistics do not exist in many subsectors, CIS data refer to a too long period and tend to overestimate persistence), the present approach suggests a new solution to test persistence, which applies in the short run. This seems not too unrealistic in the financial sector, where the financial product's life cycle is short and imitation occurs very quickly, forcing competitors to renew their product lines on a regular basis. Admittedly, our results on persistence however seem to capture only one part of the phenomenon and further empirical evidence is needed on this topic.

In addition, the model, because of the econometric approach chosen, fails in capturing the unobserved individual heterogeneity, which Peters [25] has proven to explain persistence of innovation across sectors.

The empirical analysis proposed can be seen as a first step in the still poorly researched field of financial innovation. As such, very basic questions have been addressed, like the relationship between firm size and innovation as well as evidence about innovation persistent behaviour at firm level. An interesting point, which would be worth analysing, would be to find the threshold upon which positive firm size effects expire, and a further increase in the firm dimension has negative impact on the probability to innovate. This relationship emerges in the model, but the threshold remains unknown.

Furthermore, it could be interesting to distinguish between banks and insurances in the financial sector, as well as between product and process innovations. Significant differences can emerge with respect to firm size, in the sense, for example, that smaller firms could chose different strategies of product/process innovation with respect to larger firms. Also, it can be distinguished between firms which aim at internalize the results of their innovation activities, i.e. innovate for themselves, and firms that innovate for other players. The former are more likely to develop process innovations, the latter product innovation.

## REFERENCES

- [1] C. Perez, "Technological revolutions and financial capital: The dynamics of bubbles and golden ages," *Celtenham*, Elgar, 2002.
- [2] C. Perez, "Finance and technical change: A neo-schumpeterian perspective," to appear in H. Hanusch and A. Pyka eds., *The Elgar Companion to Neo-Schumpeterian Economic*, Edward Elgar, Celtenham, 2004.
- [3] J. Schumpeter, "The theory of economic development," Harvard University Press, Cambridge, Massachusetts, 1934.
- [4] J. Schumpeter, "Capitalism, socialism and democracy," New York, 1942.
- [5] F. Malerba, L. Orsenigo, and P. Peretto, "Persistence of innovative activities, sectoral patterns of innovation and international technological specialization," *International Journal of Industrial Organization* No. 15, pp. 801–826, 1997.
- [6] J. Lerner, "Where does state street lead?" *A First Look at Financial Patents, 1971–2000*, *Journal of Finance*, 57, pp. 901–930, 2002.
- [7] F. Malerba and L. Orsenigo, "Schumpeterian patterns of innovation," *Cambridge Journal of Economics*, 19, pp. 47–65, 1995.
- [8] S. G. Winter, "Schumpeterian competition in alternative technological regimes", *Journal of Economic Behaviour and Organization*, Vol. 5, pp. 287–320, 1984.
- [9] B. S. Tether, "The sources and aims of innovation in services: Variety between and within sectors," *Econ. Innov. New Techn.*, Vol. 12 (6), pp. 481–505, 2003.
- [10] G. Dosi, L. Marengo, and C. Pasquali, "How much should society fuel the greed of innovators?" *On the relations between appropriability, opportunities and rates of innovation*, *Research Policy*, Vol. 35, pp. 1110–1121 controlla anno, 2006.
- [11] P. Geroski, J. V. Reenen, and C. F. WALTERS, "How persistently do firms innovate?" *Research Policy*, 26, pp. 33–48, 1997.
- [12] S. Roper and N. Hewitt-Dundas, "Innovation persistence: Survey and case-study evidence," *Research Policy* No. 37, pp. 149–162, 2008.
- [13] A. Cabbagnol and C. G. Le Bas, "How persistently do firms innovate?" *An Evolutionary View, An Empirical Application of Duration Models*, presented at the European Meeting of Applied Evolutionary Economics, 7–9 June 1999, Grenoble, France, 1999.
- [14] C. Le Bas, A. Cabagnols, and C. Gay, "An evolutionary view on persistence in innovation: An empirical application of duration model," in P. Saviotti (ed.) *Applied Evolutionary Economics*, Northampton, MA: Edward Elgar, 2003.
- [15] C. Le Bas and W. R. Latham, "Persistence of firm innovative behaviour: Towards an evolutionary theory," working paper, 2004.
- [16] F. Malerba and L. Orsenigo, "Schumpeterian patterns of innovation," *Cambridge Journal of Economics*, 19, pp. 47–65, 1995.
- [17] F. Malerba and L. Orsenigo, "Schumpeterian patterns of innovation are technology-specific," *Research Policy* 25, pp. 451–478, 1996.
- [18] F. Malerba and L. Orsenigo, "Technological entry, exit and survival," *Research Policy* 28, pp. 643–660, 1999.
- [19] E. Cefis and L. Orsenigo, "The persistence of innovative activities: A cross-countries and cross-sectors comparative analysis," *Research Policy* No. 30, pp. 1139–1158, 2001.
- [20] R. Nelson and S. Winter, "An evolutionary theory of economic change," *The Bellknap Press of Harvard University Press*, Cambridge, MA, 1982.
- [21] G. Dosi, "Sources, procedures and microeconomic effects of innovation," *The Journal of Economic Literature*, pp. 1120–1171, 1988.
- [22] E. Cefis, "Is there persistence in innovative activities," *International Journal of Industrial Organization* No. 30, pp. 489–515, 2003.
- [23] G. Bottazzi, G. Dosi, M. Lippi, F. Pammolli, and M. Riccaboni, "Innovation and corporate growth in the evolution of the drug industry," *International Journal of Industrial Organization* No. 19, pp. 1161–1187, 2001.
- [24] E. Duguet and S. Monjon, "Is innovation persistent at firm level?" *An Econometric Examination Comparing the Propensity Score and Regression Methods*, University of Paris I-Cahiers de la MSE Working Paper No. 2004(75), 2004.
- [25] Peters, "Persistence of innovation: Stylised facts and panel data evidence," *Discussion Paper* No. 05–81, 2006.
- [26] J. Sutton, "Sunk costs and market structure," Cambridge, 1991.
- [27] R. Napoli, "L'innovazione finanziaria in Italia e Germania: Statistiche descrittive e considerazioni teoriche", in *Bancaria*, n. 2008/1, pp. 90–102, 2008.
- [28] P. Tufano, "Financial innovation and first mover advantages," *Journal of Financial Economics*, No. 25, pp. 213–240, 1989.
- [29] D. Teece, "Reflections on "Profiting from innovation," *Research Policy* No. 35, pp. 1131–1146, 2006.
- [30] R. Gilbert and D. Newberry, "Pre-emptive patenting and the persistence of monopoly," *American Economic Review*, 72 (3), pp. 514–526, 1982.

- [31] G. Symeonidis, "Innovation, firm size and market structure: Schumpeterian hypotheses and some new themes," OECD 1996.
- [32] G. Gellatly and V. Peters, "Understanding the innovation process: Innovation in dynamic service industries," Statistics Canada, Working Paper 127, 1999.
- [33] K. Kamien and N. Schwartz, "Market structure and innovation," Cambridge University Press, 1982.
- [34] W. Cohen, "Empirical studies of innovation activity," Oxford: Blackwell, 1995.
- [35] W. M. Cohen and R. C. Levin, "Empirical studies of innovation and market structure," in Handbook of Industrial Organization, eds R. Schmalensee and R. D. Willig, New York: North-Holland, 1989.
- [36] A. Vaona and M. Pianta, "Firm size and innovation in European manufacturing," Kiel Working Paper No. 1284, 2006.
- [37] Z. J. Acs and D. B. Audretsch, "Innovation in large and small firms: An empirical analysis," The American Economic Review, Vol. 78, No. 4, pp. 678–690, 1988.
- [38] N. S. Siddharthan, "In-house r & d, imported technology and firm size: Lessons from Indian experience," Developing Economies, No. 26, pp. 212–21, 1988.
- [39] N. Kumar and A. Aggarwal, "Liberalization, outward orientation and in-house r&d activity of multinational and local firms: A quantitative exploration for indian manufacturing," unpublished paper, 2000.
- [40] S. Heffernan, "Financial Innovation in the UK," Working Paper No 4, Cass Business School, City University, London, 2008.
- [41] J. Lerner, "The new new financial thing: The origins of financial innovations," Journal of Financial Economics, 79, pp. 233–255, 2006.
- [42] W. S. Frame and L. J. White, "Empirical studies of financial innovation: Lots of talk, little action?" Federal Reserve Bank of Atlanta, Working Paper, No.12, July 2002.
- [43] P. Tufano, "Financial innovation," working paper, 2002.
- [44] F. M. Scherer and D. Ross, "Industrial market structure and economic performance," 3rd ed. Boston: Houghton-Mifflin, 1990.
- [45] R. C. Merton, "Financial innovation and economic performance," Journal of Applied Corporate Finance 4(4), pp. 12–22, 1992.

## Appendix

correlation matrix

	<i>INNO</i>	<i>INNOEXPS</i>	<i>EMPL (ln)</i>	<i>EMPL (ln, squared)</i>	<i>OLIG</i>	<i>EAST</i>	<i>EXP</i>
<i>INNO</i>	1.0000						
<i>INNOEXPS</i>	0.1139	1.0000					
<i>EMPL (ln)</i>	0.2478	-0.0991	1.0000				
<i>EMPL (ln, squared)</i>	0.2569	-0.1026	0.9644	1.0000			
<i>OLIG</i>	-0.0801	-0.0716	0.0823	0.0185	1.0000		
<i>EAST</i>	-0.0786	-0.0045	-0.1760	-0.1487	-0.0324	1.0000	
<i>EXP</i>	-0.0578	-0.0279	0.1634	0.1941	0.0939	-0.0486	1.0000

# A Way to Improve Knowledge Sharing: from the Perspective of Knowledge Potential

Lingling Zhang<sup>1,2\*</sup>, Xiuyu Zheng<sup>1,2</sup>, Jun Li<sup>1,2</sup>, Guangli Nie<sup>1,2</sup>, Guoqing Huo<sup>1</sup> & Yong Shi<sup>2,3</sup>

<sup>1</sup>School of Management, Graduate University of Chinese Academy of Sciences, Beijing (100190), China, <sup>2</sup>Research Centre on Fictitious Economy and Data Science, CAS, Beijing (100190), China, <sup>3</sup>College of Information Science and Technology, University of Nebraska at Omaha, Omaha, NE 68118, USA.

Email: zhangll@gucas.ac.cn, baobaozhuxy@sina.com, buaalijun@163.com, sdungl@163.com, leohuo@gucas.ac.cn, yshi@gucas.ac.cn

Received April 18<sup>th</sup>, 2008; revised October 29<sup>th</sup>, 2008; accepted November 19<sup>th</sup>, 2008.

## ABSTRACT

*Knowledge is the most important resource in an organization, and the knowledge transfer and sharing between employees is of vital importance for organizations. "Prisoner's dilemma" exists in the process of the organizational knowledge transfer and sharing when the employees transfer their knowledge to the organization and share their knowledge with other employees. This paper analyzes the process and obstacle of the knowledge transfer and sharing in the organization and studies the game model of the knowledge transfer and sharing, and put forward the conclusion that different knowledge potential employees should be stimulated by different measures. Through analyzing principle-agent in the incentive mechanism, introducing the equity incentive method will have infinitely repeated games to the knowledge high-potential employees who are the key sources of the knowledge transfer and sharing in the incentive mechanism design. This makes it possible to break the prisoner's dilemma of the knowledge transfer and sharing.*

**Keywords:** knowledge potential, knowledge sharing, knowledge transfer, incentive mechanism, game theory

## 1. Introduction

In the era of knowledge economy, knowledge has become important resource and the most crucial strategy elements in an organization. Many researchers have pointed out that knowledge transfer and sharing among employees would lead to faster responses to customer requirement at a lower cost in operations. The knowledge creation and the ability to apply knowledge are the most important sources of the sustainable competitive advantage in an organization [1,2]. Knowledge has a characteristic which is the same as the currency, i.e. people can't find its value unless it is used or transmitted. The knowledge is useful only when it becomes social sharing [3,4]. In other words, knowledge develops in the communication and value-added in the use.

Knowledge may be held by one person or embedded in organizational employees. Knowledge between employees is complementary [5,6], so knowledge transfer and sharing is propitious to the employees' knowledge richness and growth. However, due to organizational culture environment, incentive mechanisms, the characteristics of the knowledge and other reasons, knowledge transfer and sharing between the employees and the organization often get into prisoner's dilemma. Researches indicate that the quality of the employee can only play to 20%–30% in the environment of lack of incentives, but in a good incentive environment the same person can play a

potential of 80%–90%, which indicate that 50%–60% of the gap is caused by incentives.

Currently, the incentive of knowledge transfer and sharing based on principal-agent is mainly in two aspects. On the one hand is the tacit incentive method to solve the agency problem. On the other hand is the explicit incentive method to solve the agency problem. Propose incentive measures based on the information asymmetry model in the principal-agent relationship and put the residual distribution to link to the operating performance. Since the 1980s, economics introduce dynamic game theory to principal-agent relationship and demonstrate in the repeatedly agency relationship circumstances, the tacit incentive mechanism of competition, reputation and so on can encourage the agents, which enrich the incentive theoretical content in the long-term principal-agent relationship.

It is evident that organizational knowledge seems to be highly relevant to organizations to achieve sustainable advantages. As for this problem, the following approaches are mostly adopted in the previous study at home and abroad: 1) To expand game to the N-person prisoners' dilemma, and set up the utility functions and organization scale of the players so as to study the cooperation of organization through computer simulating the process of



the game [7] To introduce the model of infinitely repeated game and to institute the rules of knowledge transfer and sharing based on the experimental results of AXELROD so as to find out solutions [8,9,10]. The limitations of the previous studies are as follows:

1) The study is mainly focused on the analysis of the game model, and specific and feasible solutions suitable for corporations haven't been given.

2) Research is generally confined to break "prisoner dilemma" between the employees and the organization or between the employees, not having a good combination of them to analysis.

3) The utility function of the players is a specific function supposed by the writer, which cannot reflect the effect of knowledge sharing on the players fully and exactly.

In this paper we focus on the different knowledge potential of the employees in the organization, combining knowledge transfer and sharing between the employees and the organization to analysis, and design incentive mechanisms for the different employees which enable the employees and organization to achieve a win-win situation. The results demonstrate research on improving efficiency of organizational knowledge transfer and sharing based on knowledge potential is effective.

The rest part of this paper is organized as follows. The analysis of obstacles of knowledge transfer and sharing in the organization is presented in the next section. In Section 3, the game model of the knowledge transfer and sharing is described. Research on incentive mechanisms in knowledge transfer and sharing is discussed in Section 4. Section 5 concludes the paper with a discussion of the implications and limitations of this study, research directions, and concluding remarks.

## 2. The Analysis of Obstacles of Knowledge Transfer and Sharing in the Organization

A large number of scholars emphasize the importance of knowledge transfer and sharing in many related areas. According to the view of some scholars, the ability to transfer knowledge is a system which is superior to other arrangements such as unique source of market organizational competitive advantage [2]. An effective knowledge transfer and sharing is considered to be the key to a series organizational process and results, including best practice transfer, new product development, learning speed and organizational survival. So how to promote the employees in the organization to transfer and share knowledge is an important research topic. Knowledge transfer and sharing cannot progress naturally in an organization. In addition to the organizational structure, enterprise culture and the application of information technology etc external factors, the most important part is still its internal factors.

The knowledge has the characteristics of public property [11,12], the high cost production, the use of non-exclusive, but low dissemination cost. Once the knowledge is transferred or shared, its exclusive rights will be lost, while the exclusive rights are always the embodiment of the employees' values in the enterprises as well as guarantee of the security of their occupations. In the process of knowledge transfer and sharing, the employees obtain profits while they pay cost, therefore, employees' utility can be regarded as the difference between the gained profits and the costs. The employees are concerned that sharing knowledge has negative effects which will reduce their original knowledge value and undermine the competitiveness in the organization. Thus, when the employee cannot access to others' exclusive knowledge or hasn't got any compensation, his proceeds will be negative. The employee will lose his special value in the organization because of losing the exclusivity of knowledge. Because the employees' core knowledge is always tacit, the knowledge sharing must be based on a voluntary basis. So there are few people who are willing to share their core knowledge with others, which create the knowledge sharing barriers.

## 3. The Game Model of the Knowledge Transfer and Sharing

### 3.1. Knowledge Transfer between the Employees and the Organization

In the organization there are two forms in the knowledge transfer and sharing, one is between the employee and the organization, the other is between the employees. If all the employees are willing to transfer and share their own knowledge, not only the employees will raise their knowledge and skills, but also it's useful for the organization. The decisive factors to the quality and effectiveness of knowledge transfer and sharing are the subjective desire of the knowledge owner. As the organizations are unable to observe and measure the employees' knowledge sharing, so there are games exist between the employees and the organization. Assuming.

$k$  : Knowledge volume of the employees transfer and sharing.

$W$  : The incentive costs which organizations pay to the employees under encouragement.

$C(k)$  : The costs of employees' knowledge transfer. It refers to the fact that the knowledge receivers pay out the amount of time, energy and opportunity cost during the knowledge transfer, and knowledge owners pay out time, energy and the loss of competitive advantages due to knowledge transfer and sharing.

$\pi(k)$  : The output of knowledge transfer and sharing to the organization.

As the employees' knowledge transfer and sharing is highly difficult to detect in the organizations, so the

organizations cannot pay out the incentive costs according to the volume of knowledge transfer, but can only to decide to take incentive methods or not. Employees' knowledge transfer costs and the outputs of knowledge transfer and sharing to the organization are the function of knowledge volume of the employees transfer and sharing. Therefore, the game matrix of the employees and the organization is as table1 shown:

From the above game matrix we can easily draw the conclusion that the best choice to the organization is not to pay out incentive costs and to the employee is not to transfer knowledge, so (not encourage, not transfer) is Nash equilibrium. Therefore, even the knowledge transfer and sharing is both good for the organization and employees, the game of the employees' knowledge transfer to the organizational knowledge still traps in the prisoner's dilemma.

### 3.2. Knowledge Sharing between the Employees

According to the situation of knowledge transfer and sharing in the organization the knowledge can be divided into two categories: the transferable knowledge and the non-transferable knowledge [13,14].

Assuming the employee A and B process the game of knowledge sharing in the organization, they are all rational players and the knowledge they transferred and shared is useful to the organization. Each of them has two choices, to share knowledge or not to share knowledge. Assuming:

$U_{A1}, U_{B1}$  : Respectively represent the non-transferable knowledge of the player A and B.

$U_{A2}, U_{B2}$  : Respectively represent the transferable knowledge of the player A and B.

$U_{A3}, U_{B3}$  : Respectively represent the synergy value of the player A and B in the knowledge sharing. Synergy value is the newly acquired knowledge value by virtue of the fusion of special knowledge when both sides share their knowledge [15].  $U_{A4}, U_{B4}$  : Respectively represent the multiplication value of the player A and B in the knowledge sharing. Multiplication value is the capacity of the knowledge receivers using the knowledge suppliers' knowledge to improve their own competitive advantages [15].

**Table 1. The game matrix of the employees and organization**

Organization	Employee	
	Transfer	Not Transfer
Encourage	$(\pi(k) - W, W - C(k))$	$(-W, W)$
Not Encourage	$(\pi(k), -C(k))$	$(0, 0)$

$U_{A5}, U_{B5}$  : The negative utility to the knowledge suppliers due to the knowledge transfer and sharing.

$\alpha_A, \alpha_B$  : Respectively represent the knowledge absorptive capacity coefficient of the player A and B.

Therefore, the profit matrix as Table 2 shown:

M. Levy. find that low synergy value is not related to high multiplication value, but synergy value is linked to the negative effects through a group of organizations' empirical research. Therefore, the game results showed two cases as follows:

a. Prisoner's dilemma. When the generated synergy value of the knowledge sharing is lower than its negative effects, that is when  $U_{A3} < U_{A5}$  and  $U_{B3} < U_{B5}$ , (not-sharing, not-sharing) is the only Nash equilibrium. The two players are in Prisoner's dilemma.

b. Trust game. When the generated synergy value of the knowledge sharing is higher than its negative effects, that is when  $U_{A3} > U_{A5}$  and  $U_{B3} > U_{B5}$ , there are two equilibriums (sharing, sharing) and (non-sharing, non-sharing). At this point, there is first-mover advantage. The player adopts a wait-and-see attitude, whose revenue maximization strategy is making the same choice to the other.

From the above process of game we can see, the best choices of the play A and B are all non-sharing, so (non-sharing, non-sharing) is a Nash equilibrium. The dilemma of knowledge sharing reflects the contradictions of individual rationality and collective rationality, and the appeared optimal strategy to individuals will make the entire organization in a disadvantageous position. Therefore, the organizations need to establish an effective mechanism of knowledge sharing to break this "dilemma".

## 4. Research on Incentive Mechanisms in Knowledge Transfer and Sharing

### 4.1. Subject and Knowledge Potential in the Knowledge Transfer and Sharing

In organization the knowledge quantity of the employees is different. Someone has the low potential knowledge, and someone has the high potential knowledge. Therefore, knowledge transfer and sharing is a tripartite game process of the knowledge suppliers, receivers and organizations. The position of the subject is not all the same in the

**Table 2. The game matrix of knowledge sharing**

Player A	Player B	
	Sharing	Not-sharing
Sharing	$(\alpha_A U_{B2} + U_{A3} + U_{A4} - U_{A5}, \alpha_B U_{A2} + U_{B3} + U_{B4} - U_{B5})$	$(-U_{A5}, \alpha_B U_{A2} + U_{B4})$
Not-sharing	$(\alpha_A U_{B2} + U_{A4}, -U_{B5})$	$(0, 0)$

Knowledge potential	High	Strong willingness to share knowledge I High degree of organizational loyalty	Weak willingness to share knowledge IV Low degree of organizational loyalty
	Low	Strong motive of receiving knowledge II High degree of organizational loyalty	Weak motive of receiving knowledge III Low degree of organizational loyalty

**Figure 1. Ypes of the knowledge employees**

knowledge sharing process. The employees can be divided into four types according to the knowledge potential level and mobility of the employees as Figure1 shown. The organization must recognize the quadrant of the employees in the incentive mechanism design of knowledge transfer and sharing, the adopted incentive mechanism to different employees should be different.

#### 4.2. Organizational Knowledge Transfer and Sharing Flow Map based on the Knowledge Potential

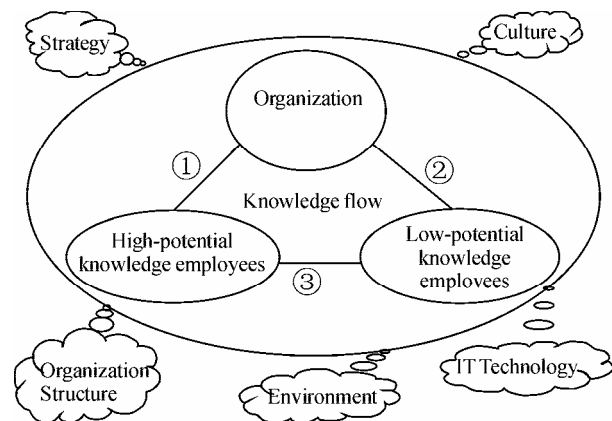
In the organizational knowledge transfer and sharing, the organizations are clients who must make games with one or more agents namely employees. As the employees have different knowledge quantity, the design of incentive mechanism to high-potential knowledge employees and low-potential knowledge employees should be different. Because the key to the knowledge sharing lies in the willingness of high-potential knowledge employees' sharing his personal core knowledge, the mechanism design to the high-potential knowledge employees is the key.

Figure 2 shows the knowledge transfer and sharing flow map between the employees and organization. It can be seen there are two main paths to achieve the employees' knowledge transferring to the organizational knowledge, respectively ① and ② of Figure 2. So in this process the key knowledge source is the high-potential knowledge employees. The most important point to the knowledge transfer and sharing is the willingness and ability of the employee dispersing his individual knowledge.

After breaking the "prisoner's dilemma" of high-potential knowledge employees' knowledge transferring to the organizational knowledge, the knowledge sharing between the employees has become simple. In Figure 2, if the high-potential knowledge employees are voluntary to transfer and share knowledge, not only path ① can be achieved, but also can promote path ② completed through the high-potential knowledge employees transferring knowledge to the low-potential knowledge employees.

#### 4.3. The Premise and Assumption

If the employees in the organization could transfer and share



**Figure 2. Organizational knowledge transfer and sharing flow map**

the knowledge, it will have more complex and higher value knowledge [16], and will greatly increase the output. However, the knowledge owner does not always have the will to share the knowledge, such as the individuals refuse to transfer and share knowledge with others in order to maintain the expert influence power, and could not have a reasonable reward during the knowledge sharing.

In the organization, the information is asymmetrical, and the output is the result of joint efforts of all the employees. So the research of knowledge transfer and sharing is usually based on the following basic assumptions:

First of all, the economic man hypothesis. The employees are all the rational egocentrics, and they all have a very good definition of preferences. In the face of a given condition, they will maximize their own preferences.

Secondly, the employees have greater autonomy in the process of knowledge transfer and sharing. They can choose the best knowledge independently to achieve a certain output.

Third, there is information asymmetry between the knowledge employees and the organization. The employees are at a comparative advantage in the information, and it is incomplete information game between the employees and the organization.

Transferring and sharing knowledge can improve productivity and competitiveness in the organization. But the employees contribute their knowledge will reduce the economic interest, reputation and status due to the monopoly of knowledge. Therefore, the employees often hope to share other people's knowledge and not to contribute their knowledge.

In addition to meet the three basic assumptions above, the paper also subject to the following economic assumption:

1) Assuming that the employees are risk-neutral in the organization.

2) The output is not subject to external uncertainty factors, and is the function of the level of all the employees' knowledge transfer and sharing.

## 5. Research of the Incentive Mechanism in the Knowledge Transfer and Sharing Based on the Knowledge Potential

The incentive contract of knowledge transfer and sharing in the organization is under the asymmetric information. Assuming that the clients cannot observe the action choice of the agents which is the shared knowledge volume  $k$  and external variable  $k$ , only can observe the output  $\pi$ . At this time, the agent's incentive compatibility constraint is contributing, because regardless of how the clients reward and punish the agents, the agents always choose the action maximizing their own utility level. In other words, the clients cannot use "mandatory contract" to force the agents to choose the clients' favourite action, but only through the incentive contract to induce the agents to choose the clients' favourite action. The clients' problem is to choose incentive contract which meets the agents' participation constraint and incentive compatibility constraint simultaneously to maximize their own expected utility functions.

Assuming that the minimum possible value of the shared knowledge volume is  $k_L$ , the maximum possible value is  $k_H$ , and the minimum possible value of the organizational output produced by the knowledge transfer and sharing is  $\pi_{\min}$ , the maximum possible value is  $\pi_{\max}$ . If the agents actively share knowledge,  $k = k_H$ , distributing function and distributing density of  $\pi$  is respectively  $F_H(\pi)$  and  $f_H(\pi)$ ; if the agents do not actively share knowledge,  $k = k_L$ , and distributing density of  $\pi$  is respectively  $F_L(\pi)$  and  $f_L(\pi)$ . Supposed to regard the output  $\pi$  of knowledge transfer and sharing to the organization as a random variable,  $\pi(k, \theta)$  changes along with  $k$  in the same direction. That is to say the more knowledge quantity the agents share, the higher outputs are produced. Distributing function satisfies the first-order stochastic dominance condition, i.e. with all  $\pi \in [\pi_{\min}, \pi_{\max}]$ ,  $F_H(\pi) \leq F_L(\pi)$  and strict inequality comes into existence at least to some  $\pi$ . In other words, compared to not actively sharing knowledge, the actively sharing knowledge has the higher probability of producing higher output.

Further assuming  $C(k_H) > C(k_L)$  which means the cost of actively sharing knowledge is higher than the one of not actively sharing, and the clients wish the agents choose  $k = k_H$ . At this time the agents' incentive compatibility constraint means  $\partial W / \partial \pi \neq 0$ . To enable the agents consciously have sufficient enthusiasm for choosing to share knowledge, the clients must abandon

Pareto optimal risk-sharing contracts. The clients' problems are to choose incentive contract  $W(\pi)$  to resolve the following optimization problem:

$$\max \int v(\pi - W(\pi)) f_H(\pi) d\pi \quad (1)$$

$$\text{s.t. } (IR) \int u(W(\pi)) f_H(\pi) d\pi - C(H) \geq \bar{u} \quad (2)$$

$$(IC) \int u(W(\pi)) f_H(\pi) d\pi - C(H) \geq \int u(W(\pi)) f_L(\pi) d\pi - C(L) \quad (3)$$

$\bar{u}$  is the agents' reservation utility. The most important result of principle-agent model is that it can predict what kind of observational variables should enter incentive contract. So in the design of incentive mechanism the key is to bring the organizational output  $\pi$  produced by the knowledge transfer and sharing to the incentive contract. This is because  $\pi$  is the function of the shared knowledge volume  $k$ , we can indirectly observe  $k$  through observing the value of  $\pi$ .

We can imagine that, if the promise the organizations make to encourage the high-potential knowledge employees for knowledge sharing and the commitment is believed. Let incentive cost  $W = W(\pi)$ , i.e. the incentive cost  $W$  which the organization pays to the high-potential knowledge employees is the function of the organizational output  $\pi$  produced by knowledge sharing. To be simple, consider two extreme cases, if the employees do not transfer knowledge, then  $k = 0$ ,  $\pi(k) = 0$ ,  $W(\pi(k)) = 0$ . If the employees transfer knowledge, then  $W = W(\pi(k))$  and deem  $\pi$  increases along with  $k$  increases,  $W$  increases along with  $\pi$  increases. Put these values to the game matrix of the employees and the organization in table1, and then we get the improved game matrix as shown in Table 3:

From the above improved game matrix we can see that when the organization is committed to encourage and the incentive cost  $W = W(\pi(k))$ , the best choice of the high-potential knowledge employees is knowledge sharing. In this way, the organization can monitor the high-potential knowledge employees' knowledge sharing without any actions.

The above incentive mechanism is designed to the high-potential knowledge employees, and the incentive mechanism for the high-potential knowledge employees should satisfy the following constraints. Firstly, the

**Table 3. The improved game matrix**

Organization	Employees	
	Transfer	Not Transfer
Encourage	$(\pi(k) - W(\pi(k)), W(\pi(k)) - C(k))$	(0,0)
Not Encourage	$(\pi(k), -C(k))$	(0,0)

high-potential knowledge employees should undertake some organizational operational risk, so they have indirectly the future profit view to the organization; Secondly, the rewards of the high-potential knowledge employees should be changed along with the organizational profits; thirdly, when the organization evaluate the knowledge sharing performances of the high-potential knowledge employees, it should add some exogenous variables which can be observed.

## 6. The Incentive Mechanism Application to "Prisoner's Dilemma" in the Organization Knowledge Transfer and Sharing

The specific encouragement can adopt equity incentive. As a consequence of the character of equity incentive, it very well suites to the required condition of the high-potential knowledge employees' encouragement. Taking equity incentive to the high-potential knowledge employees makes them undertake some organizational operational risk, so they can put their own profits closely with the organizational profits. It introduces repeated games virtually. In repeated games the players will consider that their current actions will influence other players' future actions, i.e. consider the current profits as well as the future profits. So the knowledge sharing is possible in repeated games. And the rewards of the high-potential knowledge employees will be changed along with the organizational profits. When the organizational profits increase, their profits will increase correspondingly. It makes there are exogenous variables which can be observed in the evaluating of the high - potential knowledge employees' knowledge sharing.

From the above game of knowledge sharing we know there does trust game exist. When the generated synergy value is higher than its negative effects, there are two equilibriums, i.e. (sharing, sharing) and (non-sharing, non-sharing). At this point, the player adopts a wait-and-see attitude. If the other player chooses sharing, then sharing. If the other player chooses non-sharing, then the most optimum choice is non-sharing. As by improving their own knowledge through the knowledge sharing process, the low-potential knowledge employees can get more proceeds than the costs of learning they pay, and the organizations' adopting equity incentive to the high-potential knowledge employees makes their optimal choices are knowledge sharing, so the low-potential knowledge employees' optimal choices are knowledge sharing. The more proceeds of the low-potential knowledge employees' learning knowledge, the stronger learning motivation will be.

The organization can adopt appropriate bonus incentive to the low-potential knowledge employees to make sure they could not only enhance their own value when actively participating in knowledge sharing, but also enjoy the utility of economic incentive. The organization can take equity incentive after the low-potential

knowledge employees' knowledge values arise to the high potential. In this way organization cannot only save the incentive costs of organizational knowledge transfer and sharing, but also can reduce the unfairness sense among the employees who have different knowledge potential in the process of knowledge transfer and sharing.

## 7. Conclusions

In virtue of the complementary knowledge, the knowledge transfer and sharing is beneficial to every employee, but the knowledge game between the organization and the employees as well as between the employees are still trapped in prisoner's dilemma. By encouraging the high-potential knowledge employees to transfer and share their core knowledge voluntarily to drive the low-potential knowledge employees to participate in knowledge sharing to obtain their own additional utility we can make the employees and organization achieve a win-win situation. Knowledge sharing incentive should be a step-by-step process. Excessive incentive will enhance the organizational costs and create unsatisfied sense. Little incentive cannot inspire the employees' enthusiasm. Analyzing the game process of knowledge transfer and sharing can find that the application of principle-agent into infinitely repeated game, increasing the employees' expected utility of knowledge transfer and sharing, avoiding the employees' short-term actions and strengthening the consistent interests between the organization and the employees will play a positive role to promote the knowledge transfer and sharing.

As the knowledge measurement is extremely difficult, the paper measures the sharing knowledge through the output, and proposes an equity incentive mechanism to the knowledge transfer and sharing in the organization, but do not have a deep exploration to the mode of equity incentive. Further research includes the empirical test of analysis of this paper and making the variables of influencing knowledge transfer and sharing endogenous.

## 8. Acknowledgements

The original form of this paper has been appeared at the Journal Publication Meeting (2007), September 24, 2007, Toronto, Canada. This research has been partially supported by a grant from National Natural Science Foundation of China (#70501030, #70621001) and Beijing Natural Science Foundation (#9073020).

## REFERENCES

- [1] H. Tao and C. J Dai, "Game analyses of organization's tacit knowledge," *Journal of Information*, Chinese, 2006.
- [2] B. Kought and U. Zander, "Knowledge of the firm, combinative capabilities, and the replication of technology," *Organization Science*, 3: pp. 383-397, 1992.

- [3] K. Mertins, P. Heisig, and J. Vorbeck, "Knowledge management concepts and best practices," Beijing, Tsinghua University Press, 2004.
- [4] T. H. Davenport and L. Prusak, "Working knowledge," Boston MA, Harvard Business School Press, 1998.
- [5] D. D. Wang, "The complementariness of knowledge along time and space and related economics," *Economic Research Journal*. Chinese, 1997.
- [6] R. M. Grant, "Toward a knowledge-based theory of the firm," *strategic management journal* 17 (special issues), pp. 109–122, 1996.
- [7] A. Robert, "More effective choice in the prisoner's dilemma," *Journal of Conflict Resolution*, (24): pp. 379–403, 1980.
- [8] N. M. Dixon, "Common knowledge: How company thrive by sharing what they know," Harvard Business School Press, Boston, MA 2000.
- [9] M. M. Lynne, "Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success," *Journal of Management Information System*, 18 (1), 2001.
- [10] J. W. Wu and H. Y. Xuan, "A game-theoretic analysis of knowledge barter betwe chinese.en staffs within firm," *Studies in Science of Science*, 2004.
- [11] K. J. Arrow, "The economic implications of learning by doing," *Review of Economic studies*, 29, pp. 155–173, 1962.
- [12] P. M. Romer, "Increasing returns and long run growth," *Journal of political Economy*, 94(5), pp. 1002–1037, 1986.
- [13] J. H. Chen and C. Lin, "The process of knowledge sharing and game analysis in the virtual enterprise," *Journal of Information*, Chinese, 2005.
- [14] L. L. Zhang, J. Li, and Y. Shi, "Study on improving efficiency of knowledge sharing in knowledge-intensive organization," *WINE 2005*: pp. 816–825, 2005.
- [15] C. Loebbecke and P. C. van Fenema, "Knowledge transfer under coopetition," 2001.  
<http://www.bi.no/dep2/infomgt/wg82-86/proceedings/loebbe1.pdf>.
- [16] I. Nonaka and H. Takeuchi, "The knowledge creating company: how japanese companies create the dynamics of innovation," New York: Oxford University Press, 1995

# Combining Keyword Search Advertisement and Site-Targeted Advertisement in Search Engine Advertising

Li Chen

Department of Operations and Information Management, School of Business, University of Connecticut  
Email: lchen@business.uconn.edu

Received June 3<sup>rd</sup>, 2008; revised November 17<sup>th</sup>, 2008; accepted November 27<sup>th</sup>, 2008.

## ABSTRACT

Internet advertising has seen a strong growth in recent years and search engine advertising has played an important role in that growth. Search engines continue to expand their business by providing new options of advertising. For example, Google provides a new advertisement mechanism based on cost-per-thousand-impression (CPM) payment called "site-targeted advertisement" in addition to the famous "keyword search advertisement" based on cost-per-click (CPC). While keyword search advertisement is a cost-efficient way of advertising, site-targeted advertisement provides a quicker alternative to expose the ad to a mass population at a higher expense. This paper studies a mixed strategy of optimization by combining these mechanisms to exploit their corresponding advantages. We set up a general model to find optimal starting and ending times for both methods. Closed-form solutions are calculated for two applications: (1) Advertising of perishable information with click-based revenue only; and (2) Site-targeted advertisement first and keyword search advertisement last. Comparative static analysis provides an analysis of properties of each application. We also develop a computational experiment based on Google AdWords to illustrate the application of the model.

**Keywords:** online advertising, search engine, keyword auction, cost per click

## 1. Introduction

An increasingly popular approach for firms to develop e-commerce is to advertise on search engines such as Google, Yahoo, and MSN. For most of people using Internet, search engine websites are a "must-see" when surfing online. A survey from *iProspect* shows that 56 percent of respondents used search engines at least once a day [1]. Although search engine is a relatively new medium for advertising compared with newspaper, TV and radio, search engines provide potential buyers and sellers a worldwide and 24-7 access to each other. Search engine advertising has seen a solid and continuous growth in recent years. During the second quarter of 2008, search engine advertising witnessed an increase of 24 percent from the second quarter of 2007 to reach \$2.5 billion, around 44 percent of total US online advertising spending [2]. The search engine advertising market can be seen as a duopoly. According to comScore.com, Google took 57 percent of the market in September 2008 and Yahoo took 23.7 percent [3].

In this paper, we base our research on advertising mechanism of Google AdWords, the advertising program of the largest player in this market. Currently Google AdWords provides two types of advertisements: keyword search advertisement and site-targeted advertisement [4]. Keyword search advertisement (hereinafter as "keyword-ad") refers to advertisements that appear side by side with search results on the Google web pages. Advertisers who are interested in putting this type of ad on Google web

pages need to participate in a keyword-ad campaign and win ad slots (see a simplified process in Figure 1).

For each ad campaign, Google adopts a modified second-price auction mechanism to allocate the ad slots. In ranking the advertisements, Google not only considers the bid amount but also the quality of the advertisements. Under this auction mechanism, an advertiser with a higher ranking wins an ad slot but only needs to pay the necessary amount to rank over the advertiser with the next highest ranking. Since keyword-ad is based on a cost-per-click (CPC) payment mechanism, an advertiser only needs to pay for every click on the sponsored link. If no click ever occurs, no payment will be charged. (See Figure 2 for an example of keyword-ad).

Google AdWords began Site-targeted advertisement (hereinafter "site-ad") in April 2005, allowing advertisers to choose individual sites in the Google network where they would like their ads to appear. (see Figure 3 for an example).

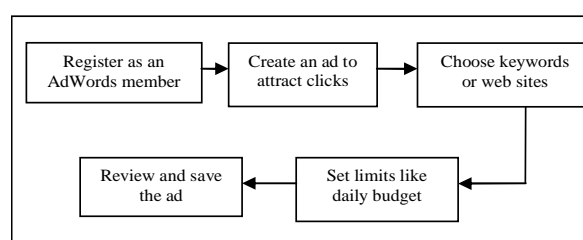


Figure 1. Process flow of ad campaign setup

Before the site-ad mechanism was introduced, the only way to advertise through Google was through CPC-based keyword-ad auction. This method was simple and easy to apply, but it had two drawbacks: (1) only visitors who search information through Google had a chance to see the ad and click on it; (2) click fraud may increase the advertising cost [4]. After the introduction of site-ads, Google expected to provide AdWords users with more choices of location of their ads and increase Google's revenues at the same time. With site-ads available, advertisers can be seen more widely over the Internet. The allocation of ad slots in a site-ads auction also uses the modified second-price auction mechanism. However, unlike a keyword auction, advertisers are charged based on a cost-per-thousand- impressions (CPM) payment method rather than cost-per-click (CPC) method. Each time an ad is displayed on a web page, the advertiser will

be charged [5].

With the presence of site-ads, advertisers can realize their revenues through CPC-based keyword-ads, CPM-based site-ads, or both. Given these choices, our goal in this paper is to study whether a mixed strategy of combining the two ad mechanisms will help advertisers increase their revenues. We develop a general budget- constrained, nonlinear optimization model to maximize advertisers' revenues using this "mixed strategy". In the model, we determine starting and ending points for the time intervals during which keyword-ad and site-ad advertising campaigns should hold. These intervals might overlap. After formulating the optimality conditions of the model, we concentrate on the study of two particular scenarios: (1) advertising of perishable information with click-based revenue only; and (2) site-ads first and keyword-ads last.

Web Images Maps News Shopping Gmail more ▼

Google

Google Adword Search Advanced Search Preferences

Web Results 1 - 10 of about 7,550,000 for Google Adword. (0.17 s)

**Google AdWords**  
adwords.google.com Advertise on Google In Just Minutes Instant Traffic. Immediate Results. Sponsored Link

Welcome to AdWords  
Google's advertising platform offering both cost-per-click and cost-per- impression pricing for advertisements served on Google.com and partner sites.  
adwords.google.com/ - 24k - Cached - Similar pages

Keyword Tool Google Analytics  
Help Learning Center  
Accounts Advertising Professionals  
Advertising Programs Google Maps  
More results from google.com »

**AdWords Help Center**  
Search AdWords Help articles, user discussions, and more. ... create account · conversion

**Profit from AdWords**  
PPC Management That Ties Ad Spend to Profits! Isn't That the Point?  
www.WorkingPlanet.com

**Certified AdWords Experts**  
Special Offer: No Setup Fee & \$100  
In Free Click Funding! 877.226.9  
www.SwiftSolution.com

**"Beat The AdWords System"**  
Access 100 Million People in 10 Min  
- The Definitive Fast Start Guide  
AdwordsStrategy.com

*Keyword-ad*

Figure 2. Example of Google AdWords keyword search ads

TimesPeople Share and Discover the Best of NYTimes.com 7:57 PM Log in or Register No, thanks

Survival  
Vital Signs: Nostrums: Supplements Seem to Be No Help to Knees

HOME & GARDEN »  
In Turkey, an Island Retreat  
In the Region | New Jersey: Honors for Restored Wright House  
The Feud and the Fakes

EDUCATION »  
With Free Bikes, Challenging Car Culture on Campus  
Officials Meet With Families About Virginia Tech Shootings  
Safety Concerns Eclipse Civic Lessons as Schools Cancel Classes on Election Day

FASHION & STYLE »  
Vital Signs Under the Viaduct  
Completely Unplugged, Fully Green  
Free to Be His Own Buckley

REAL ESTATE »  
Mixing Money and Family  
Living In | Koreatown: Exotic Flavor, Beyond Just the Food  
The Hunt: Life After Rent Control

AUTOMOBILES »  
Behind the Wheel | 2009 Honda Fit: Sized for the Times  
Wheelspin: Car Insurance May Not Cover You at the Track  
Handlebars: Borrowing From a Sporty Sibling

Pretty Smart  
The Nation: In the Conservative Commentariat, Unease

MAGAZINE »  
Working for the Working-Class Vote  
Reaching an Autistic Teenager  
What's Really Wrong With the Price of Oil

T MAGAZINE »  
Bonnie Fuller on the War of the Ritchies  
Perfume Review: Danielle by Danielle Steel  
Art: London's Frieze Fair

Privacy Policy Events

Add New York Times headlines to your site  
Add New York Times RSS Feeds NY RSS

Get home delivery of The New York Times, as little as \$3.45 a week.

The New York Times Small Business GO

The New York Times Small Business Summit nytimes.com/bizsummit

Ads by Google what's this?

**Want Repo Motorcycles?**  
The secret to getting motorcycles dirt cheap - must see to believe!  
Car-Auction.com/Motorcycle\_Auction

**My 3 Free Credit Reports**  
View your Free Credit Reports and Credit Scores from all 3 bureaus.  
FreeCreditReportsInstantly.com

**Batteries To Go**  
Sealed Motorcycle Batteries With Free Shipping  
www.eBatteriesToGo.com

*Site-ad*

Figure 3. Example of google adwords site-targeted ads (New York Times website)



In the first scenario, we assume that the advertiser only makes revenues from people visiting the advertiser's web site, which is generated by clicks. Also, we assume that the content on the web site is perishable and its value decreases over time. The scenario is typical of newspaper websites, where just looking at the ads linking to a news story page does not generate revenues for the newspaper.

In the second scenario, we assume that the advertiser starts using site-ads and then switches to keyword-ads without overlapping. This situation applies to advertising of new products or services where the advertiser wants to aggressively expose her ads to a large population of potential customers, even though it might be more expensive.

We find closed-form solutions for both scenarios, and then we perform comparative static analysis to study how the changing model parameters affect the solution. Based on this analysis, we provide some managerial insights under two conditions: non-binding budget and binding budget. Under the condition of non-binding budget, for a manager facing a high content depreciation rate and low number of visits and click-through rate, it would be a better strategy to open new advertising campaigns and compete for more popular keyword-ads or site-ads on a more popular web site or both.

Under the condition of binding budget, for the first scenario, advertisers who see a higher depreciation rate, a lower number of visits of both types of ad, or a lower click-through rate of both types of ads would have to extend the keyword-ad duration. For the second scenario, we find that advertisers who see a higher number of visits on site-ad would have to move the switch time earlier, and they would have to delay the switch time while seeing a lower number of visits of site-ad.

The remainder of this paper proceeds as follows. Section 2 provides a review of recent literature in related areas. Section 3 discusses the general model for the mixed strategy of combining both types of ad mechanisms. Section 4 details the analysis of the two scenarios discussed above and offers numerical examples. Section 5 contains the conclusion. In the Appendix, we present the optimality conditions for the general model.

## 2. Literature Review

Search engine advertising has become a hot topic and attracted significant research interests [4,6,7,10,14,15,16]. There are multiple streams of related literature studying search engine advertising. In particular, Jansen and Mullen [6] provided an extensive review discussing issues such as auction properties, competitive landscape, how to rank ads, how to set slot prices, payment mechanisms, etc. Even though they considered three types of participants: content providers, search engines and searchers; we only consider the point of view of the content providers who want to optimize their revenues subject to a budget constraint. We review the following areas that are more related to our research.

### 2.1. Advertising Allocation Mechanisms

In search engine advertising, auction is the major mechanism to allocate ad slots. Feng and Bhargava [7] used simulation to study four allocation mechanisms of ad slots, including those used by Overture (Yahoo) and Google AdWords. They found that performance of mechanisms used by Yahoo and Google is better under certain scenario, and their performance also depended on the degree of correlation between providers' willingness to pay and relevance. Edelman *et al.* [8] focused on "generalized second-price" (GSP) auction, in which the bidder who wins an ad slot only needs to pay the next highest bidders' price plus an increment. GSP is attractive to search engines because it helps to maximize profit. In addition, Feng *et al.* [9] developed a simultaneous pooled auction (SPA) mechanism and showed that using reserve price in SPA significantly increased a search engine's revenue.

### 2.2. Payment Mechanism

There are multiple payment methods in search engine advertising, such as cost-per-click, cost-per-thousand-impression (CPM) and cost-per-action [10]. While CPM comes from traditional print media, cost-per-click and cost-per-action are based on the search engines' and advertisers' measurement. Hu [11] applied the economic theory of incentive contracts to show that performance-based pricing models improve effectiveness of advertising campaigns. After investigating the implementation of paid placement strategies, Weber and Zheng [12] found that revenue-maximizing search engines ranked ads on a weighted average of relative performance and bid amount. Kumar *et al.* [13] studied an interesting problem of the optimal advertising schedule in ads slots of the web sites based on a hybrid pricing model.

### 2.3. Bidder Strategies

Advertisers need to determine how to respond to competition in auctions. One common observation is cycle bidding, where bidders revise their bids to compete for ad slots (Edelman and Ostrovsky 2007 [14] and Zhang and Feng [15]). One possible explanation is gap jamming, which refers to the behavior of bidders' raising bids to some point just below competitors' bids. When gap jamming is present, competitors will be charged for the highest possible amount. Another important strategy for bidders is how to allocate their funds across advertising campaigns such as keyword-ad campaigns. Özlük and Cholette [16] suggested a model for advertisers who have a fixed daily budget limit to select keywords to maximize productivity and then determine the bid for each keyword selected.

### 2.4. Optimization in Internet Advertising

Internet advertising also faces optimization problems. For example, Dewan *et al.* [17] found a tradeoff between ads and content for web sites: more ads generate more revenue but may turn viewers off. Their findings suggest that websites put fewer ads and more content, and get compensated for by future profits. Another example comes from Fruchter and Dou [18], which studied how to dy-

namically assign budget of banner ads between the two types of portals (generic vs. specialized).

Our work is different from previous research in that we consider a mixed keyword-ad and site-ad strategy, we only consider the content provider's point of view, we provide a new budget-constrained, nonlinear optimization model to maximize advertisers' revenues, and we study two particular scenarios that have not been considered before: advertising of perishable information with click-based revenue only, and site-ad first and keyword-ad last.

### 3. General Model

In this paper, we develop an optimization model of a mixed strategy which combines the keyword-ad and site-ad mechanisms to help advertisers maximize their revenue. Keyword-ad is a cost-efficient advertising strategy because advertisers are only charged for clicks on their ads. Therefore, we regard keyword-ad as a "waiting strategy" because only visitors with relatively strong interest will search information through search engines and advertisers have no control of the number of search requests and clicks based on these visits. For site-ad, we regard it as a "showing strategy" because after putting an ad on a targeted website, viewers of the website are exposed to that site-ad even if they do not plan to search that information. Compared with keyword-ad, site-ad is a more aggressive approach in terms of generating a large number of impressions in a short time, but the related expense is usually also higher.

In our model, we assume that there is fixed time period of length  $T$  during which the decision maker will decide the starting and ending times of each type of ad mechanisms. We denote by  $y_1$  and  $y_3$  the starting times of keyword-ad and site-ad respectively; and by  $y_2$  and  $y_4$  the ending times of keyword-ad and site-ad, respectively. Notice that  $0 \leq y_1 \leq y_3 \leq T$ ,  $0 \leq y_2 \leq y_4 \leq T$ , and the two intervals may overlap. We denote the time vector of decision variables by  $Y = (y_1, y_2, y_3, y_4)$ .

Our model uses a set of exogenous smooth functions as described in Table 1. The idea is to use these functions to capture the behavior of the searchers as well as the payment mechanism of the advertiser. We use subscript "1" to refer to CPC-based keyword-ads and "2" to refer to CPM-based site-ads. Consistent with previous work [7], we assume that click-through rate only depends on the location of ad slot on the web page. Also we assume that the click-through rate of keyword-ad is higher than that of site-ad. This is because people who search for the information are more likely to click on the ad. Although there maybe a higher number of clicks generated from site-ads, the click-through rate may not be as high as that of keyword-ad. For example, if 100 people search information of "2008 Olympics", and there are 30 people click on the keyword-ad, then the click-through rate is 30 percent. At the same time, there maybe 1000 people see the site-ad of

**Table 1. Notation used in the general model**

Ad Type	Visits in time period $t$	Payment per ad in time period $t$	Click-through rate in time period $t$
Keyword	$S_1(t)$	$P_1(t)$	$l_1(t)$
Site	$S_2(t)$	$P_2(t)$	$l_2(t)$

the same content but only 100 people click the ad. Although the site-ad generates more clicks ( $100 > 30$ ), the click-through rate is only 10 percent, smaller than that of keyword-ad (30 percent).

Our goal is to maximize advertiser's total revenue  $R(Y)$  through the period for a given choice of starting and ending times  $Y = (y_1, y_2, y_3, y_4)$ . In addition, we assume that the advertiser is subject to a budget limit  $B$ , that is, the advertiser cannot spend more than  $B$  through the whole period.

We compute total revenue  $R(Y)$  as follows:

$$R(Y) = r_1 \left\{ \int_{y_1}^{y_2} S_1(t) l_1(t) dt + \int_{y_3}^{y_4} S_2(t) l_2(t) dt \right\} + r_2 \left\{ \int_{y_1}^{y_2} S_1(t) dt + \int_{y_3}^{y_4} S_2(t) dt \right\}$$

In the definition of total revenue,  $S_1(t)$  denotes the number of visits on the keyword-ad in time period  $t$  and  $S_2(t)$  denotes to the number of visits on site-ad in time period  $t$ .  $S_1(t)l_1(t)$  denotes the number of clicks generated from keyword-ad in time period  $t$ , and  $S_2(t)l_2(t)$  denotes to the number of clicks generated from site-ad in time period  $t$ . We take the integral to calculate the total number of clicks and total number of visits during the whole ad campaign duration. Finally, we multiply by  $r_1$  and  $r_2$  the integrals, respectively, to obtain the total revenue.

To account for the budget constraint, we use the expression:

$$\int_{y_1}^{y_2} S_1(t) l_1(t) p_1(t) dt + \int_{y_3}^{y_4} S_2(t) p_2(t) dt$$

where  $S_1(t)l_1(t)p_1(t)$  denotes the cost of keyword-ad and  $S_2(t)p_2(t)$  denotes the cost of site-ad. As before, the integrals are used to calculate the total amount of payment.

The resulting model is the following:

$$\begin{aligned} \text{Max}_Y \quad R(Y) = & r_1 \left\{ \int_{y_1}^{y_2} S_1(t) l_1(t) dt + \int_{y_3}^{y_4} S_2(t) l_2(t) dt \right\} \\ & + r_2 \left\{ \int_{y_1}^{y_2} S_1(t) dt + \int_{y_3}^{y_4} S_2(t) dt \right\} \end{aligned} \quad (1)$$

subject to:

$$0 \leq y_i \leq T, \quad i = 1, \dots, 4 \quad (2)$$

$$\int_{y_1}^{y_2} S_1(t) l_1(t) p_1(t) dt + \int_{y_3}^{y_4} S_2(t) p_2(t) dt \leq B \quad (3)$$

Constraint (2) implies that the time points should be nonnegative and do not exceed the length of the decision period; and constraint (3) is the firm's budget constraint  $T$ .

We solve this nonlinear optimization problem by using Lagrangian multipliers, Karush-Kuhn-Tucker conditions and Leibniz rules (see the details in Appendix).

## 4. Two Specific Applications

In this section, we discuss the application of the general model on two specific applications of search engine advertising: (1) Advertising of perishable information with click-based revenue only; and (2) Site-targeted ad first and keyword search ad last.

### 4.1. Advertising of Perishable Information with Click-Based Revenue Only

For the first application, advertisers are advertising perishable information and their revenue comes only from

clicks generated. This scenario is common among online content providers such as online newspaper websites who have a strong incentive to attract visitors to their websites.

According to the Newspaper Association of America (NAA) ([www.naa.org](http://www.naa.org)), the audience of online newspaper websites reaches 3.6 million per month in 2007 [19]. A new method of attracting visitors to online newspaper websites is to advertise latest news stories through search engines such as Google. For example, when there is a breaking news story such as "Powell endorses Obama", online newspapers such as *New York Times* put keyword-ads on Google that might lead to their latest news story online (see Figure 4). However, news stories, like seasonal products or fashion goods, depreciate in value fairly quickly after people hear enough of them and lose interest. Therefore, online content providers such as online newspaper are willing to maximize the influence of a news story before the "news story" loses value to audience.

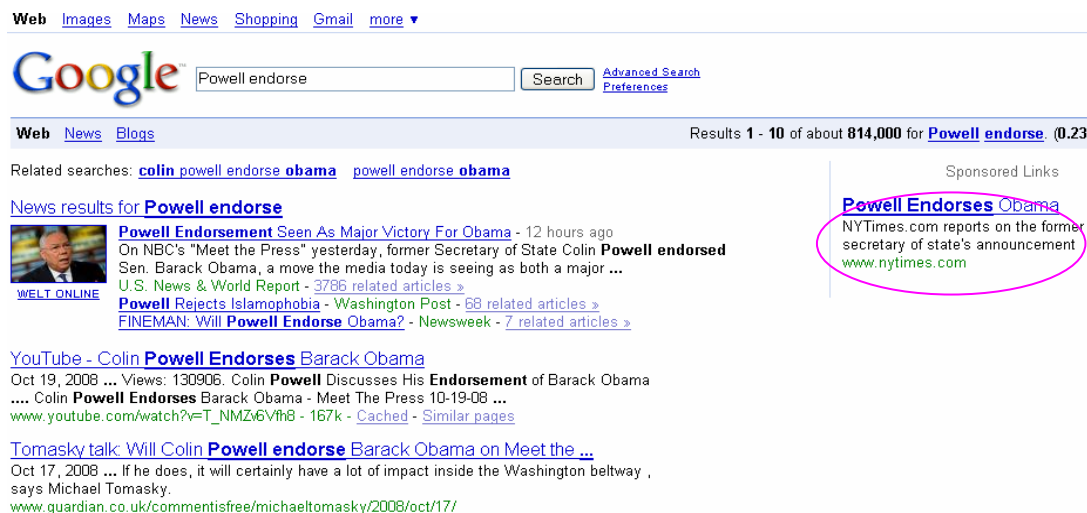


Figure 4(a). Example of New York Times' ad on Google



Figure 4(b). New York Time's news story linked by the keyword-ad in Figure 4(a)

As mentioned earlier, it is difficult for a keyword-ad strategy to rapidly generate the expected level of attention because it is a “waiting strategy” where advertisers have to wait people viewing their ads and clicks on them. In this scenario, a mixed strategy that combines site-ad with keyword-ad may be a better alternative. The rationale is that visitors of the targeted site are exposed to the ad of the news story. These visitors might not know the news or they know but have not an interest strong enough to go to Google.com to search “Powell endorses Obama”. However, they are likely to click the ad linked to the online newspaper to read the story. In other words, such a strategy is more likely to impress potential visitors whose search cost of news is relatively high but are willing to read the news when a link is in front of them.

Under this scenario, our main concern is to find the optimal ending time of both types of ads to maximize of the advertisers’ revenue given a budget constraint.

#### 4.1.1. Modified Mathematical Model

We make two modifications on the general model from Section 3 to meet requirements of this specific application: (1) we assume that advertisers want to begin both keyword-ad and site-ad at the beginning of the advertising campaign. Thus, we set the starting time of both keyword-ad and site-ad to zero, which implies that both types of ad will be adopted from the beginning; and (2) we assume that the advertisers in this specific application are interested only in the number of the clicks on their website. Impressions of both keyword-ad and site-ad on viewers will not bring value to the advertisers. For that reason, we only consider the first expression in the definition of  $R(Z)$  in (1).

We use  $z_1$  and  $z_2$  to denote the ending time of keyword-ad and site-ad respectively, and  $Z = (z_1, z_2)$  to denote the corresponding decision vector.

The general model is modified as follows:

$$\text{Max } R(Z) = r_1 \left\{ \int_0^{z_1} S_1(t) l_1(t) dt + \int_0^{z_2} S_2(t) l_2(t) dt \right\} \quad (4)$$

subject to:

$$0 \leq z_1, z_2 \leq T \quad (5)$$

$$p_1 \int_0^{z_1} S_1(t) l_1(t) dt + p_2 \int_0^{z_2} S_2(t) dt \leq B \quad (6)$$

We assume that  $S_1(t)$  decreases exponentially with a depreciation rate  $\theta$  so that  $S_1(t) = s_1 e^{-\theta t}$ , where  $s_1$  denotes to the number of initial visits. A higher value of  $\theta$  indicates a fast speed people lose interest on the information. We also assume  $S_2(t)$  is fairly stable in terms of the number of visits per unit of time so we can use average number of impressions  $s_2$  as an approximation. As before, the click-through rate is assumed to be relatively stable at  $l_1$  and  $l_2$  during the whole advertising period.

In this study, we only consider situations under which both  $z_1^*$  and  $z_2^*$  are positive. Using Karush-Kuhn-Tucker condition, we get closed forms for three optimal solutions  $(z_1^*, z_2^*)$  with economic meaning.

a)  $z_1^* = z_2^* = T$  In this case, the optimal ending time for both types of ads is the end of the advertising period, which implies that the budget constraint is not binding.

$$b) \quad z_1^* = T \quad \text{and} \quad z_2^* = \{B - \theta^{-1}(1 - e^{-\theta T})p_1 s_1 l_1\} / p_2 s_2.$$

In this case, the best choice is to stop the site-ad before the end of the advertising period, and let keyword-ad continue to the end.

$$c) \quad z_1^* = -\theta^{-1} \log\{1 - \frac{\theta(B - p_2 s_2 T)}{p_1 s_1 l_1}\} \quad \text{and} \quad z_2^* = T.$$

In this case, the best choice is to stop keyword-ad before the end of the advertising period and let site-ads continue to the end.

#### 4.1.2. Comparative Static Analysis

We investigate the impact of changes of parameter values under two conditions: whether budget constraint is binding or not. The rationale to discuss the scenario of non-binding budget is that for large online content providers such as New York Times, their interests are probably not saving money but fully utilizing the budget to generate attention and clicks, especially when a news story is still of interesting to public.

(1) The budget constraint is not binding

Let  $\tilde{B} = B - \theta^{-1}(1 - e^{-\theta T})p_1 s_1 l_1 - p_2 s_2 T$  denotes remaining budget. As mentioned above, we are interested to see how changes of parameters will influence advertisers’ revenue and remaining budget. Comparative analysis on parameters can shed lights on what decision to make (see Table 2).

We see that an increase in  $\theta$  leads to less revenue and a higher remaining budget, while an increase in  $l_1$ ,  $s_1$  and  $s_2$  leads to more revenue and less remaining budget. The result is intuitive because an increase in  $\theta$  implies that public’s interest on the online content such as a news story depreciates faster, and fewer clicks are generated. On the other hand, increases in  $l_1$ ,  $s_1$  and  $s_2$  imply that either more people are interested in the news story or more people are exposed to the site-ad or people who see the keyword-ad are more likely to click on the ad and observe the information linked to the ad. Obviously, all these changes will lead to more revenue and less remaining budget.

Using these results, a manager can determine which strategy to choose. For example, in case of a high depreciation rate and low number of visits and click-through rate, advertisers should bid aggressively so as to open new advertising

**Table 2. Comparative static analysis under non-binding budget**

Variable of interest	Effects of an increase in					
	$\theta$	$s_1$	$s_2$	$p_1, p_2$	$l_1$	$l_2$
$R$	-	+	+	N/A	+	+
$\tilde{B}$	+	-	-	-	-	N/A

campaigns and compete for more popular keyword-ads or site-ads on a more popular website or both.

(2) The budget constraint is binding

In this scenario, advertisers have consumed all their advertising budget resource before the end of advertising period (case (b) and case (c)). The difference is keyword-ad stops before the end of advertising period in case (c) and site-ad stops before the end of advertising period in case (b). Advertisers' interest here is what changing model parameters' influence will be. Table 3 shows the effect of parameters change on revenue, and optimal ending time of  $z_1^*$  in case (a) and  $z_2^*$  in case (b).

**Table 3. Comparative static analysis under binding budget**

Variable of interest	Effects of an increase in parameters					
	$\theta$	$s_1$	$s_2$	$p_1, p_2$	$l_1$	$l_2$
$R$	-	+	+	N/A	+	+
$z_1^*$ in case (a)	+	-	-	-	-	N/A
$z_2^*$ in case (b)	+	-	-	-	-	N/A

We find that under this scenario an increase in  $\theta$ ,  $l_1$ ,  $s_1$  and  $s_2$  have the same impact on revenue as in the non-binding budget case. As for the impact on duration of both types of ad, our analysis yielded two interesting findings: (1) lower revenue may happen after a longer advertising period. This result seems counter-intuitive. However, the rationale here is that if people lose interest on the news story faster, then keyword-ads fail to generate the expected number of clicks. However, since advertisers are charged based on clicks, the budget is consumed slower, which leads to a longer advertising period; (2) The click-through rate of site-ad has no certain impact on the duration. This is because site-ad is charged based on impression. Whether click-through rate is high or not will not influence the speed advertising budget is consumed.

The managerial insight here is that advertisers who see a higher depreciation rate, a lower number of visits of both types of ad, or a lower click-through rate of both types of ads would have to extend keyword-ad duration. On the other side, advertisers who see a lower depreciation rate, a higher number of visits, or a higher click-through rate would have to extend site-ad duration.



**Figure 5. Example of combining keyword-ad and site-ad**

We have observed real examples of using both types of ads (see Figure 5). However, it will be easier to show the implication of our model using real data from advertising campaigns.

To illustrate our model to advertisers who want to apply the “mixed strategy”, we use a numerical experiment to explain the general solution methodology. We put an ad of “Enjoy everyday in Shanghai<sup>1</sup>” which links to the top news story on entertainment in Shanghai to an online version of a local newspaper [www.shanghaistar.com](http://www.shanghaistar.com). In order to maximize the value of the news story, we combine keyword-ad and site-ad campaign through Google AdWords. Following the model, we begin both types of ads at the same time and our purpose is to obtain the optimal stop time for both advertising strategies.

Data are collected from Jan 9<sup>th</sup>, to Feb 22<sup>nd</sup>, 2006 from Google AdWords records of both ads. Before we apply the model, we use the data to validate our exponential decay assumption of the keyword-ad visit  $S_1(t)$  and site-ad visit  $S_2(t)$  functions. Results of MS Excel to estimate  $S_1(t)$  using the best-fit exponential curve (red curve in Figure 6) and estimate  $S_2(t)$  using average (red curve in Figure 7) fit our assumptions well. (See Figure 6 and Figure 7).

Then we used the detailed information about both keyword-ad and site-ad to calculate values of parameters in our model (see Table 4).

Finally, we set time period  $T=100$  days and the budget limit  $B=\$40$ . Advertisers who want to apply our model just need to set their own numerical values of these exogenous variables. The optimal solution is  $y_1^*=100$ ,  $y_2^*=76.32$  and the estimated number of clicks is 44. Thus, the optimal decision for advertisers is to hold the keyword-ad to the end of the advertising period, but end the site-ad around two and half months.

## 4.2. Site-Targeted Advertisement First, Keyword Search Advertisement Last

For second specific application, advertisers want to start their advertising campaign by using site-ad to aggressively impress the public and then switch to keyword-ad without overlapping. This strategy can be applied to advertisement of new products and services where the advertiser wants to expose her ads to a mass population of potential customers more quickly.

In a global economy with intense competition, firms face

**Table 4. Summary of google adwords' report**

Ad Type	Clicks	Impression	CTR (percent)	CPC	CPM
Keyword	9	1276	0.69	\$0.05	-
Site	8	2959	0.27	-	\$7.66

<sup>1</sup>Shanghai is a big city in China.

strong pressure to continuously exploit new product or services and effectively advertising those new products and services to potential consumers. At the beginning period of advertising, companies are not only interested in how many purchases have been made, but also how many consumers are aware of the new product or services. In other words, either a visit to the advertiser's website or an impression on the ad to potential customers brings benefits to the company. Consumers who visit firms' website may immediately make a purchase, while potential consumers exposed to the ad without clicking may come back and purchase the product or service later.

Site-ad meets the requirement to wildly and quickly impress the public in the early stage of advertising period. Although keyword-ad is not an effective marketing tool at the beginning because potential customers are unfamiliar with the new product or brand, it is a cost-efficient marketing method after certain level of awareness is achieved. Therefore, we suggest that site-ad first and keyword-ad last might be a better strategy for advertisers. Then the question for advertisers is how to determine optimal switching time from site-ad to keyword-ad.

### 4.2.1. Modified Mathematical Model

For this specific application, we assume that advertisers will begin with site-ad only, and then switch to keyword-ad without overlapping. To meet requirements of this specific application, we modify the general model in section 3 as follows: (1) we set the start time of site-ad as zero and the ending time of keyword-ad as the end of the whole advertising period  $T$ ; (2) we use  $z$  as the switching time of a site-ad and a keyword-ad.

The general model is modified as follows:

$$\begin{aligned} \text{Max } R(z) = & r_1 \left\{ \int_z^T S_1(t) l_1 dt + \int_0^z S_2(t) l_2 dt \right\} \\ & + r_2 \left\{ \int_z^T S_1(t) dt + \int_0^z S_2(t) dt \right\} \end{aligned} \quad (7)$$

subject to:

$$0 \leq z \leq T \quad (8)$$

$$p_1 \int_z^T S_1(t) l_1 dt + p_2 \int_0^z S_2(t) dt \leq B \quad (9)$$

Similar to the application of advertising perishable information with click-based revenue only, we estimate the visit of site-ads using average visits  $S_2(t) = s_2$  and steady click-through rate  $l_1$  and  $l_2$  over the advertising period  $[0, z^*]$ . We also assume that  $S_1(t) = s' + g(z)$ , where  $g(z)$  refers to visits due to awareness of site-ad. A linear function  $g(z) = az + b$  is used to estimate the value of  $S_1(t)$ . The trade-off in this model is that longer site-ads duration leads to a higher number of impressions at the beginning of keyword-ad, but runs out the budget more rapidly at the same time.

We get two optimal solutions for switching time  $z^*$ .



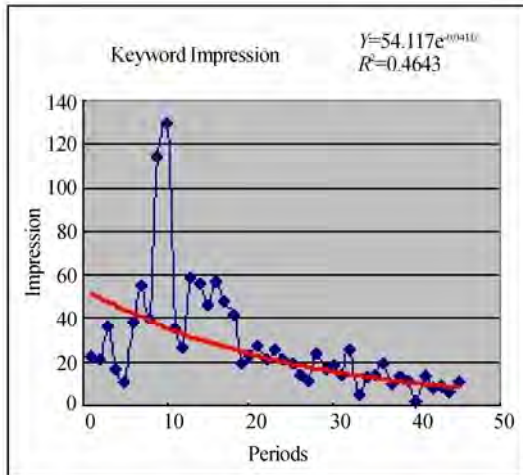


Figure 6. Summary of keyword-ad impression

$$a) \quad z^* = \frac{s_2 p_2 - p_1 l_1 (b + s' - aT) - \sqrt{\Delta}}{2ap_1 l_1}$$

$$b) \quad z^* = \frac{s_2 p_2 - p_1 l_1 (b + s' - aT) + \sqrt{\Delta}}{2ap_1 l_1}$$

where  $\Delta = 4ap_1 l_1 ((b + s')p_1 l_1 T - B) + (s_2 p_2 - p_1 l_1 (b + s' - aT))^2$ .

In both cases, advertisers switch from site-ads towards keyword-ads at  $z^*$ .

#### 4.2.2. Comparative Static Analysis

Similar to the case in section 4.1, we investigate the impact of changing parameter values under two conditions: whether the budget constraint is binding or not. We also assume here that advertisers want to make full use of the budget to maximize the advantages of both keyword-ad and site-ad.

(1) The budget constraint is not binding

$$\text{Let } \tilde{B} = B - p_1 l_1 ((s' + b)(T - z) + \frac{a}{2}(T^2 - z^2)) - p_2 s_2 z$$

denotes remaining budget. In this scenario, advertisers fail to use up their advertising budget on site-ad. Table 5 shows how an increase in value of parameters will influence both revenue and remaining budget.

We see from the table that only  $s_2$  has impact on both revenue and remaining budget. Intuitively, when the budget is not used up, an increase in visits of site-ad  $s_2$  to some value but lower than  $B/p_2 T$  will consume more advertising budget and increase revenue. Therefore,

Table 5. Comparative static analysis under non-binding budget

Variable of interest	Effects of an increase in					
	$a$	$s', b$	$s_2$	$p_1, p_2$	$l_1$	$l_2$
$R$	N/A	N/A	+	N/A	N/A	+
$\tilde{B}$	N/A	N/A	-	-	N/A	N/A

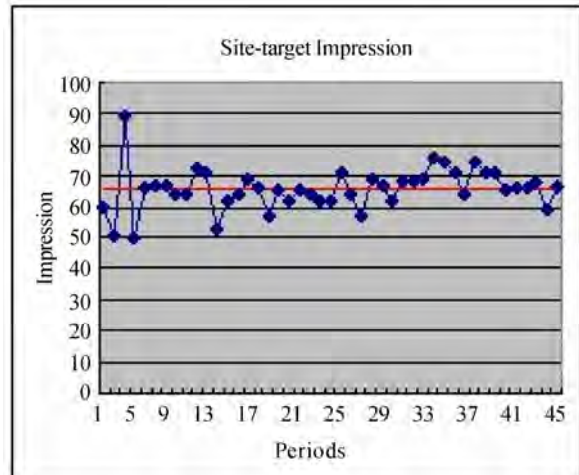


Figure 7. Summary of site-ad impression

advertisers with advertising dollar left can transform budget resource into revenue by selecting a more popular web site or start new site-ads.

(2) The budget constraint is binding

Under the condition that budget constraint is binding, advertiser is interested to see how change in value of one parameter will influence the optimal switching time, such as moving the optimal switching time earlier or later (see Table 6).

We can see from Table 6 that an increase in site-ads visits  $s_2$  will force the advertiser to move the optimal switch time earlier because the expected level of awareness is achieved earlier in time and advertising resources is used more quickly. On the other hand, an increase of value of parameters  $s'$  or  $b$  have the opposite effect on switch time although they have the same effect on revenue as  $S_2$ . However, an increase in value of parameters like payment for each click  $p_1$  and payment for each impression  $p_2$  has uncertain effect on switch time because the closed-form solution is not available.

Table 6. Comparative static analysis under binding budget

Variable of interest	Effects of an increase in parameters					
	$a$	$s', b$	$s_2$	$p_1, p_2$	$l_1$	$l_2$
$R$	+	+	+	N/A	+	+
$z^*$	?	+	-	?	?	N/A

\*Question mark means no close-form solution available

Table 7. Parameters used in numerical example

Parameter	Parameter	Parameter
$r_1 = \$0.25$	$r_2 = \$10.00$ per thousand	$a = 20$
$s' = 50$	$s_2 = 7500$	$B = \$250$
$p_1 = \$0.20$	$p_2 = \$8.00$ per thousand	$b = 40$
$l_1 = 1\%$	$l_2 = 0.03\%$	$T = 20$

We use a numerical example to illustrate the application of our model. Similar to the application of advertising perishable information with click-based revenue only, we can calculate values of parameters using the information from search engines. Using the values in Table 7, we obtain that the switch time is  $z^* = 4.1$  and the revenue is \$342.10 for the advertiser.

## 5. Conclusions

Internet advertising, especially search engine advertising has quickly become vital for businesses to succeed in e-commerce. Google, the largest search engine, provides site-targeted ads to advertisers in addition to its traditional pay-per-click model. In spite of the fact that more and more firms put site-targeted ads on Google-networked websites, very little research has attempted to analyze how advertisers can make use of this new type of ads such as a mixed strategy of combining it with the CPC-based keyword search advertisement. This research attempts to fill this gap by suggesting and formally modeling the strategy of combining both types of ads.

We developed a general model to address the research problem of what is the optimal time to start and end both keyword-ads and site-ads. This model would help advertisers to maximize their revenues. We modify the general model for two specific scenarios: (1) Advertising of perishable content with click-based revenue only; and (2) Site-ads first and keyword-ads last. We provide closed-form solutions for these two applications and provide managerial insights under the situation of binding-budget and non-binding budget. Computational experiment and numerical example is also provided to illustrate the implementation of the model.

This research focuses on the mixed strategy of both keyword-ad and site-ad in Google AdWords framework. As for future research, it will be interesting to study strategies of advertising across different search engines when they adopt different mechanisms. Another promising area will be how search engines can help advertisers when they observe that advertisers using a mixed strategy.

## 6. Acknowledgment

The author would like to thank Professor Manuel Nunez of University of Connecticut for his great help on improving the early versions of this paper.

## REFERENCES

- [1] Internet Resource, "Search engine user attitudes", April-May 2004, [www.iProspect.com](http://www.iProspect.com)
- [2] Internet Resource, "IAB internet advertising revenue report," [http://www.iab.net/media/file/IAB\\_PWC\\_2008\\_6m.pdf](http://www.iab.net/media/file/IAB_PWC_2008_6m.pdf).
- [3] Internet resource, "comScore releases september U.S. search engine rankings," [www.comscore.com](http://www.comscore.com), <http://www.comscore.com/press/release.asp?press=1805>.
- [4] B. Jansen, "Click fraud," *IEEE Computer*, Vol. 40, No. 7, pp. 85–86.
- [5] Internet resource, "Google adwords online introduction," <http://adwords.google.com/support/bin/answer.py?answer=19657&topic=342>.
- [6] B. Jansen and T. Mullen, "Sponsored search: An overview of the concept, history, and technology," *International Journal of Electronic Business*, Vol. 6, No 2, pp. 114–131, 2008.
- [7] J. Feng, H. K. Bhargava, and D. M. Pennock, "Implementing sponsored search in web search engines: Computational evaluation of alternative mechanisms," *INFORMS Journal on Computing*, Vol. 19, No. 1, pp. 137–148. Winter 2007.
- [8] B. Edelman, M. Ostrovsky, and M. Schwarz, "Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords," *American Economic Review*, Vol. 97, No. 1, pp 242–259, March 2007.
- [9] J. Feng, Z. M. Shen, and R. L. Zhan, "Ranked items auctions and online advertisement," *Production and Operation Management*, Vol. 16, No. 4, pp. 510–522. July–August 2007.
- [10] K. Asdemir, N. Kumar, and V. Jacob, "Internet advertising pricing models," Technique report, School of Management, The University of Texas at Dallas, 2002.
- [11] Y. Hu, "Performance-based pricing models in online advertising," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2005, AAT 0807658.
- [12] T. A. Weber and Z. Zheng, "A model of search intermediaries and paid referrals," *Information Systems Research*, Vol. 18, No. 4, pp. 414–437, December 2007.
- [13] S. Kumar, M. Dawande, and V. S. Mookerjee, "Optimal scheduling and placement of internet banner advertisements," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 19, No. 11, pp. 1571–1584, November 2007.
- [14] B. Edelman and M. Ostrovsky, "Strategic bidder behavior in sponsored search auctions," *Decision Support Systems*, Vol. 43, No. 1, pp. 192–198.
- [15] J. Feng and M. Zhang, "Price cycles in online advertising auctions," *Proceedings of the 26<sup>th</sup> International Conference on Information Systems (ICIS)*, Las Vegas, NV, December 2005.
- [16] Ö. Özlük and S. Cholette, "Allocating expenditures across keywords in search advertising," *Journal of Revenue and Pricing Management*, Vol. 6, pp. 347–356.
- [17] R. M. Dewan, M. L. Freimer, and J. Zhang, "Management and valuation of advertisement-supported web sites," *Journal of Management Information Systems*, Vol. 19, No. 3, pp. 87–98, Winter 2002/2003.
- [18] G. E. Fruchter and W. Dou, "Optimal budget allocation over time for keyword ads in web portals," *Journal of Optimization Theory and Applications*, Vol. 124, No. 1, pp. 157–174, January 2005.
- [19] Internet resource, "Online newspaper viewership reaches record in 2007," [www.naa.org](http://www.naa.org), <http://www.naa.org/PressCenter/SearchPressReleases/2008/Online-Newspaper-Viewership.aspx>.



## Appendix

We use the Karush-Kuhn-Tucker conditions to solve the general model of Equation (1) on page 3.

Step 1: We get the following Lagrangian equation:

$$\begin{aligned} L(y_1, y_2, y_3, y_4, \lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \\ r_1 \left\{ \int_{y_1}^{y_2} S_1(t) l_1(t) dt + \int_{y_3}^{y_4} S_2(t) l_2(t) dt \right\} + r_2 \left\{ \int_{y_1}^{y_2} S_1(t) dt + \int_{y_3}^{y_4} S_2(t) dt \right\} \\ + \lambda_0 \left\{ B - \int_{y_1}^{y_2} S_1(t) l_1(t) p_1(t) dt + \int_{y_3}^{y_4} S_2(t) p_2(t) dt \right\} \\ + \lambda_1 (T - y_1) + \lambda_2 (T - y_2) + \lambda_3 (T - y_3) + \lambda_4 (T - y_4) \end{aligned}$$

Step 2: Using Leibniz rule, we get the Kuhn-Tucker conditions:

$$\frac{\partial L}{\partial y_1} = S_1(y_1) l_1(y_1) \{ \lambda_0 p_1(y_1) - r_1 \} - S_1(y_1) r_2 - \lambda_1 \leq 0; \quad y_1 \geq 0; \quad y_1 \frac{\partial L}{\partial y_1} = 0;$$

$$\frac{\partial L}{\partial y_2} = S_1(y_2) l_1(y_2) \{ r_1 - \lambda_0 p_1(y_2) \} + S_1(y_2) r_2 - \lambda_2 \leq 0; \quad y_2 \geq 0; \quad y_2 \frac{\partial L}{\partial y_2} = 0;$$

$$\frac{\partial L}{\partial y_3} = S_2(y_3) \{ \lambda_0 p_2(y_3) - r_1 l_2(y_3) \} - S_2(y_3) r_2 - \lambda_3 \leq 0; \quad y_3 \geq 0; \quad y_3 \frac{\partial L}{\partial y_3} = 0;$$

$$\frac{\partial L}{\partial y_4} = S_2(y_4) \{ r_1 l_2(y_4) - \lambda_0 p_2(y_4) \} + S_2(y_4) r_2 - \lambda_4 \leq 0; \quad y_4 \geq 0; \quad y_4 \frac{\partial L}{\partial y_4} = 0;$$

$$\frac{\partial L}{\partial \lambda_0} = B - \int_{y_1}^{y_2} S_1(t) l_1(t) p_1(t) dt - \int_{y_3}^{y_4} S_2(t) p_2(t) dt \geq 0; \quad \lambda_0 \geq 0; \quad \lambda_0 \frac{\partial L}{\partial \lambda_0} = 0;$$

$$\frac{\partial L}{\partial \lambda_1} = T - y_1 \geq 0; \quad \lambda_1 \geq 0; \quad \lambda_1 \frac{\partial L}{\partial \lambda_1} = 0;$$

$$\frac{\partial L}{\partial \lambda_2} = T - y_2 \geq 0; \quad \lambda_2 \geq 0; \quad \lambda_2 \frac{\partial L}{\partial \lambda_2} = 0;$$

$$\frac{\partial L}{\partial \lambda_3} = T - y_3 \geq 0; \quad \lambda_3 \geq 0; \quad \lambda_3 \frac{\partial L}{\partial \lambda_3} = 0;$$

$$\frac{\partial L}{\partial \lambda_4} = T - y_4 \geq 0; \quad \lambda_4 \geq 0; \quad \lambda_4 \frac{\partial L}{\partial \lambda_4} = 0;$$

Since we only consider solutions with economic meanings, we make several restrictions on both  $y$ 's and  $\lambda$ 's:

(1)  $\lambda_1$  and  $\lambda_3$  cannot be positive, which means no advertisement.

(2)  $y_1$  is always less than or equal to  $y_2$ , and  $y_3$  is always equal to or less than  $y_4$ .

(3) One of  $y_1$  and  $y_3$  must equal to zero, which makes sure that at least one type of ads begin.

$$\begin{aligned} \text{Let } \frac{\partial h(t)}{\partial t} = S_1(t) l_1(t) p_1(t), \quad \frac{\partial g(t)}{\partial t} = S_2(t) p_2(t), \\ f_1(t) = l_1(t) (\lambda_0 p_1(t) - r_1) + r_2, \quad f_2(t) = l_1(t) (r_1 - \lambda_0 p_1(t)) + r_2, \\ f_3(t) = \lambda_0 p_2(t) - r_1 l_2(t) + r_2, \quad f_4(t) = r_1 l_2(t) - \lambda_0 p_2(t) + r_2. \end{aligned}$$

Step 3: Assume the reverse function of  $f_i(t)$ ,  $i=1,2,3,4$ , exist and  $g(t)$  and  $h(t)$  exist, we get closed-form solution as followings:

When the budget condition is not binding:

$$1) \quad y_1 = 0 \quad y_2 = T \quad y_3 = 0 \quad y_4 = T;$$

When the budget condition is binding, there are 14 possible cases:

$$2) \quad y_1 = y_2 = y_3 = 0 \quad y_4 = h^{-1}(B + h(0));$$

$$3) \quad y_1 = y_3 = y_4 = 0 \quad y_2 = g^{-1}(B + g(0));$$

$$4) \quad y_1 = y_3 = 0 \quad y_2 = T \quad y_4 = h^{-1}(B - g(T) + g(0) + h(0));$$

$$5) \quad y_1 = y_3 = 0 \quad y_2 = g^{-1}(B + g(0) + h(0) - h(T)) \quad y_4 = T;$$

$$6) \quad y_1 = 0 \quad y_2 = y_4 = T \quad y_3 = h^{-1}(B + g(0) - g(T) - h(T));$$

$$7) \quad y_1 = g^{-1}(B + h(0) - g(T) - h(T)) \quad y_2 = y_4 = T \quad y_3 = 0;$$

$$8) \quad y_1 = y_3 = 0 \quad y_i = f_i^{-1}(0) \quad i = 2, 4$$

$$9) \quad y_1 = 0 \quad y_i = f_i^{-1}(0), i = 2, 3, 4$$

$$10) \quad y_1 = 0 \quad y_i = f_i^{-1}(0), i = 2, 3 \quad y_4 = T$$

$$11) \quad y_1 = 0 \quad y_2 = T \quad y_i = f_i^{-1}(0), i = 3, 4$$

$$12) \quad y_i = f_i^{-1}(0), i = 1, 4 \quad y_2 = T \quad y_3 = 0$$

$$13) \quad y_i = f_i^{-1}(0), i = 1, 2, 4 \quad y_3 = 0$$

$$14) \quad y_i = f_i^{-1}(0), i = 1, 2 \quad y_3 = 0 \quad y_4 = T$$

$$15) \quad y_i = f_i^{-1}(0), i = 1, 2, 3 \quad y_4 = T$$

The revenue function:

$$\begin{aligned} R = r_1 \left\{ \int_{y_1}^{y_2} S_1(t) l_1(t) dt + \int_{y_3}^{y_4} S_2(t) l_2(t) dt \right\} \\ + r_2 \left\{ \int_{y_1}^{y_2} S_1(t) dt + \int_{y_3}^{y_4} S_2(t) dt \right\} \end{aligned}$$

The remaining budget

$$\begin{aligned} \tilde{B} = B - \int_{y_1}^{y_2} S_1(t) l_1(t) p_1(t) dt - \int_{y_3}^{y_4} S_2(t) p_2(t) dt \\ = B + g(y_1) + h(y_3) - g(y_2) - h(y_4) \end{aligned}$$

(a) Starting time

$$\frac{\partial R}{\partial y_1} = -S_1(y_1) l_1(y_1) r_1 - S_1(y_1) r_2 < 0;$$

$$\frac{\partial \tilde{B}}{\partial y_1} = \frac{\partial g(y_1)}{\partial y_1} = S_1(y_1) l_1(y_1) p_1(y_1) > 0;$$

$$\frac{\partial R}{\partial y_3} = -S_2(y_3) l_1(y_3) r_1 - S_2(y_3) r_2 < 0;$$

$$\frac{\partial \tilde{B}}{\partial y_3} = \frac{\partial h(y_3)}{\partial y_3} = S_2(y_3) p_2(y_3) > 0;$$

which shows an increase in starting time decreases revenue but increases the remaining budget.

(b) Ending time

$$\frac{\partial R}{\partial y_2} = S_1(y_2) l_1(y_2) r_1 + S_1(y_2) r_2 > 0;$$

$$\frac{\partial \tilde{B}}{\partial y_2} = -\frac{\partial g(y_2)}{\partial y_2} = -S_1(y_2) l_1(y_2) p_1(y_2) < 0;$$

$$\frac{\partial R}{\partial y_4} = S_2(y_4) l_2(y_4) r_1 + S_2(y_4) r_2 > 0;$$

$$\frac{\partial \tilde{B}}{\partial y_4} = -\frac{\partial h(y_4)}{\partial y_4} = -S_2(y_4) p_2(y_4) < 0;$$

which shows that increase in ending time increase revenue but decrease the remaining budget.

Specific problems like the two problems discussed in Section 4 can be solved using the approach above.

# Market Segmentation for Mobile TV Content on Public Transportation by Integrating Innovation Adoption Model and Lifestyle Theory

Chi-Chung Tao

Tamkang University, Taipei County, Taiwan, R.O.C.  
Email: cctao@mail.tku.edu.tw

Received June 20<sup>th</sup>, 2008; revised August 10<sup>th</sup>, 2008; accepted September 16<sup>th</sup>, 2008.

## ABSTRACT

*An integrated approach based on innovation adoption model and lifestyle theory for customer segmentation of mobile TV content on public transportation using multivariate statistical analysis is proposed. Due to high daily trips and different train types Taiwan Railway Administration is chosen as the case study. Firstly, the content of mobile TV on the train are identified as the segmentation variable and key factor facets for mobile TV content are renamed by using factor analysis. Then, the cluster analysis is used to classify customer groups which are named by analysis of variance (ANOVA) and market segmentations are described with demographic, lifestyle and train patronage variables by using cross analysis and Chi-squared independence tests. Finally, this paper discusses empirical results to provide valuable implications for better mobile TV content marketing strategies in the future.*

**Keywords:** mobile TV, market segmentation, multivariate statistical analysis

## 1. Introduction

Mobile TV has been widely discussed among different players in the telecommunications and media industry. Mobile operators, which face the saturation of voice services and a declining ARPU (Average Revenue Per User), hope that the TV concept in the mobile phone will be the next killer application. According to analysts' predictions, mobile TV will become a service with significant market size. The valuations of the global mobile TV market vary across different analysts, from US\$ 5.5 billion in 2009 to US\$ 28 billion in 2010. A list of different analysts' predictions is shown in Table 1[1,2].

There are many trial projects confirming future success of mobile TV worldwide using different technologies and business models [3,4,5]. Currently three standards are competing with each other: Digital Video Broadcasting Transmission System for Handheld Terminals (DVB-H), MediaFLO and Digital Multimedia Broadcasting (DMB)

which are listed in Table 2[6]. AT&T has announced that its MediaFLO based mobile TV service will be going live in May 2008. While Qualcomm is in a position to leverage other technologies or use it for open access technologies such as WiMAX or use them for mobile services. It looks like that despite the EU having embraced a single standard for mobile TV, the US market will remain fragmented with multiple technologies.

According to mobile TV usage patterns identified in many worldwide trials TV content must adjust to mobile context of use [3,4,5]. Results of a Finnish study show that mobile TV users spent approximately 20 minutes a day watching mobile TV and more active users watched between 30 to 40 minutes per session [7]. Typical usage environments include transportation terminals (airport, train station, bus stop, etc.), in the moving vehicles, working places or at home. It is also found that smaller screens and the duration of usage may have significant influences on the types of mobile TV content as well as the way users' willingness to pay for mobile TV.

In Taiwan, a handheld TV experimental project was launched in October 2006. There were five teams participating in mobile TV trials in North and South Taiwan. The MediaFLO was tested in North Taiwan, while the DVB-H was chosen for South Taiwan. The experiment is expected

**Table 1. Global mobile TV market values**

Analyst firm	Year	Prediction of market size
Datamonitor	2009	US\$ 5.5 billion
Strategy Analytics	2009	US\$ 6.4 billion
ABI Research	2010	US\$ 27 billion
Pyramid Research	2010	US\$ 13 billion – 28 billion
Frost & Sullivan	2011	US\$ 8.1 billion

**Table 2. Overview of mobile TV solutions**

Technology	Major market	Standard Group	Industrial players	Comparisons
DVB-H	Europe, Asia, North America, Australia	DVB OMA	Nokia, BenQ - Siemens, Motorola, Samsung, LG, Alcatel, ... etc.	IP based; 90% power saving design, either use independent MUX, OR
T-DMB	South Korea	DAB	LG, Samsung	some proprietary; bit rate 1.5Mbps, non-IP based; external antenna; less power saving; be constructed based on existing DAB network
S-DMB	South Korea, Japan	DAB	LG, Samsung, Alcatel	Some proprietary; non-IP based, big antenna; power saving; satellite transmission with terrestrial repeaters
MediaFLO	North America	Qualcomm	Qualcomm & CDMA manufacturers	Proprietary; non-IP based; bandwidth efficiency
ISDB-T	Japan only			Proprietary
MBMS	3G service regions	3GPP, 3GPP2	Most 3G manufacturer	Bit rate 345kbps~2Mbps, IP based; based on 3G network; standardization still in progress
IPTV over WiMAX	WiMAX service regions	WiMAX Forum	Intel...etc.	standardization still in progress

to end no later than June 2008 and the formal licenses for mobile TV operators will be permitted after NCC's (National Communications Commission) official evaluations. Preliminary results of this experiment show that end-users will use mobile TV to fill in gaps in their daily schedules: waiting for the bus or subway, sitting in the train etc. In these situations mobile TV competes with other possibilities such as reading a book, listening to radio, playing a mobile game or just watching out the window. In these scenarios mobile TV might be an appealing choice, but only if it does not inflict significant costs. On the other hand, if the pricing is low enough, there might be quite large audiences awaiting the launch of mobile TV services. A logical choice might be to keep subscription prices as low as possible, thus maximize the popularity of mobile TV, and subsidize the lower subscription income with higher advertising revenue. In summary, these five teams together with three main telecommunication operators (CHT, TMT, FET) reach a consensus that a user-centered content design will contribute to future success of Taiwan's mobile TV market.

To survive in competitive mobile TV services markets, the operators need to determine who the target customers are, what motivates them and why they pay for the mobile TV content. This process is called market segmentation, by which companies are able to understand their loyal customers and concentrate their limited resources into them. Although there are many studies covering critical variables for market segmentation of mobile TV content, until now only few papers focus on public transportation systems with high speed, especially on the train [8]. Currently Taiwan Railway Administration's (TRA) is conducting a BOT project to implement mobile commerce services on the train. In addition, TRA's network is across north and south Taiwan, content diversity of mobile TV can be tested on the train countrywide. These advantages

may attract the five teams to focus on certain killer applications for mobile TV.

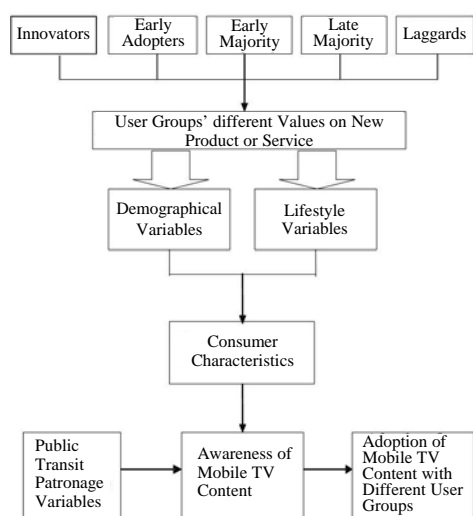
The concept of segmentation in mobile TV marketing recognizes that consumers differ not only in the price they will pay, but also in a wide range of benefits they expect from the content. Good mobile TV content with compelling value-added services are provided by tight business and strategic partnership arrangements and by involving a large number of companies, with each influencing other parties in the value chain. Wang [9] identifies that powerful actors, such as carriers and the media industry's content providers must agree on business models that support the new ecosystem of mobile TV. Carlsson and Walden [7] conclude that mobile TV content is the key factor to determine the adoption and usage of mobile TV, especially when traveling with public transportation to and from work in order to relax or to keep up to date with the latest news.

This paper is aimed on proposing an integrated approach for market segmentation of mobile TV content on the train by integrating innovation adoption model and lifestyle theory. Figure 1 shows the conceptual framework of this integrated approach.

First, the mobile TV content is identified as the segmentation variable. And key factor facets for mobile TV content are redefined by using factor analysis. Then, the cluster analysis is used to classify consumer groups which are named by analysis of variance (ANOVA) and market segmentations are described with demographic, lifestyle and train patronage variables by using cross analysis and Chi-squared independence tests. Finally, empirical results are analyzed and the conclusion follows.

## 2. Literature Review

It is the first step for mobile TV companies to identify the



**Figure 1. Conceptual framework of the integrated approach**

primary variables of market segmentation to understand their customers' requirements, attitudes and habits. These primary variables called segmentation variables which can be derived from the technological, demographical and psychological, behavioral perspective, such as age, income, gender, occupation, attributes of product or service, personal interest, consumer awareness, perception. A review of prior studies suggests the theoretical foundations of the hypotheses formulations [10]. To achieve this goal, this paper examines two prevalent theories (lifestyle theory and innovation diffusion theory) for identifying individual acceptance of mobile TV content on the train.

Lifestyle is a key factor in determining the adoption rate for mobile TV content. Individuals will adopt given behavior patterns representative of their lifestyles, and as a consequence will purchase different types of mobile TV content. Many studies verified that behavioral variations in purchases even if there is no question of a mix of socio-demographical variables coming into play, lead to a need for research into lifestyle as a potentially influential factor. Market segmentation according to features of lifestyle divides the market into segments based on activities, interests, and opinions [11]. The lifestyle segmentation in this paper is defined by including variables like activities, referring to the way in which individuals spend their time and money; interests, which are those things in their immediate surroundings they consider more or less important; and opinions, the view they have of themselves and of the world around them. One of well-known lifestyle analysis methods is proposed by Wind and Green [12]. First, factor-groups influencing lifestyle patterns are reduced by using factor analysis. Then, the cluster analysis is used to classify consumer groups. The interrelationships among lifestyle patterns and other consumer behavioral variables are verified by cross analysis.

The innovation diffusion theory (IDT) is a well-known theory proposed by Rogers [13]. The adoption of mobile TV content on the train could be studied from the perspective of information technology innovations. IDT states that diffusion of an innovation depends on five general attributes including relative advantage, compatibility, complexity, observability, and trial ability. These characteristics are used to explain the user adoption and decision making process. They are also used to predict the implementation of new technological innovations and clarify how these variables interact with one another. The central concept of innovation diffusion is "the process in which an innovation is communicated through certain channels, over time, among the members of a social system." However, research has suggested that only the relative advantage, compatibility, and complexity are consistently related to innovation adoption [14]. Relative advantage is similar to perceived usefulness, whereas complexity is similar to perceived ease of use. Compatibility is the degree to which the innovation is perceived to be consistent with the potential users' existing values, previous experiences, and needs [15]. High compatibility will lead to preferable adoption. Holak and Lehmann modified Rogers's model for measuring relative advantages [16]. They provided empirical evidence that relative advantage and compatibility "directly" affect consumers' purchase intention. The two factors have received favoritism in recent published articles in the area of product innovation. The two innovation characteristics are also peculiar in that they are defined in relation to existing products, while others (complexity, trial ability, and observability) indicate the innate characteristics of an innovation.

Originating in IDT research, different adopter groups perceive innovations and thus behave differently. Miller [17] finds that prior knowledge of potential adopters can focus the use of resources to prevent an innovation from failing. Innovation theory can be applied to identify the attitudes and behavior of early adopters, as a dynamic basis for a market segmentation model. Therefore, Holak's new product adoption model is used to examine the acceptance of mobile TV for public transportation users. Through the review of relevant literature, following variables of three dimensions for mobile TV content segmentation analysis shown in Table 3 are examined in this paper [18,19,20]:

**Table 3. Segmentation dimensions**

Demographic	Psychographic	Behavioral
Age, occupation, gender, income, religion, nationality, education, marital status, ethnicity	Values, lifestyle, interests, opinions, activities	Product use patterns, perceived benefits (value-added, ease of use, usability, usefulness)

To identify links among what the mobile TV operators would know about their customers and the bundles of content they could offer, clustering algorithms are generally used as the primary methodology for market segmentation. Clustering analysis techniques have been discussed in details in the literature [21,22,23]. The most popular is k-means algorithm which together with its modifications was broadly reviewed by different authors [24,25,26]. It is also found that algorithms using computational intelligence did not show better results than k-means, the combinations of several algorithms are very often recommended as the conclusion. Zakrwska and Murlewski [27] investigated the shortcomings and advantages of three algorithms of clustering analysis: k-means, two-step clustering and density based spatial clustering of applications with noise. Their numerical tests showed that k-means is very efficient for large multidimensional data sets, however depends strongly on the choice of input parameter k. However, it is not recommended in the case of data sets with noise.

### 3. Research Methodology

#### 3.1. Research Design

The framework of research design is shown in Figure 2. After conducting an interview survey (face-to-face) with questionnaires, the mobile TV content on the train is identified as the segmentation variable and key factor facets for mobile TV content are redefined by using factor analysis. Then, the cluster analysis (k-means) is used to classify consumer groups which are named by ANOVA. The market segmentations are described with content usage, demographic, lifestyle and train patronage variables by using cross analysis and Chi-squared independence tests. Finally, each segment market can be targeted with precise customer characteristics and those results are used as the starting point for providing the market strategies.

The first round survey with 90 questionnaires was conducted on the three train types from Taipei station to Taichung station from 22<sup>nd</sup> to 28<sup>th</sup> December 2007. After reviewing preliminary results, some question items were modified and new question items of lifestyle variables were supplemented. The second round survey with 500 questionnaires was conducted on the train from 18<sup>th</sup> to 25<sup>th</sup> January 2008. Two types of handheld devices NOKIA N77 and N92 were used to demonstrate mobile TV content. The valid sample consisted of 462 respondents.

#### 3.2. Research Model and Hypotheses

The research model tested in this paper is shown in Figure 3. With this integrated approach the mobile TV content may be regarded as segmentation variables including willingness to use, time of usage, price of willingness to pay,

type of payment, incentives to take train. The demographic variables (gender, occupation, age, income, education), lifestyle variables (knowledge-oriented, recreation-oriented, high living quality-oriented, favorite information-oriented, price sensitive-oriented, fashion-oriented), train patronage variables (frequency, travel time, train type, trip purpose) are chosen as descriptive variables to depict customer characteristics.

The following hypotheses of the proposed constructs are based on prior studies in the relevant literature [10,14,15]:

H<sub>1a</sub>: No significant difference exists between each segment and gender.

H<sub>1b</sub>: No significant difference exists between each segment and occupation.

H<sub>1c</sub>: No significant difference exists between each segment and age.

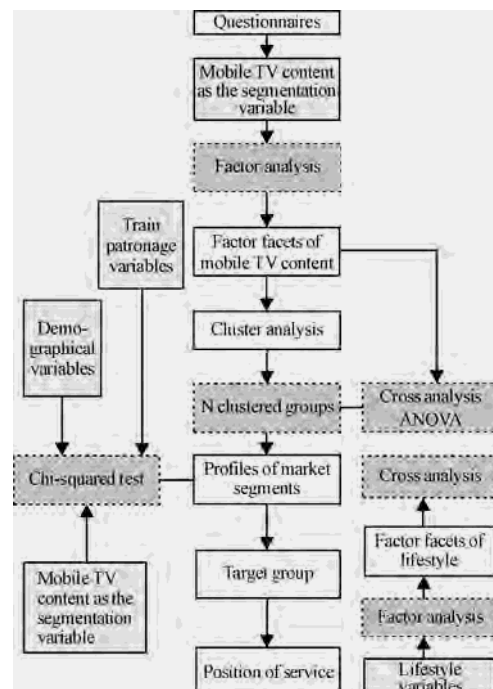


Figure 2. Framework of research design

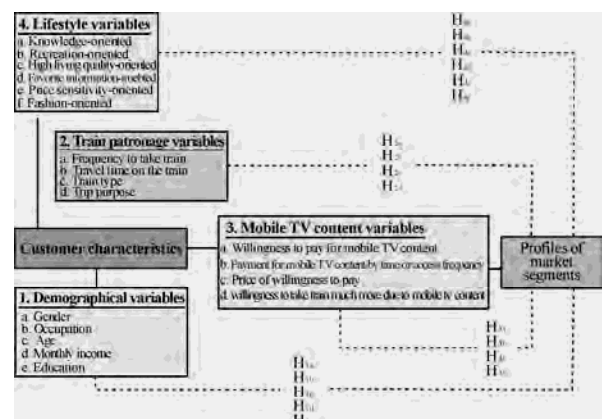


Figure 3. Research model and hypotheses

H<sub>1d</sub>: No significant difference exists between each segment and income.

H<sub>1e</sub>: No significant difference exists between each segment and education.

H<sub>2a</sub>: No significant difference exists between each segment and frequency to take train.

H<sub>2b</sub>: No significant difference exists between each segment and travel time on the train.

H<sub>2c</sub>: No significant difference exists between each segment and train type.

H<sub>2d</sub>: No significant difference exists between each segment and trip purpose.

H<sub>3a</sub>: No significant difference exists between each segment and willingness to pay for mobile TV content.

H<sub>3b</sub>: No significant difference exists between each segment and payment for mobile TV content by time or access frequency.

H<sub>3c</sub>: No significant difference exists between each segment and price of willingness to pay.

H<sub>3d</sub>: No significant difference exists between each segment and willingness to take train much more due to mobile TV content.

H<sub>4a</sub>: No significant difference exists between each segment and knowledge-oriented lifestyle.

H<sub>4b</sub>: No significant difference exists between each segment and recreation-oriented lifestyle.

H<sub>4c</sub>: No significant difference exists between each segment and high living quality-oriented lifestyle.

H<sub>4d</sub>: No significant difference exists between each segment and favorite information-oriented style.

H<sub>4e</sub>: No significant difference exists between each segment and price sensitive-oriented lifestyle.

H<sub>4f</sub>: No significant difference exists between each segment and fashion-oriented lifestyle.

#### 4. Results Analysis

The descriptive statistics of demographical variables and train patronage variables are summarized as follows:

1) A total of 48% of the respondents were male. The age of most respondents was from 19 to 33 years old with a total of 64%. 70% of them were college educated, 29% were students and 55% working in service industry and trade. 31% of respondents earn under US \$ 300 per month and 21% earning US \$ 800 – 1000 per month.

2) 23% of total respondents take train at least one time daily and 31% taking at least one time per week. 36% of them were commuters for work and school, and 25% were leisure or tourism travelers. The percentage of travel time from 31-60 minutes was 33% due to short-trip distance. A total of 34% and 36% of the respondents most often take the EMU commuter train type and “Zhi-Chiang” express train type, respectively.

Factor analysis utilized the principal axis factoring method in order to identify underlying constructs. The KMO measure of sampling adequacy and Barlett’s tests of sphericity provided support for the validity of the factor analysis of the data set. The varimax rotation facilitated interpretability.

The major mobile TV content consists of seven factor facets which together explained 72.1% of the total variation shown in Table 4.

Factor facets are thus named: films, news, documentaries, sports, leisure & tourism, entertainment and drama. Measure validation was also examined for internal consistency by computing Cronbach’s  $\alpha$  coefficient. The average Cronbach’s  $\alpha$  value was found to be greater than 0.8, in accordance with Nunnally’s standard [28]. The same procedure of factor analysis for lifestyle variables was also conducted to obtain following factor facets with a satisfied average Cronbach’s  $\alpha$  value 0.7: knowledge-oriented type, recreation-oriented type, high living quality-oriented type, favorite information-oriented type, price sensitive-oriented type and fashion-oriented type.

**Table 4. Results of mobile TV content by factor analysis**

Component	Initial Eigen-values			Extraction Sums of Squared Loading			Rotation Sums of Squared Loading		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1(Films)	12.71	27.80	27.80	12.71	27.80	9.32	23.21	23.21	23.21
2(News)	7.24	13.41	41.21	7.24	13.41	7.57	12.16	12.16	35.37
3(Documentaries)	4.18	10.06	51.27	4.18	10.06	3.98	9.89	9.89	45.26
4(Sports)	2.26	7.02	58.29	2.26	7.02	3.70	9.34	9.34	54.60
5(Leisure & Tourism)	1.60	6.22	64.51	1.60	6.22	2.53	8.22	8.22	62.82
6(Entertainment)	1.25	4.55	69.05	1.25	4.55	1.95	6.19	6.19	69.02
7(Drama)	1.04	3.01	72.07	1.04	3.01	1.22	3.06	3.06	72.07
8(Weather forecasts)	0.99	2.79	74.85						
9(Stock reporting)	0.97	2.59	77.44						
10(Religion talking)	0.936	2.516	79.96						

Then based on cluster analysis of six factor scores for lifestyle variables, the whole sample was classified into three groups: Group 1 (Group of being fashion and optimistic), Group 2 (Group of being leisured and living information-oriented) and Group 3 (Group of being conservative and traditional). ANOVA was conducted and the results reveal that there are significant differences among three groups on each factor scores.

The hypothesized relationships in Figure 3 are tested using ANOVA, cross analysis and Chi-squared test to examine if there is enough evidence to infer that two research variables are related. Results of hypotheses tests with P-value, Chi-squared values or F-value are summarized in Table 5.

As shown in Table 5, approximately 60% of all hypotheses are proven to have significant differences at the 95% level.

To identify characteristics and the structure of the market segments, a summarized analysis was conducted to compare results of mobile TV content, demographical, train patronage and lifestyle variables including gender, occupation, age, income, education, frequency to take train, travel time on the train, train type, trip purpose, willingness to pay for mobile TV content, payment by time or frequency, price of willingness to pay, type of lifestyle, favorite content and willingness to take train due to mobile TV content. The summarized profiles are shown in Table 6.

As to the key information in the target group, the gender is male, monthly income is under US \$ 300, the age

is from 19 to 33 years old. They are high school and college students and spend 31-90 minutes on the train per trip. They will pay either US \$ 0.25/ h or US \$ 0.5 for one usage during the travel time between 61 and 121 minutes and over 121 minutes. Their lifestyle types include entertainment-oriented, fashion-oriented, price sensitivity-oriented. Their favorite content includes films, news, entertainment, documentaries and leisure & tourism. They are willing to take train much more if TRA offers more interesting mobile TV content.

**Table 5. Results of hypotheses tests**

Hypothesis	P-value	Chi-squared ( $x^2$ ) / F-value	Accept	Reject
H <sub>1a</sub>	0.005	10.502 ( $x^2$ )	√	
H <sub>1b</sub>	0.010	29.287 ( $x^2$ )		×
H <sub>1c</sub>	0.021	28.165 ( $x^2$ )		×
H <sub>1d</sub>	0.012	22.760 ( $x^2$ )		×
H <sub>1e</sub>	0.467	7.666 ( $x^2$ )	√	
H <sub>2a</sub>	0.596	4.602 ( $x^2$ )	√	
H <sub>2b</sub>	0.002	24.890 ( $x^2$ )		×
H <sub>2c</sub>	0.668	4.068 ( $x^2$ )	√	
H <sub>2d</sub>	0.043	11.946 ( $x^2$ )	√	
H <sub>3a</sub>	0.022	11.463 ( $x^2$ )	√	
H <sub>3b</sub>	0.096	4.692 ( $x^2$ )	√	
H <sub>3c</sub>	0.032	19.678 ( $x^2$ )		×
H <sub>3d</sub>	0.000	22.980 ( $x^2$ )		×
H <sub>4a</sub>	0.597	0.516 (F)	√	
H <sub>4b</sub>	0.000	21.63 (F)		×
H <sub>4c</sub>	0.020	11.37 (F)	√	
H <sub>4d</sub>	0.179	1.72 (F)	√	
H <sub>4e</sub>	0.048	2.05 (F)	√	
H <sub>4f</sub>	0.000	5.76 (F)		×

**Table 6. Profiles of market segments**

	Group 1 (fashion and optimistic)	Group2 (leisured and living information-oriented)	Group3 (conservative and traditional)
Sample size	193	94	175
Gender	Male > Female	Female > Male	Male = Female
Occupation	College students	Service industry and trade	Uniformly distributed
Age	19-33 years old	14-42 years old	Uniformly distributed
Average income per month	Under US \$ 300	US \$ 800-100	US \$ 800-1500
Education	College and high school	College	High school
Frequency to take train	No distinct segments	No distinct segments	No distinct segments
Travel time on the train	More than 90 minutes	31-90 minutes	31-60 minutes
Train type	Zhi-Chiang and EMU	Zhi-Chiang and EMU	Zhi-Chiang and EMU
Trip purpose	Commuter, Homebound traveler, leisure traveler	Commuter, leisure traveler	Uniformly distributed
Willingness to use mobile TV	Yes for travel time between 61 and 121 minutes and over 121 minutes	Yes for travel time between 61 and 121 minutes and over 121 minutes	Yes for travel time over 121 minutes
Payment by time or frequency	both	Prefer by frequency	Prefer by frequency
Price of willingness to pay	US \$ 0.25/h and US \$ 0.5 for one usage	US \$ 0.5 for one usage	US \$ 0.3 for one usage
Type of lifestyle	Entertainment-oriented, high living quality-oriented and fashion-oriented	High living quality-oriented	Price sensitive-oriented
Favorite content	News, films, entertainment, documentaries	News, films, leisure & tourism, entertainment	No preference
Duration of watching mobile TV content	More than 30 minutes	20-30 minutes	No preference
Willingness to take train due to mobile TV content	Yes for the travel time over 90 minutes	Yes for the travel time over 90 minutes	Yes for cheaper fares only

## 5. Conclusions

This paper identifies the new primary factors for mobile TV content on the train which may not found in the previous studies. Additionally, this paper proposes a concise framework of research methodology for market segmentation and user preferences for mobile TV content on the train.

However, the proposed approach is applicable to the case of "on the train". The case of "railway networks" which fully represents mobile TV content anywhere and anytime in stations and trains needs to be researched further. It is also recognized that many other facets of individual differences (e.g., psychological type, cognitive processing skills etc.) may be candidate variables for lifestyle consideration.

## REFERENCES

- [1] N. Holland, *et al.*, "Rescuing 3G with Mobile TV: Business Models and Monetizing 3G," Pyramid Research, March 2006.
- [2] Wireless World Forum, "Mobile Youth 06 Video," mobileYouth06-part one, July 2006.
- [3] J. Trefzger, "Mobile TV-launch in germany-challenges and implications," working paper No. 209, Institute for Broadcasting Economics, Cologne University, Germany, 2005.
- [4] S. Orgad, "How will mobile tv transform viewer's experience and change advertising," Final report, Dept. of Media and Communications, London School of Economics and Political Science, November 2006.
- [5] QuickPlayMedia, "Mobile TV and video survey 2008," Toronto, Canada, 2008.
- [6] M. P. Shih, "Analysis of mobile TV and its key success factors: from the perspective of mobile operator," Proceedings of International Symposium on HDTV and Mobile TV, Taipei, Taiwan, 2007.
- [7] C. Carlsson and P. Walden, "Mobile TV-to live or die by content," Proceedings of the 40th Hawaii International Conference on System Sciences, IEEE, Hawaii, USA, 2007.
- [8] C. M. Tan and C. C. Wong, "Mobile broadband race: Friend or foe," Proceedings of the International Conference on Mobile Business, IEEE, ICMB'06, 2006.
- [9] G. Wang, "Mobile TV value chain and operator strategies," Master's Thesis, Dept. of Communication Systems, School of Information and Communication Technology, KTH, Finland, February 2007.
- [10] T. M. Lee and J. K. Jun, "Contextual perceived usefulness? toward an understanding of mobile commerce acceptance," Proceedings of the International Conference on Mobile Business, IEEE, ICMB'05, 2005.
- [11] J. T. Plummer, "The concept and application of lifestyle segmentation," Journal of Marketing, pp. 33-74, January 1974.
- [12] Y. Wind and P. E. Green, "Some conceptual measurement and analytical problem in life style research," Life style and Psychographics, Chicago, AMA, 1974.
- [13] E. Rogers, Diffusion of Innovation, Free Press, New York, 1962.
- [14] J. H. Wu and S. C. Wang, "What drives mobile commerce? An empirical evaluation of the revised technology acceptance model," Information & Management 42, pp.719-729, 2005.
- [15] Y. Shin, H. Jeon, and M. Choi, "Analysis of the consumer preferences toward m-commerce applications based on an empirical study," Proceedings of International Conference on Hybrid Information Technology, IEEE, ICHIT'06, 2006.
- [16] S. L. Holak and D. R. Lehmann, "Purchase intentions and the dimensions of innovation: An exploratory model," Journal of Product Innovation Management, 7 (1), pp. 59-73, 1990.
- [17] R. N. Miller, "Target marketing," Multinational Marketing, Vol. 13, No. 10, 1993.
- [18] A. L. Gilbert and J. D. Kendall, "A marketing model for mobile wireless services," Proceedings of the 36th Hawaii International Conference on System Science (HICSS'03), 2003.
- [19] H. H. Lin and Y. S. Wang, "Predicting consumer intention to use mobile commerce in taiwan," Proceedings of the International Conference on Mobile Business (ICMB'05), 2005.
- [20] H. Feng, T. Hoegler, and W. Stucky, "Exploring the critical success factors for mobile commerce," Proceedings of the International Conference on Mobile Business (ICMB'06), 2006.
- [21] C. C. Aggarwal, C. Procopiuc, J. S. Wolf, P. S. Yu, and J. S. Park, "Fast algorithms for projected clustering," Proceedings of SIGMOD Conference, Philadelphia, 1999.
- [22] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: A review," ACM Computing Surveys, Vol. 31, No. 3, September 1999.
- [23] M. Zait and H. Messatfa, "A comparative study of clustering methods," FGCS Journal, Special Issue on Data Mining, 1997.
- [24] P. V. Balakrishnan, M. C. Cooper, V. S. Jacob, and P. A. Lewis, "Comparative performance of the fscl neural net and k-means algorithm for market segmentation," European Journal of Operational Research, No. 93, 1996.
- [25] H. Hruschka and M. Natter, "Comparing performance of feedforward neural nets and k-means for cluster-based market segmentation," European Journal of Operational Research, No. 114, 1999.
- [26] C. Y. Tsai and C. C. Chiu, "A purchased-based market segmentation methodology", Expert Systems with Applications, No. 27, 2004.
- [27] D. Zakrzewska and J. Murlewski, "Clustering algorithms for bank customer segmentation", Proceedings of the 5th International Conference on Intelligent Systems Design and Applications (ISDA'05), 2005.
- [28] J. C. Nunnally, Psychometric Theory, McGraw-Hill, New York, 1967.



# Holonic Production System to Obtain Flexibility for Customer Satisfaction

Gandolfo Dominici

University of Palermo, Faculty of Economy, Italy  
Email: gandolfodominici@unipa.it

Received December 3<sup>rd</sup>, 2008; revised December 10<sup>th</sup>, 2008; accepted December 15<sup>th</sup>, 2008.

## ABSTRACT

*The Holonic Production System (HPS) can be a valid choice to overcome the problems of traditional production systems' architectures, thanks to its capability to adapt and react to changes in the business environment whilst being able to maintain systemic synergies and coordination. The HPS is made of holons seen as functional production units which are simultaneously autonomous and cooperative. Although the holonic approach could represent a valid solution in order to pursue the necessary levels of agility of production systems, they have been scarcely implemented in practice and even less studied from a business studies perspective. The purpose of this discussion paper is to show the benefits of further research on cases of implementation of HPS from a business organization studies perspective. Very little research on this topic has been done outside the field of business engineering and computer science; the study of this topic from a different perspective can shed the light on new aspects and new applications of the theory.*

**Keywords:** holonic production systems, production agility, customer satisfaction

## 1. Introduction

Mass production showed its effectiveness in stable environments and with continuous growth trends<sup>1</sup> until the end of the 80's. Since the beginning of the 90's, it has begun to show its weaknesses due to the growing instability of business environments and of systemic complexity. The spread of Internet made it possible for firms the use of a low cost, worldwide extended, informative infrastructure which can bring profound changes in the market. In mature markets it is necessary to supply a wide variety of products in order to adhere to the need of customers whose role has changed from "consumer" to "prosumer"<sup>2</sup>. These changes caused the shift from "mass production" to "mass-customisation". In order to fulfil these new needs for agility, it becomes unavoidable for firms to develop an extremely flexible production structure able to:

- duly react to the market environment's turbulences;
- survive production system changes through the adoption of new technologies;
- adapt to the uncertainties of production systems in such environments.

Neither hierarchical or heterarchical systems are able

to fulfil these requirements [3,4]. Hierarchical systems have a typically rigid structure which makes it very hard for them to react to turbulences in an agile way. Heterarchical systems are networks of elements with common aims in which each element shares with the others the same "horizontal" position of power and authority. Though heterarchical systems can easily adapt to environmental changes and turbulences, their control system cannot assure the high level of performance and the predictable organization behaviours needed for the industrial production of goods.

## 2. Theoretical Framework

The growing power of IT opened new possibilities in the worldwide arena and supplied management new and effective instruments for planning, budgeting, design and customer care. The central role of the customer thrived to the point that the supply chain has begun to be defined as the "demand" chain [5]. Literature on this topic shows several trends which manufacturing and supply chain systems have to adapt to [6,7]:

- a) the paradigm shift from mass production to *semi-personalized* production;
- b) the opening to collaboration with other agents in order to speed up production innovation and processes;
- c) the critical role of effective and efficient cooperation inside the network;
- d) understanding the problems connected to the implementation of a centralized control system between different entities with different information, experiences, activities, objectives and decisional authorities.

<sup>1</sup> The hierarchical pattern on which mass production is founded presumes the steadiness of social, economic and technological factors.

<sup>2</sup> In 1972 McLuhan & Nevitt [1] in "Take Today" suggested that electronic technologies would transform consumers into producers. Some years later, in 1980, the futurologist Alvin Toffler [2] in "The Third Wave" coined the term "prosumer", predicting the blurring of the distinction between producer and consumer due to the saturation of markets with standardized products which would have pushed towards the search for higher levels of differentiation and personalization of products.

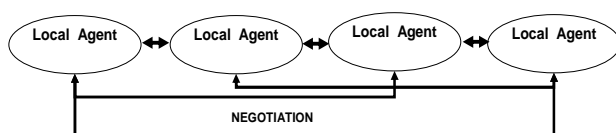
These changes call for new organization structures. Traditional hierarchical systems show several inadequacies to work in these new business environments: a) they strongly limit the reconfiguration capacity, the reliability and the growth capacity of the organization [7]; b) their complexity grows together with the size of the organization [8]; c) communication among the elements of the system is strictly determined *ex ante* and vertically limited [9]; d) the structure's modules may not take initiatives, therefore reducing the system's readiness to react thus resulting not agile in turbulent environments environment [10]; e) the structure is expensive to build and to maintain. Heterarchical systems do not have the limits of hierarchical systems, as they are able to obtain flexibility and adaptability to external stimuli. In heterarchical systems every hierarchy is banned and power is given to the single "agents"<sup>3</sup> of the system. Agents interact with their environment and with other agents according to their own attributes and aims. Control is based on negotiation due the lack of hierarchy.

In the field of *artificial intelligence*, the term agent is used to define the intelligent elements of a system who observe and act in the environment as entities capable of awareness and purposive behaviors; such agents must have the following attributes [11,12]:

- Autonomy - they act without the help or guide of any superior entity;
- Social ability-they interact with other agents;
- Reactivity-they perceive their environment and respond rapidly to changes;
- Pro-activity-they are able to have initiative and specific behaviors for a specific scope.

For example, in a heterarchical manufacturing system, the relation between the work station and supply orders is such that every supplier has direct contact with the work station in order to exploit all possible options to face unexpected fluctuations in supply and/or demand.

In spite of their agility, heterarchical systems are not able to operate following predefined plans, hence their behavior is hardly predictable, increasing variability in systemic dynamics. Heterarchical structures work well in simple, non complex and homogeneous environments



**Figure 1. Heterarchical control architecture**

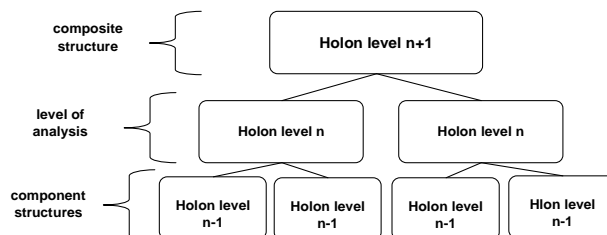
<sup>3</sup> Most of the system architectures based on agents are heterarchical. Nevertheless there are also agent based systems which do not adopt heterarchical control.

<sup>4</sup> Holistic scientific paradigm focusing on the study of Complex Adaptive Systems (C.A.S.).

with abundance of resources [10], while in complex environments they can bring to instability because of their unpredictability; moreover, with scarcity of resources, they are not able to act efficiently due to the lack of planning.

It is therefore necessary to conceptualize and implement a system able to assure both performance and reactivity at the same time. The answer to this challenge could come from the theories on living organisms and social organizations, which, if applied to the business, present a representation of the firm as a living system. The *holonic* paradigm emerges in this research stream, amidst the holistic<sup>4</sup> approach and the vital systemic approach [13]. The holonic paradigm stems from the thoughts of Arthur Koestler [14] who underlined how complex systems can originate only if they are composed by stable and autonomous sub-systems, which are able to survive turbulences and, at the same time, can cooperate forming a more complex system. Koestler underscores that analyzing both the biological and the physical universe shows that, it is necessary to take into account the relations between the whole and the part of the entities we observe. To understand the abearance of the world, according to Koestler, is not enough to study atoms, molecules, cells individuals or systems as independent entities, but it is crucial to consider such unities as simultaneously part of a larger whole; in other words, we have to consider it as a *holon*. The term holon is a combination of the ancient Greek "ὅλος" with the meaning of "whole" and the suffix "ῶν" meaning "entity" or part; thus the whole is made of parts which unlike atoms are also entities. The holon is, indeed, a whole which includes, simultaneously, the elements or the subparts which form it and give it structural and functional meaning. Holons act as intelligent, autonomous and cooperative entities working together inside temporary hierarchies called "*holarchies*". A *holarchy* is a hierarchy of self-regulating holons working, in coordination with their environment, as autonomous wholes which are hierarchically superior to their own parts and, at the same time, are parts dependent by the control of superior levels. Figure 2 shows the general relationship between holon and holarchy.

Holons of the same level process elements and information coming from lower level holons and they transfer the results to higher level ones for further processing. Processes of holons belonging to level 'n' hence originate from



**Figure 2. Holons and holarchies**

process of 'n-1' level subordinated holons and at the same time are the input for the processes of 'n+1' superior holons. [15,16]. The strength of the holonic approach resides in the concept of holarchy, which allows the development and implementation of extremely complex systems which are able to use resources efficiently, are resilient to disturbances and, at the same time, adaptable to changes of the environment. What makes the holonic system extremely effective in turbulent environments is that, inside a holarchy, holons are able to dynamically create and change hierarchies and also to participate to different hierarchies simultaneously. The holonic system can therefore be defined as a global and organized entity made of interrelations among highly self-regulating operative units which are able to cooperate with each other, keeping their autonomy, seeking shared results and common aims. It is possible to find the three pillars of holonic systems [17]:

1) the *shared-value system* in the organization allows the spontaneous and continuous interaction among groups of people who are far from each other and are not linked by legal or ownership ties, in order to take advantage of the economies of cooperation and of the increased stability of the system. Examples of shared value systems are some of the elements of lean production, that are often embedded in the company's vision, such as the principle of continuous improvement (kaizen);

2) the *distributed network information system* which is the neural sub-system [18] supporting real time supply of information between operating units which consents the pursuit of maximum income by better exploiting the coming business opportunities;

3) the *autonomous distributed hierarchy* which is based on the ability of each autonomous part to become leader according to requirements of specific situations caused by the turbulent changes in the environment. Every entity is able to directly interact with other entities without mediation. Due to this property in a holonic system every holon has potentially the same importance and the same responsibility; the involvement of a holon as

operative unit is based on its knowledge and competencies and is not a consequence of predefined leadership.

### 3. The Holonic Production System

The Holonic Production System (HPS) can be a valid choice to overcome the problems of traditional production systems' architectures, thanks to its capability to adapt and react to changes in the business environment whilst being able to maintain systemic synergies and coordination. The HPS is made of holons seen as functional production units which are simultaneously autonomous and cooperative. These holons can be represented as networked agents which define different levels of a system [19].

Every element represented in Figure 4 is a holon (work cell, factory, firm, supply chain). At the supply chain level the interaction among firms, their suppliers and their clients takes place. It is possible to determine a subsystem for each firm in the supply chain level, this subsystem is an enterprise level holon. In the enterprise there is cooperation among factories and sales departments. Inside each factory there are several working cells which interact with each other; the working cell is the basic level of the holarchy described which is self-controlled by the interaction among men and machines [20].

### 4. Cases of Application of Theory and Further Possible Developments of Research

Although the holonic approach could represent a valid solution in order to pursue the necessary levels of agility of production systems, they have been scarcely implemented in practice and even less studied from a business studies perspective. Furthermore few studies of implementation of holonic-like systems can be found in the literature. Shen [21] noted that IBM has been one of the first firms to adopt a system based on intelligent agents to avoid bottlenecks and smooth production. Jennings & Bussman [22] developed a way to implement a standard modules system, where each module is flanked by an

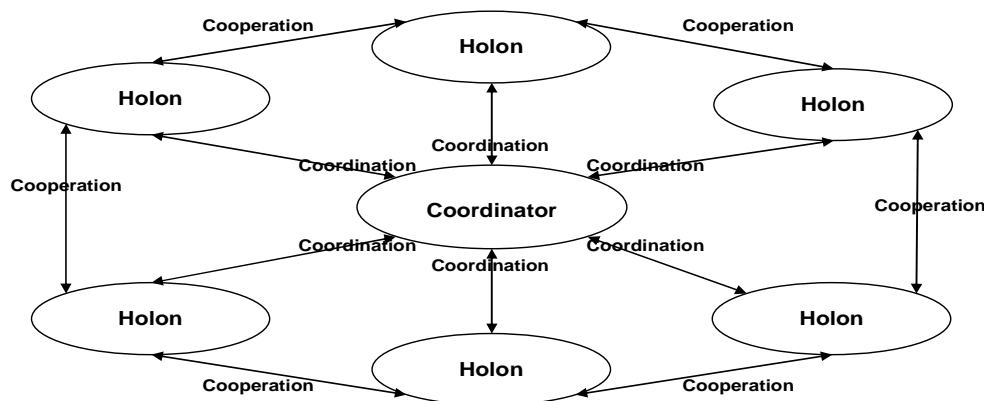


Figure 3. Architecture of an autonomous distributed system

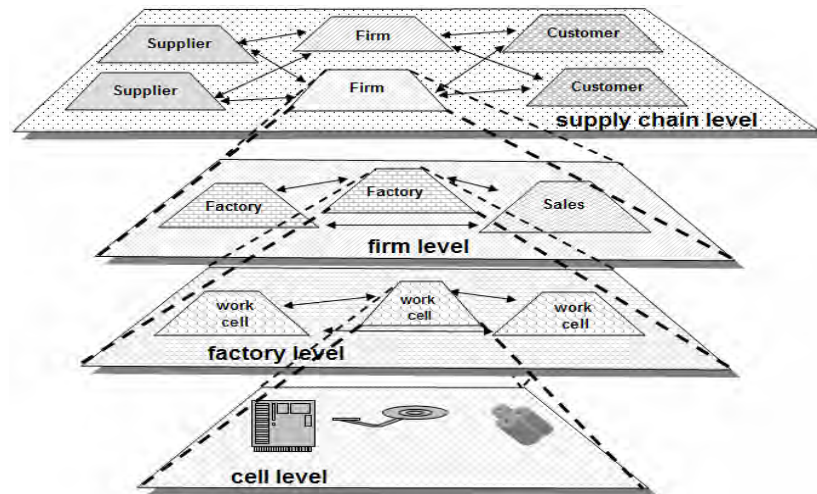


Figure 4. Production holarchy and relation among the different holon-levels

intelligent agent in order to compose a holon which becomes the building block of the system; this system has been tested by Daimler-Chrysler in order to evaluate its resilience of the system. The result obtained was of 99,7% of the theoretical optimum and the system has been adopted in the Factory of Stuttgart-Untertürkheim in Germany.

The purpose of these notes is to show the benefits of further research on cases of implementation of HPS from a business organization studies perspective. Very little research on this topic has been done outside the field of business engineering and computer science; the study of this topic from a different perspective can shed the light on new aspects and new applications of the theory.

The HPS is surely not easy to implement in a real factory, nevertheless a step-by-step approach for the introduction of this system in those industries where the need for flexibility goes together with the scarcity of resources and margins, can become the way for the factory of the XXI century.

## REFERENCES

- [1] M. McLuhan and B. Nevitt, "Take today: The executive as dropout. B. J. Harcourt, 1972.
- [2] A. Toffler, "The third wave," Bantam, 1980.
- [3] D. M. Dilts, N. P. Boyd, and H. H. Whorms, "The evolution of control architectures for automated manufacturing systems," *Journal of Manufacturing Systems* 10(1), pp. 79–93, 1991.
- [4] T. J. Crowe and E. J. Stahlman, "A proposed structure for distributed shopfloor control," *Integrated Manufacturing Systems* 6 (6), pp. 31–36, 1995.
- [5] R. Blackwell and K. D. Blackwell, "The century of the consumer: Converting supply chains into demand chains," *Supply Chain Management Review*, pp. 22–32, 1999.
- [6] F. Frederix, "From production to a product perspective," *New Industrial Scenario*, In Yoon S. *et al.* *Evolution of Supply Chain Management, Symbiosis of Adaptive Value Networks and ICT*, Kluwer Academic Publishers, Norwell, 2004.
- [7] L. Gou, P. B. Luh, and Y. Kyoya, "Holonc manufacturing scheduling: Architecture, cooperation mechanism, and implementation," *Computers in Industry*, 37 (3), pp. 213–231, 1998.
- [8] J. Hatvany, "Intelligence and cooperation in heterarchic manufacturing systems," *Robotics & Computer-Integrated Manufacturing*, 2 (2), pp. 101–104, 1985.
- [9] H. V. Brussel, L. Bongaerts, J. Wyna, P. Valckenaers, and T. V. Ginderachter, "A conceptual framework for Holonic manufacturing: Identification of manufacturing holons," *Journal of Manufacturing Systems*, 18 (1), pp. 35–52, 1999.
- [10] P. Valckenaers, E. Bonneville, H. V. Brussel, L. Bongaerts, and J. Wyna, "Results of the holonic control system benchmark at the kuleven," *Proceedings of CIMAT*, pp. 128–133, 1994.
- [11] T. Moyaux, B. Chaib-draa, and S. D'Amours, "Supply chain management and multiagent systems: An overview," In B. Chaib-draa, J. P. Müller (editor), *Multiagent based Supply Chain Management*. Springer, 2006.
- [12] M. Paolucci and R. Sacile, "Agent-based manufacturing and control systems," CRC Press, 2005.
- [13] G. M. Golinelli, *L'approccio sistemico al governo dell'impresa-L'impresa sistema vitale Vol. 1*. CEDAM, Italy, 2000.
- [14] A. Koestler, "The ghost in the machine," Arkana, 1967.
- [15] M. Mesarovic, D. Macko, and Y. Takahara, "Theory of hierarchical, multi-level systems," Academic Press, 1970.
- [16] P. Mella, "La Rivoluzione Olonica: Oloni, olarchie e reti oloniche," *Il fantasma del kosmos produttivo*, FrancoAngeli, Italy, 2005.
- [17] C. Sacconi, "Il Sistema Olonico," *Sistemi&Impresa*, 2, pp. 29–45, 1996.
- [18] M. A. Arbib, "The handbook of brain theory and neural networks," MIT Press, Boston, 1995.
- [19] M. Ulieru and M. Cobzaru, "Building holonic supply chain management systems: An e-logistics application for the telephone manufacturing industry," *IEEE Transactions on Industrial Informatics* 1 (1), pp. 18–30, 2005.
- [20] G. Dominici, "Il contesto istituzionale nipponico e l'evoluzione della "lean production," Aracne, Roma, Italy, 2007.
- [21] W. Shen, "Distributed manufacturing scheduling using intelligent agents," *IEEE Intelligent Systems*, pp. 88, 2002.
- [22] N. R. Jennings and S. Bussmann, "Agent-based control systems," *Why are they suited to engineering complex systems?*. *IEEE Control Systems Magazine*, 61, 2003.

# The Evaluation of Enterprises' Sustainable Superiority Degree of Technical Innovation Based on DEA Method

Bing Jiang, Lu Yang & Jun Fang

School of Management, Hefei University of Technology, 230009, China

Email: bing-jiang@sohu.com, yanglu\_suosuome@tom.com, ice1029@gmail.com

Received August 31<sup>st</sup>, 2008; received November 20<sup>th</sup>, 2008; accepted December 5<sup>th</sup>, 2008.

## ABSTRACT

*The paper at first expounds two new concepts of sustainability degree and superiority degree, and finds a method of DEA of relative efficiency index to rank all decision-making-units having been identified efficient. On the basis of them, the definition and arithmetic are given to analyze them. Then seven input and output techno-innovation indexes which belong to six automobile listed companies from 2002 to 2005 are used to evaluate enterprises' sustainability degree and superiority degree by vertical and horizontal calculation based on DEA method. At last, we get general conclusions about enterprises' sustainable superiority degree by weighting calculation.*

**Keywords:** technical innovation, DEA, relative efficiency, sustainable superiority evaluation

## 1. Introduction

DEA model (Data Envelopment Analysis) is one of mathematical programming approaches to analyze decision-making-units' (DMUs) relative efficiency. To some enterprises' inputs and outputs which are observed in a period of time, we apply DEA model and acquire not only vertical evaluation of enterprises' sustainable innovation based on time as DMU, but also horizontal relative evaluation of enterprises' innovation performance based on enterprises as DMU.

However, in traditional DEA model, DMU can only be examined as "effective" or "ineffective" but can not be ranged by their weights. Furthermore, each DMU' weight is computed from the most beneficial angle, which, obviously, will give birth to that most of DMU even all will be described as effective. Consequently, the traditional DEA model will affect the further comparable analysis. The purpose of this paper is to calculate degree of technical innovation sustainability and superiority according to the definition of relative efficiency index and to present a concept of sustainable superiority degree. In this paper, we apply the ideal DMU and obtain a standard weights vector as referenced ones, on the basis of these, the values of relative efficiency index are available. We select six listed companies in automobile industry as our research objects, and the analysis data are based on seven indexes, which reflect the innovations and originate from the 2002–2005 years' reports of relating companies.

## 2. DEA Model Based On Relative Efficiency Indexes

### 2.1. The Basic Principle of Traditional DEA Model

Technical innovation in a specific enterprise can be described by a group of innovation indexes of input-output. We use vector  $A_j$  to stand for input of decision-making-unit  $j$  (DMU $_j$ ) and vector  $B_j$  for its output, i.e.:  $A_j = (a_{1j}, a_{2j}, \dots, a_{mj})^T$ ,  $B_j = (b_{1j}, b_{2j}, \dots, b_{sj})^T$ . The dual programming model of DMU $_j$  can be constructed as below:

$$\begin{aligned} D: \quad & \min \quad V_D = y \\ \text{s.t.} \quad & \sum_{j=1}^n a_{ij} x_j + s_i = a_{ij_0} y, i = 1, 2, \dots, m \\ & \sum_{j=1}^n b_{kj} x_j - t_k = b_{kj_0}, k = 1, 2, \dots, s \\ & x_j \geq 0, j = 1, 2, \dots, n; s_i \geq 0, i = 1, 2, \dots, m; t_k \geq 0, k = 1, 2, \dots, s \end{aligned} \quad (1)$$

Where  $x_j$  stands for the  $j$ th DMU' decision-making variable;  $s_i$  for the slack variable of the  $i$ th input index;  $t_k$  for the surplus variable of the  $k$ th output index; and  $y$  for the input proportional variable.

If non-Archimedes infinitesimal quantity is applied, then the equation will appear to:

$$\begin{aligned} D\epsilon: \quad & \min \quad V_D(\epsilon) = y - \epsilon \left( \sum_{i=1}^m s_i + \sum_{k=1}^s t_k \right) \\ \text{s.t.} \quad & \sum_{j=1}^n a_{ij} x_j + s_i = a_{ij_0} y, i = 1, 2, \dots, m \\ & \sum_{j=1}^n b_{kj} x_j - t_k = b_{kj_0}, k = 1, 2, \dots, s \end{aligned} \quad (2)$$

$$x_j \geq 0, j = 1, 2, \dots, n; s_i \geq 0, i = 1, 2, \dots, m; t_k \geq 0, k = 1, 2, \dots, s$$

On the assumption that optimum solutions of equation (2) are  $x_j^*$  ( $j=1, 2, \dots, n$ ),  $s_i^*$  ( $i=1, 2, \dots, m$ ),  $t_k^*$  ( $k=1, 2, \dots, s$ ),  $y^*$ , some effective judgment rules will go like this:

(1)  $y^*=1$ , and  $s_i^*=0$  ( $i=1, 2, \dots, m$ ),  $t_k^*=0$  ( $k=1, 2, \dots, s$ ), then DMU $_{j_0}$  is DEA effective, and its economic significance means the optimal innovative efficiency and constant return to scale simultaneously.

(2)  $y^*=1$ , we can conclude that DMU $_{j_0}$  is DEA ineffective. Its economic significance means it is not simultaneous to reach optimal innovative efficiency and constant return to scale.

## 2.2. DEA Model Based on Relative Efficiency Index

To overcome the deficiency of traditional model which is unable to discriminate the differences among different efficient units, we apply ideal DMU. To each input index, if minimum input vectors are made of minimums of all DMU on the assumption, then  $A_{\min}=(a_{1\min}, a_{2\min}, \dots, a_{m\min})$ ; likewise, if maximum output vectors are made of maximum of all DMU, we mark  $B_{\max}=(b_{1\max}, b_{2\max}, \dots, b_{s\max})$ , then we can say  $(A_{\min}, B_{\max})$  stand for innovation activity related to ideal DMU. We add the ideal DMU to DEA model, and under the thought of using DEA method to calculate weight, an efficient index model can be constructed as below:

$$\begin{aligned} \max \quad & h_{\max} = \frac{u^T B_{\max}}{v^T A_{\min}} \\ \text{s.t.} \quad & \frac{u^T B_j}{v^T A_j} \leq 1, \quad j = 1, \dots, n \\ & \frac{u^T B_{\max}}{v^T A_{\min}} \leq 1 \\ & u_k \geq 0 \quad k = 1, \dots, s \\ & v_i \geq 0 \quad i = 1, \dots, m \end{aligned} \quad (3)$$

The model is used to obtain a series of weights  $v_1, v_2, \dots, v_m; u_1, u_2, \dots, u_m$ , which can optimize the efficiency indexes of ideal DMU if all observed DMU' indexes meet the inequality:  $h_j \leq 1$ . We take these weights as referenced weights to calculate the relative efficiency indexes of other units. Because of their inferior efficiencies to the ideal DMU, we can tell the difference between excellent DMU and bad ones easily.

Charnes-Cooper transformation is used, ordered:  $t=1/(v^T A_{\min})$ ,  $\omega^T = tv^T$ ,  $\mu^T = tu^T$ . There were:  $h_{\max} = (u^T B_{\max}) / (v^T A_{\min}) = t(u^T B_{\max}) = \mu^T B_{\max}$

But:  $h_j = (u^T B_j) / (v^T A_j) = t(u^T B_j) / t(v^T A_j) = \mu^T B_j / \omega^T A_j \leq 1$ ,  $j=1, 2, \dots, n$

After transforming, it becomes an equal model of linear programming:

$$\begin{aligned} \max \quad & \mu^T B_{\max} \\ \text{s.t.} \quad & \mu^T B_j \leq \omega^T A_j \quad j = 1, 2, \dots, n \end{aligned}$$

$$\begin{aligned} \mu^T B_{\max} &\leq \omega^T A_{\min} \\ \omega^T A_{\min} &= 1 \end{aligned} \quad (4)$$

$$\omega_k \geq 0, \quad k=1, 2, \dots, m; \quad \mu_i \geq 0, \quad i=1, 2, \dots, s$$

We call  $h_{\max}^*$  ( $h_{\max}^* = \mu^{*T} B_{\max}$ ) as efficient index of ideal DMU and  $h_j^*$  ( $h_j^* = (\mu^{*T} B_j) / (\omega^{*T} A_j)$ ) as relative efficient index of ideal DMU. Consequently, all DMU can be ranked based on its relative efficiency index.

## 3. Sustainable Superiority Degree of Technical Innovation in Enterprises

Sustainability degree of technical innovation is a new concept that is mainly used to describe stability and durative of technical innovation process in an enterprise when it is considered as the main unit of technical innovation activities. It isn't only a matter of time, but also refers to development of the technical innovation activities in enterprise and furtherance of the innovative spirit in corporate culture.

Technical innovation, which plays a role of encouraging and promoting innovations in other departments, is the inexhaustible motive force for sustainable development of enterprises. The relation between techno-innovation and techno-innovation sustainability just likes the relation between qualitative change and quantitative change, techno-innovation is the unity of gradual change and abrupt change. On one hand, techno-innovation can lay a solid foundation and offer technical supports for innovations in other areas gradually and also the guidance for future development. On the other hand, when technical innovation is promoted to a certain phase, it will break through the outdated formation and captures the new technology primacy, which is the demonstration of techno-innovation sustainability. Continuous technical innovation presents a premise to techno-innovation sustainability, and in turn, techno-innovation sustainability provides assurance for techno-innovation.

Superiority degree is used to describe the degree of primacy for a techno-innovation performance in a certain domain. To an enterprise, higher sustainable superiority means larger market share and more powerful competitive advantage. Technical innovation is the only access for the enterprise to be the leader in a market, and then, the innovation performance depends on its superiority.

According to the description above, sustainable superiority degree of technical innovation could be defined as bellow:

Sustainability degree of technical innovation describes how techno-innovation performance keeps going with the time changing, measured by the slope of linear regression model which uses enterprises' vertical efficiency indexes as the vertical axis and years as the horizontal axis. The least

square estimator is applied to calculate the parameter of linear regression model, the formula appears to be:

$$b1_j = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(y_{ij} - \bar{y}_j)}{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2} \quad j = 1, 2, \dots, m \quad (5)$$

Where,  $y_{ij}$  stands for vertical efficiency index of the  $j$  enterprise in year  $x_i$ ,  $n$  means how many years.

Superiority degree of technical innovation represents relative level of technical innovation performance among all observed enterprises. If I use the sum of differences between horizontal efficiency indexes of each year and average relative efficient index as the measurement. The formula is as follows:

$$b2_j = \sum_{i=1}^n (H_{ij} - \bar{H}_i) \quad j = 1, 2, \dots, m \quad (6)$$

Where,  $H_{ij}$  stands for horizontal efficiency indexes of the  $j$  enterprise in year  $i$ .

Sustainable superiority degree of technical innovation represents comprehensive performance of an enterprise considering both the time and its counterparts. We use the method of normalized weighted average of sustainability and superiority degree of technical

innovation to calculate them:

$b = w_1 b1 + w_2 b2$ , weight  $w_1$  and  $w_2$  can be evaluated based on your preference. If you pay more importance to growing character, weight  $w_1$  should be higher; or smaller.

#### 4. Empirical Study

We select six automobile listed companies as our research objects, and apply DEA model as mentioned above to analyze the indicator of innovation input-output during 2002 to 2005 from vertical and horizontal two aspects. On the vertical aspect, we take each year as DMU and then learn how is technical innovation performance and returns to scale changing with time, which, in fact, reflects the superiority characteristic of techno-innovation; in contrast, On the horizontal aspect, we use individual enterprise as DMU and relatively rank each enterprise with their techno-innovation efficiency and returns to the scale, which, correspondently reflects the primacy of innovative enterprise among its counterparts. To the convenience of building the model, we categorize the data to two types: data of each year for all companies (vertical section) as in table 1, and data of each company for all years (transect).

**Table 1. The values of techno-innovation indexes of each enterprise in four years (vertical section)**

	DFAC				JAC			
indexes\years	2002	2003	2004	2005	2002	2003	2004	2005
ratio of technical staffs (%)	22.8	21	20.1	20.84	10.58	11.57	8.2	9.24
ratio of expenditure in technical development in the main business income (%)	0.09	5.71	0.07	0.01	0.12	0.28	0.55	0.53
net profits of fixed assets (10million)	179.48	207.95	110.82	132.83	22.99	68.21	117.16	140.88
main business income (10million)	443.08	700.06	585.14	610.01	198.94	344.3	545.01	806.89
income taxes (10million)	9.23	10.39	5.5	8.1	1.53	2.97	8.09	12.99
return on total assets (%)	10.86	9.32	9.66	6.55	6.49	7.28	8.28	8.54
rate of net profit (%)	12.13	8.81	10.76	7.64	4.47	3.85	3.8	3.98
	CHANGHE				YXMC			
indexes\years	2002	2003	2004	2005	2002	2003	2004	2005
ratio of technical staffs (%)	10.69	12.62	12.62	13.6	7.78	6.4	6.2	7.33
ratio of expenditure in technical development in the main business income (%)	0.44	1.56	0.67	1.53	0.24	0.35	1.1	0.16
net profits of fixed assets (10million)	88.87	86.35	95.37	96.34	29.6	29.36	32.12	31.34
main business income (10million)	332.85	437.3	415.05	313.87	101.98	87.65	79.87	79.53
income taxes (10million)	1.96	2.79	1.49	1.4	0.38	0.63	0	0.02
return on total assets (%)	2.64	2.19	1.31	-1.32	4.14	1.93	-11.84	-5.68
rate of net profit (%)	2.61	1.94	1.22	-1.53	5.22	2.54	-18.47	-8.71
	ANKAI				JMC			
indexes\years	2002	2003	2004	2005	2002	2003	2004	2005
ratio of technical staffs (%)	12.16	14.14	12.2	9.46	11.98	11.53	11.54	11.97
ratio of expenditure in technical development in the main business income (%)	0.23	0.57	0.92	0.26	1.32	1.56	1.6	1.11
net profits of fixed assets (10million)	30.41	32.87	28.07	31.13	166.82	155.47	145.42	130.55
main business income (10million)	44.59	48.8	73.25	91.53	337.91	427.09	509.49	577.07
income taxes (10million)	0.11	0.07	0	0	0.48	2.33	4.84	6.79
return on total assets (%)	1.01	-4.09	1.82	0.87	2.75	8.01	11.8	9.36
rate of net profit (%)	2.3	-8.14	2.71	1.11	2.98	6.71	8.81	6.7

The origin of the data: Shanghai stock exchange; Shenzhen stock exchange; China finance online and so on

**Table 2. The results of traditional DEA model**

years	vertical section↓						transect→					
	DFAC	JAC	CHANGHE	YXMC	ANKAI	JMC	DFAC	JAC	CHANGHE	YXMC	ANKAI	JMC
02	1	1	1	1	1	0.585	1	1	1	1	0.4	0.933
03	1	0.895	1	1	0.893	0.819	1	1	1	0.591	0.293	1
04	1	1	1	0.94	1	1	1	1	0.869	0.47	0.494	1
05	1	1	0.671	1	1	1	1	1	0.568	0.443	0.513	1

**Table 3. The results of relative efficiency index based on DEA model**

years	vertical section↓						transect→					
	DFAC	JAC	CHANGHE	YXMC	ANKAI	JMC	DFAC	JAC	CHANGHE	YXMC	ANKAI	JMC
02	0.783	0.589	0.761	0.992	0.613	0.564	0.341	0.271	0.157	0.430	0.121	0.160
03	0.957	0.604	0.847	0.802	0.00002	0.740	0.305	0.272	0.317	0.125	0.032	0.339
04	0.529	0.970	0.804	0.185	0.998	0.882	0.308	0.377	0.189	0.054	0.227	0.442
05	0.752	0.887	0.564	0.351	0.182	0.963	0.352	0.550	0.215	0.195	0.403	0.557

**Table 4. The evaluation of sustainability and superiority degree of enterprises' techno-innovation**

	DFAC	JAC	CHANGHE	YXMC	ANKAI	JMC
sustainability degree b1(0.6)	-0.05228	0.12607	-0.0634	-0.25396	-0.02936	0.13411
superiority degree b2(0.4)	0.18289	0.34687	-0.24523	-0.31902	-0.33991	0.3744
sustainable superiority degree b	0.60459	0.97217	0.34765	0.01349	0.34726	1
the order	3	2	4	6	5	1

According to formula (2), we obtain the whole DEA validity of each enterprise in four years and their validity in each year, as displayed in table 2. From it, we can see that, regarding of the enterprise itself, DFAC is the only one which has been simultaneously on the state of optimal innovation efficiency and constant returns to scale simultaneously for four years, while JAC, CHANGHE, ANKAI and YXMC each have one year in DEA inefficiency, and JMC has two years in DEA inefficiency. Compared these enterprises with each other, we can find that in year 2002 and 2005 there were four enterprises DEA efficiency, and in year 2003 and 2005 there were three. Among them, DFAC and JAC share the leadership for four years. However, no matter from vertical or horizontal section, traditional DEA analysis will only result in many efficient DMU but can not tell the differences between them, so here model (4) is used to further discriminate degree in efficiency of those DMU. The results are listed on Table 3.

According to the arithmetic of sustainability and superiority degree mentioned above, we order their correspondent weights equal to 0.6 and 0.4, then the evaluation of all enterprises' sustainability and superiority can be acquired as in Table 4. The rank can be seen in Table 4.

As in Table 4, JMC wins the No.1 in sustainability and superiority, closely followed by JAC, and DFAC ranks the third. Comparably, CHANGHE, ANKAI, YXMC

are lagged far behind them. And the conclusion is also consistent with what really has happened. The new concept, this paper aims to present, sustainability and superiority, can not only evaluate enterprises' innovation performances among counterparts and also on the order of time in this paper. Because of the constraint of observed samples and certain historical period, the conclusion can only be used as reference. Considering of the pertinence and practicability, you can select specific sample for observation according to your own needs.

## REFERENCES

- [1] Q. L. Wei, "Data envelopment analysis," [M], Beijing: Science Press, 2005.
- [2] J. Y. Yan and B. J. Sun, "System analysis and evaluation of enterprises' technical innovation," [M], Beijing: China Financial & Economical Publishing House, 2002.
- [3] G. Li, X. Y. Shen, and Y. M. Wang, "A new method for ranking DMUS," [J], Forecasting, 2000(4): pp. 51-53.
- [4] D. Du, "The evaluation of enterprises technical innovation with Data Envelopment Analysis," [J], Systems Engineering-Theory Methodology Applications, 2002, 10(1): pp. 82-84.
- [5] B. Jiang and C. Q. Zhang, "DEA on enterprise's technological progress and example research," [J], Systems Engineering-Theory & Practice, 7(7): pp. 65-70, 2002.



# Examining Human Resource Competencies and Their Relationship to the Success Factors of HR Profession

Choi-Sang Long

School of Business & Management, Southern College Malaysia  
Email: cslong\_1@yahoo.com

Received October 4<sup>th</sup>, 2008; revised November 11<sup>th</sup>, 2008; accepted November 25<sup>th</sup>, 2008.

## ABSTRACT

*This study examines competencies of Human Resource (HR) professionals in the manufacturing companies of Malaysia. The Human Resource Competency Survey (HRCS) model is used in this study. The competencies that are examined in this study are business knowledge, strategic contribution, HR delivery, personal credibility and HR technology. All these competencies will be tested whether or not they are significantly related to a firm's performance. Furthermore, researcher wanted to study the relationship of these competencies with variables such as experience, education level, firm's size and salary of the sample. The sample employed here consists of HR professionals from Malaysian manufacturing companies in Johor, the southernmost state of Malaysia.*

**Keywords:** human resource, competencies, manufacturing, success factors, performance

## 1. Introduction

Human resource professionals needed to function strategically. To play more critical roles more effectively, HR professional must master the necessary competencies, and that mastery of HR knowledge comes only from being familiar with the concepts, language, logic, and practices of HR that are the result of research and training. Furthermore, mastery of the above abilities comes from being able to apply the knowledge within specific business settings [1].

Nowadays, competencies are used in many facets of human resource management, ranging from individual functions such as recruitment and performance management to organizational strategic planning and design of organizational structure and culture. HR competencies are said to be a set of characteristics contributing to the effective HR performance that enables an organization to carry out its business strategies in a competitive market. However, many HR executives are not invited to the strategic planning table because they have failed to display the required competencies [2]. In fact, it is suggested that the competency level of the HR manager has an influence on whether he or she is able to get into the executive board chamber [3].

## 2. Review of the Literature

There are several major studies available on HR competencies [4]. One study surveys 3000 HR professionals, consultants, line executives and academicians. That study reports that line executives thought that computer literacy was the

most critical HR competence; while academicians argue that a broad knowledge of and a clear vision for HR were the most important issues, and consultants believe that ability to change things is the most important factor in the excellent or HR performance. Another study, examining 300 HR professionals from various sectors, establishes a set of core HR competencies consisting of leadership style, management intuition, functional abilities and personal attributes [5]. One of HR competency surveys suggests that HR professionals needed to be more knowledgeable about financial management, external competition and customer demands [6].

The survey data of the Human Resource Competency Study (HRCS) were collected in 2003 under the initiative of the University of Michigan. The study was carried out online (web-based). The respondents of the European HRCS were HR professionals and line managers of multinational companies located in Europe [7]. In this survey, five domain factors emerged as making a difference in terms of performance. The domains are as follows:

### 1) Strategic Contribution

High-performing companies have HR professionals involved in the business at a strategic level. These HR professionals manage the culture, facilitate rapid change, and are involved in the strategic decision making and create market-driven connectivity of the operation [7]. In this competency area, culture management, rapid change efforts, and a business partner role along with customer focus emerged as important factors for HR professionals,

making their impact on their organizations' financial performance significant [8].

## 2) Personal Credibility

HR professionals must be credible to both their HR counterparts and the business line managers whom they serve. They need to promise and deliver results and establish a reliable track record. Furthermore, working well with others by building good relationship is vital in developing the ability to work together with others effectively. In addition, HR professionals must have effective writing and verbal communication skills [7]. The findings of the study by [7] correspond with the prior research of [9], who found that the personnel directors require professional competence in social skills to develop effective interpersonal relations with other board directors. This is one of the competencies of personal credibility.

## 3) HR Delivery

HR professionals deliver both traditional and operational HR activities to their business in four major categories. First, by designing developmental programs and challenging work experiences. This is done by offering career planning services, and facilitating internal communication processes. These efforts include both individual development as well as organisation-wide development. Second, by structuring and HR measurement: restructuring the organisation, measuring impact of HR practices, and managing global implications of HR practices. Third, by attracting, promoting, retaining, and out-placing appropriate people. Finally, by performance management in terms of designing performance-based measurements and reward systems and providing competitive benefit packages [7].

## 4) Business Knowledge

To become key players in the organisation, HR professionals must understand the business or industry of the company they serve. Key areas of knowledge include applied understanding of the integrated value chain (how the firm horizontally integrates) and the firm's value proposition (how the firm creates wealth). The labour factor, representing institutional constraints such as labour legislation, is the third factor that constitutes the domain of business knowledge [7]. Human resources professionals must understand how their business or agency operates. This includes the organization's strategy, how the organization makes money or achieves its primary purpose, its technological processes and organizational capabilities, etc. [10].

## 5) HR Technology

HR professionals need to be able to leverage technology for HR practices and use e-HR/web-based channels to deliver value to their customers [12]. [12] further argues that the pace of technological innovation will continue to accelerate. HR can take advantage of these changes by automating HR processes and becoming more effective in communicating with its internal /external customers. More importantly, by absorbing the latest tech-

nology, HR can project a forward looking image that will help it earn the respect of skeptical colleagues. According to a recent survey by Society for Human Resource Management, the top workplace trend identified was technology [12].

## 2.1. The Relationship between HR Competencies and a Firm's Performance

Researchers in the field of strategic human resource management have emphasized that human resource practices may lead to higher firm performance and be sources of sustained competitive advantages [13]. Competing in today's tumultuous global economy provides additional challenges to the HR function in creating the expected value to create and sustain competitive advantages. To function effectively, HR professionals must master the necessary competencies, and that mastery of HR knowledge comes from knowing the concepts, language, logic, research, and practices of HR [1]. Furthermore, mastery of these abilities comes from being able to apply that knowledge to specific business settings.

[14] have convincingly argued that HR professionals need to become more effective strategic business partners. [15] argue further that HR professionals must make the transition from being strategic business partners to becoming contributors in their organizations. Several studies have shown a positive relationship between certain HR competencies and firm performance [7,13,16,17,18] have stated that there is an emerging group of human resource professionals who see the opportunity to turn human capital strategy into a long-term competitive advantage. They observe that in the 1990s there was a wake-up call for the human resource profession. More than ever, organizations now seek greater creativity and productivity from people. Part of the strategy in being creative and maximizing productivity is to possess the necessary competencies for enabling these outcomes.

The research by Brockbank [1] showed that HR activities positively impact business performance by approximately 10% (defined as the financial performance of the business over the last three years compared to its major competitors). Strategic contribution accounts for 43 percent of HR total impact on business performance which is almost twice the impact of any other domain. These are all reasons why competencies are being discussed by academicians and practitioners alike as ways of creating sustainable competitive advantages.

A study by [7] indicates the domain of strategic contribution is positively correlated with financial competitiveness, while the domain of HR technology is negatively correlated with this performance outcome. This is not completely in line with the global HRCS findings. In the Europe, the study found only one domain (strategic contribution) to be positively related to financial competitiveness, in contrast to the global results that suggest four out of five domains to be positively linked to finan-

cial competitiveness. Fourth, all domains reveal relatively high correlations with each other.

## 2.2. Relationship between HR Competencies and Success Factors of HR Profession

In this study, researcher wanted to see if there are any impact on HR professionals' competencies with selected success factors like the working experience, education level, firm's size and salary. [17] study found that salary was significantly related to years of HR experience as well as to self-rated competency in six of the competency areas. The six competency area are delivering HR, technical competence, managing change, understanding business, strategic contribution and accounting. Interestingly, although experience was related to self-ratings of competencies in six areas, education was not related to any of the self-ratings of competence. This result contradicts with [19] study which found that education level of the HR practitioners does influence their competencies. [19] further claim that leading companies, such as General Electric and Amoco, seek entry-level HR/IR practitioners who have the capacity to become strategic business partners and align themselves with the other functions to ensure organizational success. Specifically, those who seek admittance to General Electric's prestigious HR management development program must possess a graduate degree or bachelor's degree from a reputable HR/IR program [20].

[17] study shows the various relationships existing among the core HR strategies, factors leading to improvement in the competencies, impact of education and the competencies on compensation, and the relationship between education and the respective competencies. The results provide a reminder to HR professionals of the value of graduate degrees and other means of developing higher levels of HR technical competence, understanding the benefits of accounting, marketing, and other different functional areas in effectively developing and implementing HR strategies. By understanding the impact of the competencies on the various organizational practices, there could be a more directed strategy in developing expertise among HR professionals, hence, a more credible and effective function [17].

[21] in their study mention that HR professionals working in small companies may be so focused on solving problems on a daily basis that they have no time for focusing being a strategic business partner. Customer service skills are more important to HR practitioners in large companies. HR Practitioners from small companies gave more importance to organizational skills. Anyway, no differences emerged between upper and middle level managers of their HR competencies in this study. Furthermore, upper managers described the HR department as designing, implementing, and evaluating HR strategies and programs, acting as an internal consultant, supporting line management, and being both a member of the executive

management team and a strategic partner with top management. HR professionals from the manufacturing sector reported two roles that HR plays in their companies, that of change agent and outsourcer of HR programs. The study from [21] shows significant differences in the task of HR practitioners from large and small companies. Small companies reported that their HR departments monitored legal compliance and motivated employees and large companies reported that their HR department outsourced HR programs.

## 3. The Study

The purpose of this study is to attempt to understand better the Human Resource (HR) professional's competencies in the manufacturing sector of Malaysia. Furthermore, these finding will be tested in order to determine if they are linked to a firm's performance. In this research, the tool of Human Resource Competency Study (HRCS), which has been designed by Wayne [8], will be used to assess HR competencies among the HR professionals. HR professionals need to endure and overcome many barriers to reach the ultimate goal of becoming a strategic partner in his of her organization. It is hoped that by making this examination, we will be able to develop a realistic picture of the competencies of the HR professionals in the manufacturing firms of Malaysia. Another purpose of this study is to determine the relationship between HR competencies and possible success factors of HR profession eg. experience, education level firm's size and salary.

## 4. Research Methodology

### 4.1. Sample

The sample employed here consists of HR professionals from Malaysian manufacturing companies. All respondents work for manufacturing companies in the southernmost state of Malaysia, Johor. These industries were chosen because of their relatively large. The list of firms in the manufacturing sector was drawn from the "FMM directory of Malaysian Manufacturers 2007." Only firms with at least 50 full-time employees were studied. This is because other studies have shown that firms with smaller employment size are less likely to have HRM departments [22]. Out of the entire list in the directory, the research focused on a sample population in the Southern region of Malaysia (State of Johor). A total of about 300 firms were included in the list for this area. The total number of firms involve in this study are 32 respondents.

### 4.2. The Instrument

The data collection instrument used in this research is the a quantitative methodology with a survey instrument developed based on the five competency domains and 17 competency factors identified in the *Human Resource Competency Study* (HRSC) [8]. To improve statistical reliability, the HR technology domain was divided into two competency factors; one was operational and one

strategic. This resulted in 18 competency factors that resided within the five competency domains. A Likert scale was used on the questionnaire with the following ratings: 1–strongly disagree, 2–disagree, 3–moderately agree, 4–agree and 5–strongly agree. The respondent was asked how well they performed the competencies identified in the HRSC. A statement describing each competency factor is listed on the questionnaire. The 18 items in the instrument are arranged in groups of five competency domain.

Firm performance was measured by the self-reported rating of the respondents concerning the indicators of financial and operational performance, sustainability of profits, staffs turnover and the opportunity for growth for staff. A Likert scale was used on the questionnaire with the same rating scale as above. However, for negative questions, the rating procedures are opposite. The respondent was asked to choose the number that accurately represented their firm's performance. There are 5 items in this section that assess a firm's performance. The research used three negative questions to ensure reliability.

No researcher can completely eliminate measurement error, but he or she can reduce it in several ways, such as by conducting a pilot study. If the measurement error is reduced, the reliability of the measurement technique is increased [23]. Therefore, a pilot study was done to test the research instrument in this study. The researcher of this study used Cronbach alpha co-efficient method for this purpose. The result of the reliability test shows that the alpha value base in each domain of the instrument is between 0.62 to 0.89. Components that are tested are

strategic contribution (alpha = .89), personnel contribution (alpha value=0.67), HR delivery (alpha value=0.62), Business knowledge (alpha value=0.76), HR technology (alpha value=0.88) and firm performance (alpha value=0.86). According to [24], any measurement instrument should have reliability value of more than 0.60. [23] stressed that a measurement instrument can be considered reliable if the results are consistent from one time to another and that the reliability value is 0.70 or greater. Therefore, from the alpha value obtained, we can conclude that the research instrument is reliable and consistent.

## 5. Analysis

Table 1 show that the top nine ranking HR competency factors are from the domain of personal credibility and HR delivery. The respondents' self-rated competency shows that personal communication, legal compliance, effective relationship and performance management rank above all other factors. It would follow that respondents are most competent in these areas.

HR Professionals need to develop a relationship of trust with their clients, i.e. management team and line managers to instill confidence. These findings show positive development to the HR professionals in Malaysia because personal credibility competency is the foundation for a HR professional to become intimately involved at the strategic level in an organization once given the opportunity. Without this foundation of trust, HR Professionals may very well find themselves excluded from the strategy table. However, this study has

**Table 1. Rank order of means of HR competency factors in each domain**

Competency	Factors	Mean	Std. Deviation	Rank
Strategic Contribution	Culture management	2.91	0.89	18
	Fast change	3.06	0.95	15
	Strategic decision-making	3.03	0.69	16
	Market driven connectivity	2.97	0.78	17
Personal Credibility	Achieving results	4.16	0.68	7
	Effective relationships	4.38	0.66	3
	Personal communication	4.63	0.49	1
HR Delivery	Staffing	4.31	0.78	5
	HR development	4.16	0.95	8
	Organization structure	4.28	0.85	6
	HR measurement	4.13	0.79	9
	Legal Compliance	4.44	0.80	2
	Performance management	4.38	0.79	4
Business Knowledge	Value chain knowledge	3.31	1.03	13
	Value proposition knowledge	3.44	0.88	12
	Labor knowledge	3.53	0.88	10
HR Technology	User of technology to deliver HR services	3.31	0.69	14
	Strategic HR technology	3.47	0.95	11

found that respondents score lowest for strategic contribution competency. Based on competency factors in each domain, all strategic contribution factors score the lowest mean score in terms of mean ranking as shown in Table 1. This shows that HR professionals in Malaysian manufacturing sector are extremely weak in culture management, market driven connectivity, strategic decision-making and fast change. This result is indeed a concern because HR professionals should be able to identify and implement organizational cultures that help firms win the marketplace and successfully implement business strategies. Furthermore, if HR professionals are not able to facilitate change management processes and adapt learning to new change initiatives, they would have problems working with key individuals to ensure decisions are made quickly and to ensure resources are aligned with desired changes [8].

As shown in Table 2, not all HR professional competencies (Strategic Contribution, Business knowledge, Personal credibility, HR delivery and HR technology) have significant correlation with a firm's performance. Competencies such as strategic contribution, business knowledge and HR technology have significant correlation with firm performance. These competences obtained Spearman's rho value at 0.542 ( $p < 0.01$ ), 0.542 ( $p < 0.01$ ) and 0.373 ( $p < 0.05$ ). However, no correlation was found between personal credibility and HR delivery in assessing

**Table 2. Relationship of HR competencies to firm performance**

	Spearman's rho analysis	Firm performance
Strategic contribution	Correlation Coefficient	0.542**
	Sig. (2-tailed)	0.001
Personal credibility	Correlation Coefficient	0.144
	Sig. (2-tailed)	0.433
HR delivery	Correlation Coefficient	0.016
	Sig. (2-tailed)	0.930
Business knowledge	Correlation Coefficient	0.542**
	Sig. (2-tailed)	0.001
HR technology	Correlation Coefficient	0.373*
	Sig. (2-tailed)	0.036

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

a firm's performance. Both of these competencies obtained a Spearman's rho value at 0.144 and 0.016.

This result partially supports the earlier research from [7]. While a study by [7] indicates that the domains of strategic contribution, business knowledge, personal credibility and HR delivery are positively correlated with financial competitiveness, although there is no correlation between HR technology and financial performance.

As shown in Table 3, only two out of five HR competencies have significant correlation with independent variables. HR competencies such personal credibility has significant correlation with HR experience, education level and salary but show no correlation with the firm's size. Business knowledge competencies have significant correlation with all success factors (HR experience, education level, firm's size and salary).

The findings from this section is supported by the study from [17] that indicates that salary was significantly related to competency such as understanding business knowledge. [17] study also shows that the impact of education and the competencies on compensation. The result provide a reminder to HR professionals of the value of education level such as graduate degrees developing higher levels of HR technical competence, understanding the benefits of accounting, marketing, and other different functional areas in effectively developing and implementing HR strategies. This study also indicates that HR experience is significantly related to HR competencies such as delivering HR, strategic contribution, understanding the business and technical competence.

**Table 3. Correlations between HR competencies and HR experience, education level, firm size and salary competencies**

	HR experience	Education level	Firm's size	Salary
Strategic contribution	0.185	0.078	0.326	0.138
Personal credibility	0.436*	0.549**	0.294	0.612**
HR delivery	0.174	0.011	0.231	0.244
Business knowledge	0.367*	0.468**	0.417*	0.443*
HR technology	0.060	0.039	0.213	-0.069
HR experience	1.000	0.442*	0.669**	0.676**
Education level	0.442*	1.000	0.430	0.613
Firm's size	0.669**	0.430*	1.000	0.576**
Salary	0.676**	0.613**	0.576**	1.000

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

The study from [19] show significant differences in the task of HR practitioners from large and small companies (firm's size). Small companies reported that their HR departments monitored legal compliance (one of the factors in the competency of HR delivery) and large companies reported that their HR department outsourced HR programs. However, this result could not be supported by this research as no significant relationship was found between firm's size and HR delivery.

## 6. Discussion

[14] have argued that HR professionals need to become more effective strategic business partners. [25] argue further that HR professionals must transition from being strategic business partners to becoming contributors in organizations. Given that numerous studies have shown a positive relationship between strategic HRM practices and firm performance, this study sought to identify the competencies needed for HR professionals to become effective business partners and contributor in organizations.

As [26] have shown strategic contribution, personal credibility, HR delivery, business knowledge, and HR technology are all pivotal to HR being effective business partners and players. This study enabled readers to determine the various relationships existing among the core HR strategies, factors leading to improvement in the competencies, impact of education and the competencies on compensation, and the relationship between education and the respective competencies. The results provide a reminder to HR professionals of the value of graduate degrees and other means of developing higher levels of HR technical competence, understanding the benefits of accounting, marketing, and other different functional areas in effectively developing and implementing HR strategies. By understanding the impact of the competencies on the various organizational practices, there could be a more directed strategy in developing expertise among HR professionals, hence, a more credible and effective function.

The findings of this study support notions of [17] and Becker et al. [27] who emphatically stated the need for metrics as a core component of the HR strategy. Although there are many positive results for the HR profession in general from this study, there are clearly many areas that are lacking. To be a true strategic partner and player, HR professionals and other executives of organizations must focus on developing the critical competencies necessary to enhance HR effectiveness and hence maximization of shareholder value.

HR professionals need to be proactive and flexible in their mind set. They should not think that they play only a supportive role but also their contribution can give impact to an organization performance. One major finding of this study is that HR professionals often lack the com-

petencies related to business. It is clearly shown that competency such as culture management, market driven connectivity, strategic decision making, rapid adaptability, value-chain knowledge and HR technological know-how are lacking and are among the weakest abilities of a HR professionals in the Malaysian manufacturing sector.

## 7. Conclusions

The findings of this research show that HR professional in the manufacturing companies of the southern region of Malaysia are lacking in business related human resource competencies. This is one of the main barriers to be surmounted if local HR professional are to become strategic partners in their organizations. As this study has shown, these competencies contribute to a firm's effective performance. Therefore, it is vital for HR professional to possess the right competency to improve overall firm's productivity and performance. Furthermore, they must take initiative to excel in many area especially knowledge beyond HR practices. This may be the only road to success for all HR profession in the future.

## 8. Acknowledgement

The author is grateful to Dr. Richard Spear for his valuable comments on this paper. The author also would like to take this opportunity to acknowledge the contribution and generous support of the firms that have participated in this study.

## REFERENCES

- [1] W. Brockbank, D. Ulrich, and R. Beatty, "The professional development: Creating the future creators at the university of michigan business school," *Human Resource Management*, 38(2), pp. 111-118, 1999.
- [2] J. B. Barney and P. M. Wright, "On becoming a strategic partner: The role of human resources in gaining competitive advantage," *Human Resource Management*, pp. 37 (1): 31-46, 1998.
- [3] J. Selmer and R. Chiu, "Required human resources competencies in the future: A framework for developing HR Executives in Hong Kong," *Journal of World Business* (39), pp. 324-336, 2004.
- [4] N. Khatri and P. Budhwar, "A study of strategic HR issues in an Asian context," *Personnel Review* 31 (2), pp. 166-88, 2002.
- [5] K. E. Scanlan, "University HRM curriculum and HRM competencies: A comparative study." Unpublished Dissertation, Dissertation Abstracts International, 2007.
- [6] D. Ulrich, W. Brockbank, A. Yeung, and D. Lake, "Human resource competencies and empirical assessment," *Human Resources Management*, 34(4), pp. 473-496, 1995.
- [7] P. Boselie and J. Paauwe, "Human resource function competencies in European Companies," 2004.
- [8] W. Brockbank and D. Ulrich, "Competencies for the new HR: Society for human resource management," Univer-

- sity of Michigan Business School, Global Consulting Alliance, 2003.
- [9] J. Kelly and J. Gennard, "The role of personnel directors on the board of directors," *Personnel Review*, Vol. 25 No. 1, pp. 7–24, 1996.
  - [10] J. W. Heisler, "Competency today required by today HRM professionals," *Link & Learn eNewsletter*. Retrieved November 21, 2007 from World Wide Web: <http://www.hra-nca.org/savedfiles/heisler.pdf>, 2003.
  - [11] A. Mukherjee, "HR transformation—can technology help," Retrieved 23 November, 2007 from World Wide Web: [http://www.humanlinks.com/manres/articles/hr\\_transformation.htm](http://www.humanlinks.com/manres/articles/hr_transformation.htm), 2001.
  - [12] R. W. Mondy and R. M. Noe, "Human resource management," International Edition, 9th Edition, Prentice Hall, 2005.
  - [13] P. M. Wright, B. Dunford, and S. Snell, "Human resources and the resource based view of the firm," *Journal of Management*, 27, pp. 701–721, 2001.
  - [14] E. E. Lawler and A. M. Mohrman, "Creating a strategic human resource organization: An assessment of trends and new directions. Stanford," CA: Stanford University Press, 2003.
  - [15] D. Ulrich, "A new mandate for human resources," *Harvard Business Review*, Vol. 76 No. 1, pp. 124–34, 1998.
  - [16] M. A. Huselid, S. E. Jackson, and R. S. Schuler, "Technical and strategic human resource management effectiveness as determinants of firm performance," *Academy of Management Journal*, Vol. 40 No. 1, pp. 171–88, 1997.
  - [17] S. J. Ramlall, "Identifying and understanding HR competencies and their relationship to organizational practices," *Applied HRM Research*, Vol. 11 (1), pp. 27–38, 2006.
  - [18] M. Losey, "Mastering the competencies of HR management," *Human Resource Management*, 38(2), pp. 99–103, 1999.
  - [19] C. D. Johnson and J. King, "Are we properly training future HR/IR practitioners: A review of the curricula," *Human Resource Management Review* 12(2002), pp. 539–554, 2002.
  - [20] J. E. Stockman, "Building a quality HR organization at GE," *Human Resource Management*, 38, pp. 143–146, 1999.
  - [21] C. M. Giannantonio and A. E. Hurley, "Executive insights into HR practices and education," *Human Resource Management Review*, 12 (2002), pp. 491–511, 2002.
  - [22] O. R. Zhan and I. Zakaria, "Strategic HRM: A comparison between selected manufacturing and service firms," *Research and Practice in Human Resource Management*, 4(1), pp. 43–65, 1996.
  - [23] L. R. Frey, B. H. Botan, and G. L. Kreps, "Investigating communication : An introduction to research methods," Allyn & Bacon, Boston, 2000.
  - [24] F. N. Kerlinger, "Foundation of behaviour research," New York: Holt, Rinehart & Wiston, Inc., 1973.
  - [25] D. Ulrich and D. Beatty, "From partners to players: Extending the HR playing," 2001.
  - [26] W. Brockbank, D. Ulrich, and C. James, "Trends in human resource competencies," Ann Arbor: University of Michigan School of Business, 1997.
  - [27] B. Becker, M. Huselid, and D. Ulrich, "The HR scorecard: Linking people, strategy, and performance," Boston: Harvard Business School Press, 2001.



# Analysis of University Science Research Capability Elements and Evaluation based on BP Neural Network

Weiwei Liu & Chunsheng Shi

School of Management, Harbin Institute of Technology

Email: challenge10760@163.com

Received October 28<sup>th</sup>, 2008; received December 5<sup>th</sup>, 2008; accepted December 14<sup>th</sup>, 2008.

## ABSTRACT

*After analyzing effect elements of University Science Research Capability (USRC) based on dynamic capability theory, combined the substance of university science research with the highly self-organized, self-adapted and self-learned characteristics of Back Propagation (BP) Neural Network, the paper conducts a research on evaluation of USRC, in which an evaluation index system of USRC is constructed and a 15-7-1-typed BP Neural Network with three layers is presented to evaluate USRC, which provides a BP Neural Network-based methodology for evaluation of USRC with multiple inputs.*

**Keywords:** BP neural network, science research, capability evaluation

## 1. Introduction

The capability of a university decides its competitive advantage and management performance in the essence [1]. Therefore, it does make sense for universities to accumulate, develop, evaluate and utilize their capabilities.

Nowadays, science research capacity is the core of universities and an important indication of any powerful university, all the universities tend to pay attention to cultivate and enhance their capabilities, especially science research capability. It's necessary to analyze and evaluate science research capability elements of universities. A number of researches have already been conducted on evaluating competitive capabilities [2,3,4,5,6]. However, few researches have been taken on evaluating scientific research capacity in universities. To fill in this gap, the paper, first, investigates into the elements of university science research capability based on dynamic capability, and then, constructs BP Neural Network model of evaluation, in the end, introduces into a simulation evaluation, aiming at providing both theoretical and empirical perspectives in the cultivation of university science research capability.

## 2. Analysis of USRC Based on the Theory of Dynamic Capability

### 2.1. The Theory of Dynamic Capability

The theory of dynamic capability is proposed firstly by Teece, Pisano and Shuen in "Firm Capability, Resource and Strategic concept" [7]. The theory of dynamic capability develops and consummate gradually in "Dynamic Capabilities and Strategic Management" by Teece, Pis-

ano and Shuen [8]. They defined dynamic capability as the capability of conforming, constructing and reconfiguring inner and exterior capabilities to adapt environment changed rapidly. The definition has two outline: "dynamic" namely, enterprises must renovate capabilities of themselves to adapt changeable environment; "capability" namely, strategic management has key function in renovating capabilities of themselves to adapt changeable environment. Some scholars consider that capability can be defined as the gather of enterprise knowledge and capability that can change capability is technology knowledge [9].

### 2.2. Analysis of USRC Based on Characteristics of Dynamic Capability

The theory of dynamic capability comes from the theory of resource base and absorbs many viewpoints of the theory of core capability, thus, its characteristic is similar with the theory of core capability, for example, the nature of value, the nature of unique. Nevertheless, dynamic capability is the capability that changes capabilities; its difference in nature from core capability is that it's the nature of deployment [10]. Since science research activities of universities can't depart from resource base, the products of universities science research activities has the value of applying and generalizing, science research characteristics of each universities are different from others and different from the advantage of science research competition rest with the nature of unique. Universities should renovate constantly their science research capa-

bilities for adapting the changes of environment and keeping the advantage of competition.

The theory of dynamic capability is production of conforming inside knowledge and absorbing outside knowledge [11]. Absorbing knowledge has bridge function in resource and capabilities between inside and outside, thus, the theory of dynamic capability emphasizes to construct special capabilities of absorbing knowledge from outside. Universities must enhance learning intercourse between internal and overseas to heighten science research capabilities. Universities can absorb advanced knowledge from outside and master international learning development trends. Universities should increase the opportunity of studying abroad and visiting, besides, should introduce into visiting professors. The communion of experiences is good for importing advanced techniques and methods.

### 2.3. Analysis of University Science Research Capability Effect Elements

Competitive advantages of universities come from science research capability. The elements of university science research capability can be elaborated from four aspects: science research input, transformation efficiency and science research output and science research management. Combined with the features of university development, science research capability can be classified into four aspects, i.e., science research input capability, transformation efficiency capability and science research output capability and science research management, science research input capability effect university science research capability, science research output capability reflect university science research capability and transformation efficiency capability effect university science research capability indirectly, a sound management mechanism makes effect elements exert supreme efficiency and makes the whole benefit exceed the summation of parts. Science research management is holistic macroscopically elements of effecting USRC, as shown in Figure 1.

Science research input capability and science research output capability affect each other in science research process. Science research input capability provides substance base and intellect sustainment for science research output capability. Science research output capability act on science research input capability and provides reliable basis for science research input capability. Transformation efficiency capability is used for transforming between science research input and science research output effectively.

### 3. Comparison and Selection of Evaluation Methods

As to the evaluation of USRC, such methods as Analytic Hierarchy Process (AHP), Fuzzy Comprehensive Evaluation (FCE) and Data Envelopment Analysis (DEA) have been adopted with certain effects [12,13,14].

#### 3.1. Analytic Hierarchy Process (AHP)

AHP is proposed by American T. L. Satty in 1970s<sup>[15]</sup>, its essential principle is that setting down evaluation scheme on the basis of objects with gradual rank, sub-objects, restriction condition and adopting the method of comparing between one and the other to make sure judgment matrix, and then, making proportion vectors of characteristic vector that maximum characteristic of judgment matrix corresponding to as coefficient, in the end, presenting the power weight of each schemes synthetically. The method is an improvement of the method of adding power average, however, when evaluation scale is very large and there are too many evaluation indexes, the method prefer appearing problems, for example, judgment matrix is difficult with meeting the requirement of conformity. USRC there are too many evaluation indexes, therefore, it can result with the difficulty in meeting the requirement of conformity by AHP.

#### 3.2. Fuzzy Comprehensive Evaluation (FCE)

Fuzzy Comprehensive Evaluation (FCE) is proposed by Peizhuang Wang [16]. It has two steps: individual evaluation according to each factor; integrative evaluation

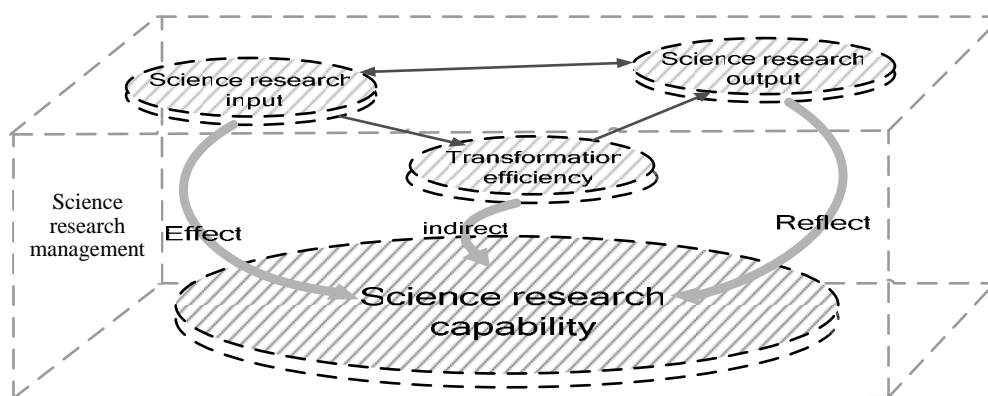


Figure 1. University scientific research capability elements system

according to all the factors. Its primary function is that quantificational evaluation to qualitative indexes. However, it has two shortages: firstly, it's difficult to compartmentalize the criterion of evaluation, for example: the limit between "very good" and "good" is difficult to ascertain, there are no specific criterions; secondly, the quality of expert mark can't be ensured. Because of intense subjectivity, FCE can't be used for evaluating USRC.

### 3.3. Data Envelopment Analysis (DEA)

DEA is a systems analysis evaluation method established on the basis of relative efficiency evaluation concept and given by A. Charnes and W. W. Cooper [17]. It supposes there are  $n$  decision making units (DMU), every DMU has  $m$  kinds of input and  $s$  kinds of output. When evaluating USRC, determinate selection rule should be transformed into data envelopment analysis model. Firstly, compartmentalize selection rule as input variable and output variable, and then, establish data envelopment analysis model and calculate relative efficiency of every university, in the end, the place of USRC can be obtained on the basis of calculation results. However, the method has relative validity because of being constructed on the basis of comparing each other. The evaluation of USRC isn't only used for getting horizontal comparison of many universities, but also getting vertical dynamic development trends. Because of disadvantages of DEA in relativity, the evaluation of USRC can't adopt DEA.

### 3.4. BP Neural Network

Among these methods, DEA is limited to the evaluation of ascertained performance indexes, AHP has restrict of evaluation scales, while FCE shows subjectivity in such aspects as data collection and index weight determination. Meantime, all these methods are based on the assumption of linear relations between indexes and can only be adopted within limits.

Characterized as self-adaptability, self-learning and large-scale parallel calculation ability, BP Neural Network is the most commonly used network model [18] in such aspects as identification, classification, evaluation, forecast, nonlinear mapping and simulating complex system [19]. BP Neural Network is a multiple-layer feedback network featured in back propagation of errors and constituted of input, hidden and output layers.

BP Neural Network is the simplification, abstraction and simulation of brain function. And it is a highly sophisticated and nonlinear dynamic system, in which nonlinear mapping is adopted to find out inner relations according to the existing training-set and through learning and training. Without the process of subjective calculation of index weight, the evaluation model proposed in the paper decreases the subjectivity and overcomes the possible linear problem between indexes effectively, which enhances the reliability and adaptability of evaluation for USRC.

## 4. Establishment of USRC Evaluation Model Based on BP Neural Network

### 4.1. Establishment of the Index System of USRC Evaluation

Competitive advantages of universities come from science research capability. A sound evaluation index system is the basis of effective evaluation of USRC. According to the establishment principles of index system [20] and based on many interrelated reference literatures [21,22,23,24], this paper selects science research input capability, transformation efficiency capability and science research input capability as the First-order indexes and further proposes the index system to evaluate synthetically USRC, as shown in Table 1.

Table 1. Evaluation index system of USRC

Object Layer	First-order indexes	Second -order indexes
USRC	Science research input capability	Science research expense ( $U_1$ )
		Science research procession ( $U_2$ )
		Subject and science research base ( $U_3$ )
		Science research equipment and condition ( $U_4$ )
		Personnel opening degree ( $U_5$ )
	Transformation efficiency capability	Learning intercommunion ( $U_6$ )
		The amount of project ( $U_7$ )
		Cooperation capability( $U_8$ )
		Science research production ( $U_9$ )
		Production application ( $U_{10}$ )
	Science research output capability	Student cultivation ( $U_{11}$ )
		Project finishing status ( $U_{12}$ )
		Science research input management capability( $U_{13}$ )
	Science research management capability	Transformation efficiency management capability( $U_{14}$ )
		Science research output management capability( $U_{15}$ )

## 4.2. Pre-treatment of evaluation index data

Evaluation index system consists of quantitative and qualitative indexes. Quantitative indexes can be further divided into positive, reverse and moderate indexes [25]. Considering the difficulties in comprehensive evaluation directly based on the data and the convergence problem of BP Neural Network, normalized and trend-assimilated pre-treatment is needed first. For qualitative indexes, we can make sure them by the method evaluation grade subjection degree, for example, 1, 0.75, 0.5, 0.25, 0 correspond respectively very good, good, common, bad, and very bad. For quantitative indexes, there are three steps:

- (1) Calculate the average value  $\bar{P}_j$  of the  $j$ th test index  $P_j$ ,

$$\bar{P}_j = \frac{\sum_{i=1}^n x_{ij}}{n}, j=1,2,k,m \quad (1)$$

- (2) Calculate middle variable of index  $M_{ij}$ :

If bigger target values indicate better effects, i.e. benefit index,

$$M_{ij} = \frac{x_{ij} - \bar{P}_j}{|P_j|} \quad (2)$$

If smaller target values indicate better effects, i.e. cost index,

$$M_{ij} = \frac{P_j - x_{ij}}{|P_j|} \quad (3)$$

If fixed scope values indicate better effects, i.e. scope index,

when

$$x_{ij} \leq A, M_{ij} = \frac{x_{ij} - A}{|A|} \quad (4)$$

when

$$x_{ij} \geq B, M_{ij} = \frac{B - x_{ij}}{|B|} \quad (5)$$

when

$$B \geq x_{ij} \geq A, M_{ij} = 6 \quad (6)$$

In Function (4), (5) and (6), A and B are best top boundary and best bottom boundary of scope indexes,  $M_{ij}$  reflects the degree  $x_{ij}$  deflects to average.

Calculate subjection degree function value  $Y_{ij}$  of indexes transforming to  $[-1, 1]$

$$Y = \frac{1 - e^{-M_{ij}}}{1 + e^{-M_{ij}}} \quad (7)$$

## 4.3. Design of Physical Structure of USRC Evaluation Model

The selection of network model structure is very important. Good selection can reduce the network training times and raise the network learning precision. The selection procedure includes:

- (1) Ascertain the layer number of BP Neural Network. Theorem A. H. Kolmogrov has proven that a nonlinear neural network with three layers and S-typed action function will approach any continuous function with any precision as long as it possesses enough hidden nodes [26]. As a result, the paper adopts three-layer BP Neural Network in constructing the evaluation model of USRC, as shown in Figure 2. As to the function between nodes, S (Sigmoid)-typed function is introduced, as shown in Function (8).

$$f(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

- (2) Ascertain the numbers of neural cell nodes both in the input layer and output layer of BP Neural Network. The numbers are dependent on outer description. The node number in the input layer is the number of characteristic factors (independent variable) while that in the output layer is the number of targets in the system. As a result, we can conclude that the input-layer node number is 15 and the output-layer is one, according to the index system of USRC evaluation as shown in Table 1.

- (3) Ascertain the node number in the hidden layer of BP Neural Network. The node number in the hidden layer has a direct relation with the requirements and the node numbers both in the input and output layers. Too few nodes will result in the failure in network training or in the identification of unseen set before, while, too many ones will result in an excessively long learning period and an unobvious reduction in errors. Therefore, an optimal node number must be existent. Up to the present, no sound method has been found to ascertain the number. On the basis of previous researches on BP Neural

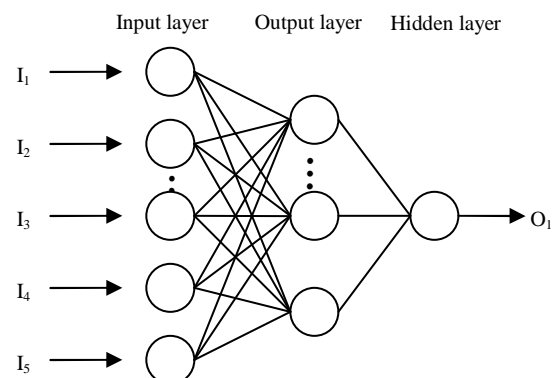


Figure 2. BP Neural network model structure

Network, two empirical formulas are concluded as follows:

$$\lambda = \sqrt{0.12m^2 + 0.43mn + 0.77m + 2.54n + 0.35} + 0.51 \quad (9)$$

$$\lambda = \sqrt{m + n} + A \quad (10)$$

In Function (9) and (10),  $m$  indicates the node number in the input layer, while  $n$  that in the output layer.  $A$  indicates any integer between one to ten.  $\lambda$  is the ascertained node number in the hidden layer.

According to Function (9) and (10) and the effectiveness of models, the node number in the hidden layer is assumed seven.

#### 4.4. The Selection of Logic Algorithm in the Evaluation Model of USRC

The BP algorithm put forward by Rumelhart provides neural network with a more practical and effective training method. However, the algorithm has limitations when applied in complex systems for the high nonlinearity results in the low efficiency and slow convergence of the initial BP algorithm. In order to overcome these obstacles, the paper adopts a refined algorithm, a combination of additional momentum and Adaptive Learning Rate Method.

According to the BP Algorithm and the Minimum Mean-square Error, we can conclude the connection weight adjusting function of additional momentum factors:

$$\Delta w_{ij}(k+1) = (1 - \alpha)\eta\delta_i x_j + \alpha\Delta w_{ij}(k) \quad (11)$$

$$\Delta b_i(k+1) = (1 - \alpha)\eta\delta_i + \alpha\Delta b_i(k) \quad (12)$$

In Function (11) and (12),  $k$  indicates the training number.  $\Delta w_{ij}$  and  $\Delta b_i$  indicate the increments of weight.  $\eta$  indicates the learning rate.  $\delta_i$  is error.  $x_j$  is network input.  $\alpha$  is momentum factor and the average value is 0.9.

The value of  $\eta$  decides the success of algorithm to some extent. Large value results in the oscillation of error function and small value the slowness of convergence. In order to solve this, the method of adapting learning rate is adopted in the network training, i. e. Adaptive Learning Rate Method and the adapted function is concluded as follows [27]:

$$\eta(k+1) = \begin{cases} 1.05\eta(k), & \text{erf}(k+1) < \text{erf}(k) \\ 0.7\eta(k), & \text{erf}(k+1) > 1.04\text{erf}(k) \\ \eta(k), & \text{else} \end{cases} \quad (13)$$

$$\text{erf} = \sum_{i=1}^N (T_i - O_i)^2 \quad (14)$$

In Function (13) and (14),  $\text{erf}$  is the error function.  $T_i$  is the expected output value of learning set.  $O_i$  is the real output value of learning set.  $N$  is the number of learning sets.

## 5. Empirical Research of Model

According to the evaluation index system above, the paper selects 12 groups of samples and divides into two parts. The former 8 groups are adopted in network training as learning sets and the latter 4 groups in network test. Besides, the paper selects 15 second-order inputting indexes and adopts the module of artificial neural network in MATLAB software. After the learning process of the evaluation of USRC, the BP Neural Network-based evaluation model of USRC is concluded. This process consists of the following steps:

(1) Select 12 groups of evaluation indexes as learning sets. The input values are the attribute values of 15 indexes after pre-treatment and the corresponding output value expected is the evaluated value of each sample.

(2) Start up the learning process after the normalization of index attributes and iterate the process till convergence. In this process, the network toolbox of MATLAB7.0 is used, with a learning rate of 0.01, a momentum factor of 0.9, a maximum training step of 2000 and an error precision of 0.0001. During the training, some obviously abnormal data have been eliminated through the displaying function of variance chart in MATLAB and the learning results have been achieved, as shown in Table 2.

(3) Restore the trained parameters of the BP Neural Network with three layers into the corresponding knowledge database and input respectively 4 groups of validating data to further examine the network model. The comprehensive examination result is comparatively ideal, which validates the effectiveness of the model, as shown in Table 3.

The simulation evaluation result with test samples is the same as experts' evaluation result basically. The application example indicates that the network has comprehensive practicability. The model can master expert knowledge by sample learning and be used evaluation with many indexes.

**Table 2. The results of learning**

Sample code	Training value	Expected value	Relative error (%)
1	0.556	0.549	1.28
2	0.117	0.119	-1.68
3	0.547	0.551	-0.73
4	0.423	0.421	0.48
5	0.113	0.116	-2.59
6	0.368	0.365	0.82
7	0.649	0.651	-0.31
8	0.235	0.227	3.52

**Table 3. The examination results**

Sample code	Training value	Expected value	Relative error (%)	Simulation place	Expert place
9	0.225	0.227	-0.88	4	4
10	0.345	0.343	0.58	3	3
11	0.435	0.433	-0.46	2	2
12	0.556	0.555	0.18	1	1

## 6. Conclusions

This paper constructs an evaluation index system of USRC combined with characteristics of universities based on dynamic capability theory, after analyzing effect elements of USRC, the paper presents a 15-7-1-typed BP Neural Network with three layers to evaluate USRC on the basis of the highly self-organized, self-adapted and self-learned BP Neural Network comprehensive evaluation method. The method constructs a comprehensive evaluation model combined with both quantitative and qualitative indexes which is close to human being thought mode better. The satisfying result is obtained by emulational test. Its advantages embodies in the following aspects: avoiding the effects of subjectivity and randomness in traditional evaluating methods and ensure the preciseness and objectivity of results; according with the empirical situation along with the increasing number of training samples; and overcoming the possible linear problems among the indexes and enhancing the reliability and adaptability of evaluation. Therefore, compared with the traditional evaluating methods of USRC, the one based on BP Neural Network is of better practicability. BP Neural Network can learn by random sample parameter and construct diverse evaluation model. It can get reliable evaluation result on the basis of practical test sample after learning successfully; meanwhile analysis result will be accurate and factual when training samples increase gradually, therefore, the method has more comprehensive applicability.

## REFERENCES

- [1] V. M. Giménez and J. L. Martínez, "Cost efficiency in the university: A departmental evaluation model," [J]. *Economics of Education Review*, Vol. 25, pp. 543-553, 2006.
- [2] P. Boccardelli and M. G. Magnusson, "Dynamic capabilities in early-phase entrepreneurship," [J]. *Knowledge and Process Management*, Vol. 13, pp. 162-174, 2006.
- [3] D. J. Teece, "Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance," [J]. *Strategic Management Journal*, Vol. 28, pp. 1319-1350, 2007.
- [4] C. E. Helfat and M. A. Peteraf, "The dynamic resource-based view: Capability lifecycles," [J]. *Strategic Management Journal*, Vol. 24, pp. 997-1010, 2003.
- [5] B. McEvily and A. Marcus, "Embedded ties and the acquisition of competitive capabilities," [J]. *Strategic Management Journal*, Vol. 26, pp. 1033-1055, 2005.
- [6] D. Hussey, "Company analysis: Determining strategic capability," [J]. *Strategic Change*, Vol. 11, pp. 43-52, 2002.
- [7] G. S. Wang, "Enterprise theory: capability theory [M]," China Economic Publishing House, PP. 69-70, 2006.
- [8] Teece, Pisano, and Shuen, "Dynamic capabilities and strategic management"[J], *Strategic Management Journal*, 18(7): pp. 509-533, 1997.
- [9] J. W. Dong, J. Z. Huang, and Z. H. Chen, "An intellectual model of the evolution of dynamic capabilities; and a Relevant Case Study of a Chinese Enterprise" [J], *Management World*, (4): pp. 117-127, 2004.
- [10] J. Z. Huang and L. W. Tan, "From capability to dynamic capability: transformation of the points of stratagem" [J], *Economy Management*, (22): pp. 13-17, 2002.
- [11] K. M. Eisenhardt and Martin, "Dynamic capabilities: what are they," [J], *Journal of Evolutionary Economics*, (3): 519-523, 2000.
- [12] T. M. Crimmins, J. E. de Steiguer, and Donald Dennis, "AHP as a means for improving public participation: A pre-post experiment with university students," [J]. *Forest Policy and Economics*, Vol. 7, pp. 501-514, May 2005.
- [13] A. Golec and E. Kahya, "A fuzzy model for competency-based employee evaluation and selection," [J]. *Computers & Industrial Engineering*, Vol. 52, pp. 143-161, February 2007.
- [14] J. C. Glassa, G. McCallionb, and D. G. McKillopc, "Implications of variant efficiency measures for policy evaluations in UK higher education," [J]. *Socio-Economic Planning Sciences*, Vol. 40, Issue 2, pp. 119-142, June 2006.
- [15] L. Bodin and S. I. Gass, "On teaching the analytic hierarchy process," [J]. *Computers & Operations research*, pp. 1487-1497, 2003.
- [16] D. Du, Q. H. Pang, and Y. Wu, "Modern synthesis evaluation methods and cases choiceness," [M], TSINGHUA University Press, pp. 34-40, 2008.
- [17] H. Hao and J. F. Zhong, "System analysis and evaluation methods," [M], Economic Science Press, pp. 185-239, 2007.
- [18] S. L. Sun, "The application study of supplier evaluation model in supply chain," [J], *Modernization of Management*, Vol. 1, pp. 62-64, 2007.
- [19] F. S. Wang and S. Z. Peng, "Study on the measurement of the activity-based value output of industry value chain based on BP neural network," [J], *Journal of Harbin Institute of Technology*, Vol. 1, pp. 91-93, 2006.
- [20] Y. J. Guo, "Theory, method and application of comprehensive evaluation," [M]. Bei Jing: Science Press, pp. 14-29, 2007.
- [21] W. Z. Zhu, "The establishment of systematized evaluation indexes to university science research capacity," [J]. *Journal of Anhui University of Technology and Science*. Vol. 18, pp. 40-44, 2003.
- [22] J. Q. Cai and L. M. Cheng, "Research on evaluation of university science research capability" [J]. *R & D Management*, Vol. 10, pp. 76-80, 1998.
- [23] X. Min, C. R. Dai, and H. Bin, "Research on evaluation of university science research capability based on theory of fuzzy mathematics," [J]. *Science and Technology Management Research*. Vol. 8, pp. 185-187, 2006.
- [24] B. Liu and X. L. Wang, "Research on evaluation method of university science research capability," [J] *Science of Science and Management of S. & T.* Vol. 12, pp. 83-85, 2003.
- [25] Y. H. Li and Y. Q. Hu, "Comprehensive evaluation based on bp artificial neural network on enterprises core competence in the old industrial base," [J], *Commercial Research*, Vol. 5, pp. 34-36, 2006.
- [26] L. C. Jiao, "Systemic theory of neural network," [M]. Xi An: Xidian University Press, 1996.
- [27] Y. F. Wu and Z. Z. Li, "Application of improved BP nerve network model in forecasting dam safety monitoring values," [J], *Design of Hydroelectric Power Station*, Vol. 2, pp. 21-24, 2002.