ISSN: 1940-9893 (Print), 1940-9907(Online) Vol.1, No.2, September 2008

Journal of Service Science and Management

Chief Editor : Samuel Mendlinger

www.SRPublishing.org/journal/jssm

Scientific Research Publishing

J. Serv. Sci. & Management, **2008, 1**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

CONTENTS

Vol. 1 No. 2 **September 2008**

Journal of Service Science and Management (JSSM)

Journal Information

SUBSCRIPTIONS

The ***Journal of Service Science and Management*** (Online at Scientific Research Publishing, www.SRPublishing.org) is published quarterly by Scientific Research Publishing, Inc. 3306 Apple Grove CT, Herndon, VA 20171, USA.

E-mail: jssm@srpublishing.org

Subscription rates:
Print: $50 per copy. Volume 1, Number 2, 2008
Electronic: free, available on www.SRPublishing.org.
To subscribe, please contact Journals Subscriptions Department, E-mail: jssm@srpublishing.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@srpublishing.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator, Scientific Research Publishing, Inc. 3306 Apple Grove CT, Herndon, VA 20171, USA.
E-mail: jssm@srpublishing.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: jssm@srpublishing.org

J. Serv. Sci. & Management, **2008, 1: 111-122**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

Strategic Capabilities, Innovation Intensity, and Performance of Service Firms

Lu-Jui Chen[1], Chun-Chung Chen[2] & Wen-Ruey Lee[3]

[1]*MingChuan University*
[2]*National Taiwan University*
[3]*National Taipei College of Business*
Email: shunyde@mcu.edu.tw, osimchen@ntu.edu.tw, wrlee@webmail.ntcb.edu.tw

ABSTRACT

This study developed and empirically tested a model examining the relationships among strategic capabilities, innovation intensity, and firm performance. Strategic capabilities include internal venturing capability and social relationship capability. Analyzing a sample of service firms from Taiwan, the study indicates that social relationships with other firms are important to facilitate innovative activities of service firms. Innovation intensity further helps service firms to improve a firm's expected performance. However, internal resources capability does not show the expected effect on innovation intensity. And innovation intensity is also not related to a firm's growth.

Keywords: *resource-based view, social relationship, innovation intensity, firm performance, service firms*

1. Introduction

This study explores the relationships between strategic capabilities, firm's innovation and the performance of the service firms. Innovations of firm create new jobs, generate new wealth for firm. However, we do not know much about value of the extent to firm's strategic capabilities on innovation, since previous studies have explored innovation without exposing its strategic capabilities. In this paper, we try to reveal the role of strategic capabilities—especially strategic in internal capabilities and in external networks—in the innovation creation process. And we also deal with the performance implication of strategy of with innovation. Two guiding theories—resource-based view and social network—were invoked to account for the value of innovation and performance.

Resource-based view (RBV) emphasizes firm with idiosyncratic resources [9] that are owned or controlled by the firm [49]. These capabilities of deploying resources made firms heterogeneous in nature. RBV regards the firm as a bundle of resources and suggests that their characteristics make a positive effect to the potential of innovation, and by implication of its performance. There are an increasing number of studies focusing on the competitive factors of firms. The studies show that intangible and tangible resources [24] and human resource management [30], among others, are elements that clearly contribute to firm's internal capabilities.

Social networks advance that external connections are the sources of competitive advantages [6]. External networks with suppliers, customers and others would facilitate the product/service mobilized and concrete. Firm transacts with outside entities in order to acquire external resources and opportunities, adjusted for the firm's potential value. Social network theory implies that its relational characteristics are embedded with creative opportunities and potentials for the process of value in firm.

Services have been increasingly providing intelligent inputs, adding products with a wide range of value and using other technological processes in this new economic time [68]. Firms create value by offering the types of services that customers need, at an acceptable price. In return, firms receive value from their internal property and their stakeholders. Unlike manufacturing firms, which rely on patented technologies or unique products, service firms gain their competitive advantage primarily through their ability of combination to make use of their proprietary knowledge. The activity of service firms is an "interaction between human and human" (or organization and organization). It is contrary to Daniel Bell's characterization, which considers pre-industrial society as a "game against nature," and industrial society as a "game against fabricated nature" [10]. Thus, the service industry needs not only technical skill, but also social skill.

Services can be the initial element in producing an innovative packaging material through R & D, or the "mediating" element in developing a major mining project. In other instances, services firms add product value by providing convenience, health and knowledge. For decades, services make up the bulk of nowadays' economy and also account for most of the growth. In fact, services now

dominate, making up about 70% of the aggregate production and employment in the Organization for Economic Cooperation and Development (OECD) nations and contributing about 75% of the GDP in the United States. Most western countries have also increased service products, and the export of these products is between 10% and 20%. The service output in Taiwan took over 70% of its GDP in 2005. This means that Taiwan is moving in a direction characterized by service-oriented output. There is considerable potential for increase in value of services.

It is well recognized that the service business plays an important role in a nation's economic development. The determinants of business components for service firms should be different from manufacturing firms. Most research in the performance literature, for example, has focused on the determinants of performance in manufacturing industries [60]. Yet the services firm literature has long argued that the nature of goods and services are not the same and that services face a unique set of challenges [18, 19]. Accordingly, it is reasonable to expect that characteristics in service settings are likely to be different and specific. Our prediction is the extent to which those determinants of strategic capabilities can be generalized to service firms.

This study pushes the envelope of research on the strategic capabilities by service firms in Taiwan by integrating RBV and social network theory. This study attempts to contribute to the knowledge and research in the service firms' management in the following ways. First, most research studies have focused exclusively on manufacturing firms but not on service firms [2, 35]. While services have become the fastest-growing component of international trade, it is important to know the extent to which competitive perspectives, theories and practices, developed for manufacturing firms, are also applicable to service firms. Second, few studies have examined the behavior of service firms by using competing theoretical perspectives [17, 19]. Although server studies attempted to integrate these two theoretical domains to explain organization development, we believe that none have done so in the context of service firms. Third, we design this study to test the theoretical relationship by combining subjective and objective data. The results inform our knowledge of the management of the service industry with a complementary view.

This study is organized as follows. In the next section, we present a brief review of relevant theoretic literature. The conceptual framework and the research hypotheses will be provided. The following section describes the methodology, samples, variables, and hypotheses testing. Finally, we will present concluding remarks and managerial implications.

2. Theory and Hypotheses

The strategic capabilities is indicated by the degree to which they can contribute the development of core competences, competitive advantage, and, ultimately, firm performance. Hence, strategic capabilities defined here are the firm's capacity to deploy internal resources and integrate external resources that have been coordinated purposely to achieve a firm's creation and a desired end state [9].

Innovation is the process of creating a commercial product and service from an invention. The innovation can be created through internal entrepreneurial mind-set [59, 71] and cooperative strategies [6]. Hence, we argue that two types of strategic capabilities will influence the process of innovative activity of a firm, and that the process is a key mediator that affects the firm's performance. We draw on theoretical perspectives from several sources: internal resource advantage from the resource-based view (RBV) of a firm [9, 49, 72], external advantage of social relationship from the social capital perspective [6, 11, 22], innovative intensity [33, 59, 71], and exploration and exploitation from organizational learning [37, 41].

The following section introduces the above theories sequentially and also introduces the research hypotheses simultaneously.

Internal venturing capability and innovation intensity

Most innovation is developed through internal R&D, exercised by corporate staffs. Thus, the most competitively successful firms reinvent their industry or develop a completely new one across time as they engage in competition with current and future rivals. In this sense, strategic entrepreneurship is about producing the innovation and encouraging innovative intensity activities that create tomorrow's business [59, 71]. Internal venturing is the set of resources and activities firms use to develop innovations.

The resource-based view [9, 49] complementing the traditional model of Porter's [52] competitive advantage, stressed the importance of the internal resources and capabilities of a firm in the context of the competitive environment. The RBV suggests that researchers devote their attention to analyzing the performance of firms in terms of their resources, rather than their product market activities [72], since distinctive organizational resources, capabilities, and competencies generate a sustainable competitive advantage and lead a firm to above-normal performance [53].

Gaining superiority in a competitive market depends on a firm's ability to identify, develop, deploy, and preserve particular resources that distinguish it from its rivals [4] [14]. Resources and capabilities contribute to improving the firm's competitive position and thus have the potential to create competitive advantage [8]. In order to produce a sustainable competitive advantage, resources and capabilities should be characterized as highly valuable, rare, inimitable (i.e., they are costly to copy by rivals), nonsubstitutable (i.e., no substitute to fulfill the same function is immediately available to competitors) [8] and nontransferable (i.e., resource cannot easily be purchased) [14].

More recent studies emphasize the important of knowledge-based resources [36, 63], which is characterized by firm employment [49]. Therefore, competitive advantage resides in the resources available to the firm [8, 63]. Recent extensions of RBV suggest that sustainable competitive advantages are not achieved through the strategic utilization of any one kind of resource, but rather through the bundling and revitalizing of multiple, distinctive firm resources and competencies in order to create valued outputs capable of becoming sustainable competitive advantages [63].

The essence of human capital is the sheer intelligence of the organizational member [9]. According to RBV, firm employment enhances the potential of internal advantage, which is most difficult to imitate and can provide a firm with sustained competitive advantage [33]. The greater the employment potential, the more they can give a specific advantage through cost savings from increased utilization, combination of resources, lower turnover and higher productivity from boosting the productivity of individual workers. According to Jackson *et al.* [33] and Ulrich and Lake [67], employment can be further analyzed into the following three dimensions: capability and potential, motivation and commitment, and innovation and learning. Capability and potential includes concepts such as educational level, professional skills, experience, attitudes, personal networks, values, and the ability of current employees to evolve within the organization. Finally, innovation shows the degree to which employments are open to create. Innovation is increased by the quality of the human capital and an enhancement of the labor productivity [70].

Hence, this study proposes that ability of employment in the service firms can represent an internal resources advantage. For example, a service firm's market orientation and strategic decision-making are bundled together with internal complementary resources, such as innovation [42, 46]. And internal resources usually mean human resources, financial property, and management know-how, etc. However, the resource-based view of the firm [49, 72] stresses the resources that are bundled by the firm, which is understood as an organization characterized by administrative routines. It is the services based on the firm's resources, rather than the resources *per se*, that constitute the firm's knowledge. Hence, the knowledge base of firms is intrinsically linked to the knowledge of their employees [38] and those that highlight the greater share of service activities and the tendency of high-skills services [50]. The production of services is almost entirely dependent on the ability of the firm to make use of the knowledge of the employees in the case of services.

Innovative capabilities can be considered as a subset of dynamic organizational capabilities. The company survived in difficult times and improved its market positioning, establishing a reputation for innovation. Personnel competencies are improved in a number of ways (e.g. multi-skilled development) that are evident when a firm's employees are engaged with customer or supplier peers. The sales per employee capture efficiency and effectiveness improvements for the firm [58, 74], which are often the central goal of restructuring a firm's process and product.

Organizational innovation is viewed as the functional systems and processes organizations utilize to upgrade a firm's existing products, services, and processes, along with the creation and introduction of new products, services, and processes [66]. Innovation represents the commercialization of new technologies or technological change [71]. Hence this study infers innovation as "a complex activity which proceeds from the conceptualization of a new idea to a solution of the problem and then to the actual utilization of economic or social value." As March [41] suggests, exploration and exploitation are essential for organization, but they compete for scarce resources. Thus, a firm's capability to allocate scarce resources that can maximize the returns from either exploration or exploitation comprizes its intangible competencies. According to Penrose [49], Cohen and Levinthal [13], and Teece *et al*. [63], a firm innovates through learning processes that enable the firm to re-bundle and revitalize existing and newly acquired resources into core competencies and competitive advantages, and by applying internally and externally created knowledge and technology to develop new products, services, and processes.

Most improvements to service activities are incremental. In effect, a firm's innovative view may create a new market. For example, FedEx Corp. redefined the package delivery market. The internal resources of firms bring about competitive advantage, innovations and efficiency [40]. Because innovative activity is characterized by the continuous improvement of products and productivity, the sudden and unpredictable changes in the threats and opportunities that a firm faces are called Schumpeterian revolutions. Schumpeterian revolutions have the effect of drastically changing the value of a firm's resources by changing the threats and opportunities that face a firm. The RBV provides a unified approach in the conceptualization of the foundation of innovation. Several researchers have extended the RBV concepts linking to innovation [14, 51, 63]. We suggest that a firm's level of overall innovation is manifested in its capability to explore new possibilities. Likewise, a firm's level of product or service quality is manifested in its capability to exploit currently established certainties. Hence,

Hypothesis 1: A firm's internal venturing capability has a positive relationship with the firm's innovation intensity.

Social relationship capability and innovation intensity

Social capital could be understood roughly as the goodwill that is engendered by the fabric of social relations and that can be mobilized to facilitate action [1]. The core of social capital is the idea that goodwill drawn from family, friends, workmates and acquaintances provides a

range of valuable resources, including information, influence and solidarity [1, 23]. Recent research has applied social relationships as external advantage to a broader range of social phenomena, including relationships within and beyond the firm [11]. They control business benefits with outside entities [5, 6].

A firm's network will consist of relationships as well as the firm's position within the whole network of relationships. There are two types of relationships for the business network. One is referred to the closeness of a firm's set of direct and dyadic relationships, which has been labeled relational embeddedness. The other is the aspect of centrality of the firm in multiple-level relationships, which has been called structural embeddedness [22, 25]. For the clarity of analysis, the study focuses on dyadic relationships.

Firms create competitive advantage and economic value through effective interfirm collaboration [16]. Social relationships build on the general idea that economic actions are influenced by the social context in which they are embedded, and that actions can be influenced by others in social interaction [25]. Relational capital, which is so important at the dyadic level, rests upon close ties at the dyadic level and can also play an important role in creating business value and learning [34]. Tsai and Ghoshal [65] identify the social capital as the essential antecedent to facilitate the activity of value creation of firms.

Absorptive capability is a critical feature that makes firms learn and assimilate outside knowledge [13]. Considering collaboration as a learning opportunity [26], a firm may initiate collaboration relationships and create new know-how [32]. Lane and Lubatkin [37] have suggested that inter-organization relationships facilitate the difficulties of assimilation of knowledge. The breadth and the depth of the relationships between firms are associated with mutual adaptation of activities and the trust existing in the relationships. Mutual trust eliminates transaction costs and also increases the opportunities to create new opportunities [16]. Firms are embedded in socio-economic networks rather than being isolated islands in the market [22] — no matter whether they are engaged in innovative activity or not. Thus collective learning, or cooperative learning, is the situation in which partners learn to work together [15], which often contributes to the increase in the stock of knowledge.

Innovation is equally important for large and small firms in the contemporary competitive and changing market. No firms—even the largest firms such as multinationals—can always undertake major innovations alone and overcome any resource barriers for innovative activities. Hence, there is an increasing trend in strategic collaborations [21, 25] and this trend is seen as an external advantage to the firm. Close contact and intense interaction between individual firms act as an effective mechanism to transfer or learn "sticky" and beneficial knowledge-how across the organizational interface. The combination of knowledge and the creation of innovation are complex social processes; much of the value of innovative concepts is fundamentally socially embedded [45].

Social relationships have become an important asset to multinational firms because of the need for appropriate resources (e.g., information, technology, knowledge, access to distribution networks, etc.) to compete effectively in the markets. For example, exchanges based on these linkages can facilitate product innovation, expedite resource exchange and create intellectual capital [45, 65]. Service-centered logic implies that value is defined by and co-created with the consumer and determined by the customer on the basis of value-in-use, rather than being embedded in predefined output [68]. Thus, from a new service development perspective, the customers become not only a necessity, but also an opportunity.

Service firms make their living by accessing, creating, and using information in ways that add value to an enterprise and its stakeholders [28]. Thus, a firm that is located in a cooperative relationship of social interaction likely has greater potential to innovate and exchange know-how with other firms, because of its specific external advantages in the network. Firms gain advantages through close cooperation, and they obtain specific information about new products. Also, they can assess their value with respect to their needs, while producers gain insight into the user or customer needs and can adjust their innovation activity accordingly. Hence,

Hypothesis 2: A firm's social relationship capability has a positive relationship with the firm's innovation intensity.

Innovation intensity and firm performance

The definition of firm performance, explored here, is based on the notion that a firm is an association of productive assets (including individuals) who voluntarily come together to obtain economic advantages [49] Owners of productive assets of a firm will make those assets available to a firm only if they are satisfied that the income they are receiving is at least as large as the income they could expect from any reasonable alternatives [9]. Depending on these insights, it is possible for us to outline a firm's performance by comparing the value that a firm creates using its productive assets with the value that managers of the firm expect to obtain.

New sources of value are generated through novel deployments of resources [59]. New ways of exchanging and combining resources are important to create a firm's value. A firm's innovation is related with organizational learning [13, 65]. The more emphatically knowledge is learned and absorbed, the higher the performance a firm can achieve through the capability of innovation [37].

Innovation is considered vital for its contribution to business performance, and the literature consistently associates it positively with performance. Empirically, this linkage for innovation and its impact on performance was

validated by Han *et al*. [27]. Higher innovation possessed by a firm causes higher organizational performance in the market competition [62]. Considering the operational complexity of a service firm, the intensity of innovation is generally manifested in the form of product modification [69]. The firm requires diverse resources inputs and combinative capabilities [36]. In the light of the growth of a firm, its ability of innovation will generate a competitive edge and business growth in the market [57]. Thus, a firm's innovation has become important for it to increase growth of development and value creation [73].

According to Leiponen's research about the Finnish Community Innovation Survey, more than 20% of service firms reported having launched new services in the previous years [39]. In other words, recent survey data indicates that innovation does occur in the service industry. Because service firms face dynamic demand and market uncertainty, they are likely to pursue more proactive and more aggressive strategies, as uncertainty increases, through innovation activities [48]. The literature states a number of strongly allied concepts of innovation and service firms [3, 54, 71]. The new products or new processes introduced in help incumbent firms to safeguard their market position and sustain growth. In essence, the more dynamic or complex the environment, the greater the compulsion to innovate and the more innovative firms are likely to be. Customer tastes or expectations fluctuate; competitors, for example, introduce new products. The pressure on firms to innovate will be great and, hence, one may anticipate that the intensity of innovation is the decisive factor for service firms.

This study intends to analyze the relationship between the intensity of organizational innovation and a firm's performance. We opted to consider, first of all, the relationship between organizational innovation and financial results, which have been the main focus of research on business strategy. Measure such as sales volumes, change in sales and market share expansion seem appropriate as measures of the firm's performance.

From a broader economic point of view, understanding innovation within service firms becomes vital as the share of the service sector in terms of GDP and employment keeps rising. But service activities' value creation and outcomes have been slow or even negative [39]. The effect of innovation does not always immediately affect economic and financial results. In other words, financial results are seen as a 'lagging indicator' for the measure of a firm [64]. Moreover, certain external factors may favor one firm over another, such as changes of government regulations or production or distribution costs [64]. For this reason, we also consider it appropriate to use perceived measures as considered in the literature. One way in which we posit performance is by examining the outcome of the firm's innovation. Innovative outcomes will materialize over time rather than instantly, so the expected performance—rather than the present performance—should constitute another dependent variable. Furthermore, if goal attainment is at the heart of a firm's performance, then we should also maintain that it is the market performance, rather than the present market performance, that should be assessed. Perceived measures are likely to reflect both enacted and potential outcome [55]. Therefore, we use expected performance as another measure of performance. Consequently, we use two performance indicators to predict a firm's performance. Hence,

Hypothesis 3a: A firm's innovation intensity has a positive relationship with its future performance.

Hypothesis 3b: A firm's innovation intensity has a positive relationship with its growth.

Hypotheses 1 to 3 are summarized in Figure 1.

3. Methods

Sample and date collection

This model is tested on samples of firms in the Taiwanese service industry. Data for this study come from two major sources. Data for social relationship capability, innovation intensity, and expected firm performance are perceptual measures; data for internal venturing capability and firm growth are from industrial secondary archives. A survey questionnaire was developed based on previous literature. We sent it to business managers familiar with the development of the service industry to verify questionnaire items and terms. Some minor changes in wording were made and the questionnaire was then adjusted.

Our sample was drawn from the *Top 5000-The Largest Corporations in Taiwan, 2006,* compiled by China Credit of Information Service, Ltd (CCIS), Taiwan. This data source not only lists the sample companies that we need but also provides data about the firms. Because of missing data for some firms, this study collected data for 1,600 firms as a sample. These questionnaires were sent to senior manager of these service firms. A t-test on the number of employees showed no significant differences in our sample and those that were not included in the sample. After several follow-ups, there were 237 responses. We exclude several incomplete questionnaires. There are 226 complete responses. The samples are comprised of 226 observations, meaning that we provide a valid sample size for the subsequent statistical analysis to be carried out. In order to ascertain that the response is effective, we sent another set of questionnaires to other managers in the responding firms. Nineteen of the second questionnaires were returned. We found a high degree of correlation between the two sets of responses. Hence, we argue that the first collection is sufficient for subsequent hypotheses testing. After the process of collecting questionnaires, we use archival data to provide information for the rest of the constructs as another source of analytic data.

Variable measurement

This study uses the LISREL (Linear Structural Relations) model as the analytical tool. Jöreskog introduced the LISREL model in 1973. The LISREL model consists of

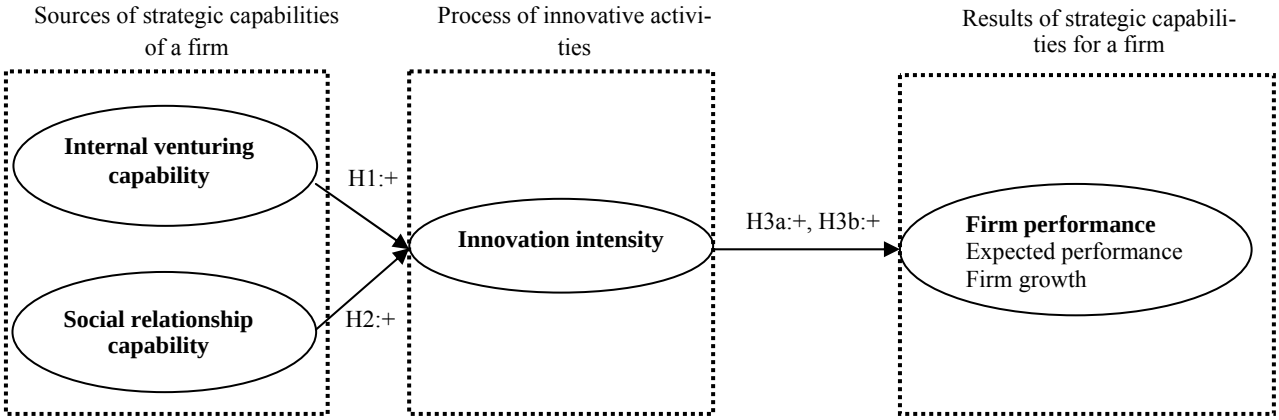


Figure 1. The Hypothesized Model

two parts: the measurem ent model and the structural equation model. The measurement model specifies how latent variables or constructs depend upon or are indicated by the observed variables.

The hypothesized model includes four constructs: internal venturing capability, social relationship capability, the intensity of innovative activity and firm performance. The operational nature of the constructs has been widely discussed in the literature. Eleven variables were developed in this research. It must be noted that two of the eleven variables are objective measures: internal resources advantage and firm growth. The study uses archival data to measure internal resources advantage as one of the measures of a firm's growth in order to avoid common method bias. We explained each measure of the variables as follows:

Internal venturing capability: The employee productivity shows a positive impact on the innovation intensity of a firm, and is the basis of sustainable competitive advantage [56]. The natural logarithm of sales per employee (X1) is a widely used measure of employee productivity [31] and was adopted here in analyses.

Social relationship capability: We measure social relationship capability by the following three areas: (1) cooperation in deciding strategic objectives and goals (X2); (2) cooperation in functional areas such as service products, R&D, purchasing, marketing, human resources, and budgeting (X3); (3) cooperation in implementing new plans for the service design, R&D, or new market entry (X4) [25][34]. We ask respondents to indicate one cooperative partner who is the most important or critical.

Innovation intensity: This study uses the following items to ask respondents how the firm is involved in innovative activities: (1) How many service product innovations per year were produced in your firm? (Y1); (2) What was the extent of formulating new service proposals, including service design and specifications? (Y2); (3) Within the firm, do managers consistently care about the innovative issues? (Y3) [39][42]

Firm performance: Firm performance includes two types of measures. First, we use perceived measures. To assess the perception of expected firm performance, we ask respondent to estimate the expected increase in sales growth (Y4), profitability (Y5) and market share (Y6) with 5-point Likert type scale (1 = very small to 5 = very high). Second, we adopt firm growth, measured as *change in sales* (Y7) [5][6], as the other proxy for firm performance.

Table 1 shows the descriptive statistics of the independent and dependent variables analyzed in the hypothesized model.

4. Results

Measuring model evolution

We use composite reliability, which is analogous to coefficient α [20], and average variance extracted to measure *internal consistency*.

Estimates of composite reliability and average variance extracted are sufficient to support internal consistency. All information is shown in Table 2. In addition, we also evaluate the discriminant validity of the model. We compared chi-square value for a measurement model, and constrained the correlation to equal one to a baseline model without this constraint. All the measures of constructs in the measurement model are significant in difference and achieve discriminant validity.

Structural model estimation

With respect to the fitness of statistics for the full model ($\chi^2_{(41)}$=112.18, p=0.00, GFI=0.92, AGFI=0.87, CFI=0.95, NNFI=0.94), the chi-square is significant, which is usually influenced by sample sizes. All the other statistics are within the acceptable ranges. All the other statistics are within the acceptable ranges, which indicate a good model fit.

In order to empirically test theoretic hypotheses, the hypothesized model is examined by using LISREL, in which the four paths between different latent variables are estimated. Our empirical results show that the internal resources construct has a non-significantly negative effect

Table 1. Descriptive statistics and pearson correlation analysis (N=226, * p<0.01; ** p<0.05)

	Mean	SD	X1	X2	X3	X4	Y1	Y2	Y3	Y4	Y5	Y6	Y7
X1	2.46	1.38	1										
X2	4.19	0.77	− 0.040	1									
X3	4.06	0.96	0.027	0.716**	1								
X4	3.99	1.07	0.047	0.598**	0.791**	1							
Y1	3.73	0.65	− 0.040	0.764**	0.551**	0.445**	1						
Y2	3.65	0.98	0.018	0.404**	0.314**	0.251**	0.592**	1					
Y3	2.46	0.98	− 0.017	0.382**	0.324**	0.259**	0.584**	0.357**	1				
Y4	3.69	0.93	− 0.002	0.447**	0.274**	0.173**	0.564**	0.314**	0.305**	1			
Y5	3.65	1.05	0.032	0.311**	0.259**	0.244**	0.422**	0.270**	0.312**	0.700**	1		
Y6	3.52	1.11	− 0.041	0.382**	0.261**	0.166*	0.459**	0.256**	0.234**	0.732**	0.510**	1	
Y7	0.07	0.27	0.212	0.075	0.026	0.131	0.038	0.100	− 0.042	0.043	0.061	0.050	1

Table 2. Matrix of latent constructs for full sample (IV: internal venturing capability , SR: social relationship advantage; II: innovation intensity; EP: expected performance, FG: firm growth; ** p<0.05)

	IV	SR	II	EP	FG	Composite reliability	AVE
IV	1					–	–
SR	0.019	1				0.85	0.66
II	− 0.012	0.517**	1			0.76	0.53
EP	0.005	0.343**	0.464**	1		0.83	0.63
FG	0.212**	0.089	0.038	0.059	1	–	–

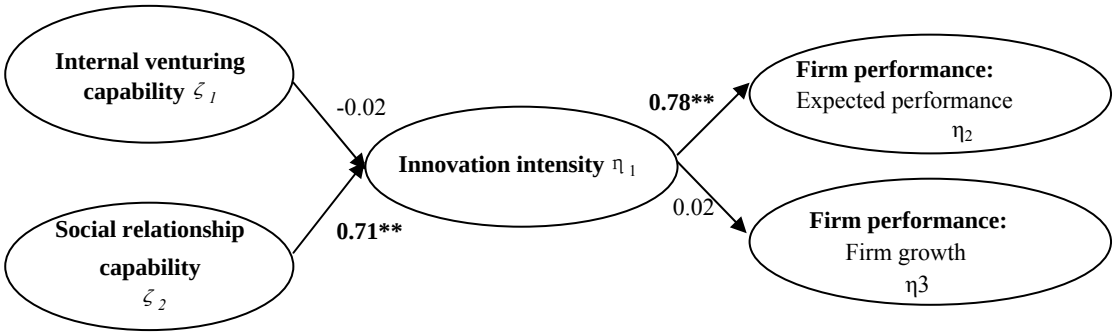


Figure 2. The results of Hypothesized model[a]

[a]: The figure depicts a structural model with maximum likelihood estimates. We set the error variances for single indicator at 0, with loadings (lambdas) fixed at 1 (that is, X1 and Y7 with each corresponding latent variable)

Table3. Analysis of competing structural model (IV: internal venturing capability, SR: social relationship advantage; II: innovation intensity; EP: expected performance, FG: firm growth; ** p<0.01)

Hypothesized model		Rival model	
Path	**Estimate**	**Path**	**Estimate**
IV→II	− 0.02	IV→EP	0.01
SR→II	0.71**	IV→FG	0.04**
II→EP	0.78**	SR→EP	− 0.02
II→FG	0.02	SR→FG	0.04
		II→EP	0.79**
		II→FG	− 0.01
$\chi^2_{(41)}$=112.18, p=0.00, GFI=0.92, AGFI=0.87, CFI=0.95 NNFI=0.94, PNFI=0.69		$\chi^2_{(37)}$=102.37, p=0.00, GFI=0.92, AGFI=0.86, CFI=0.96 NNFI=0.94, PNFI=0.63	

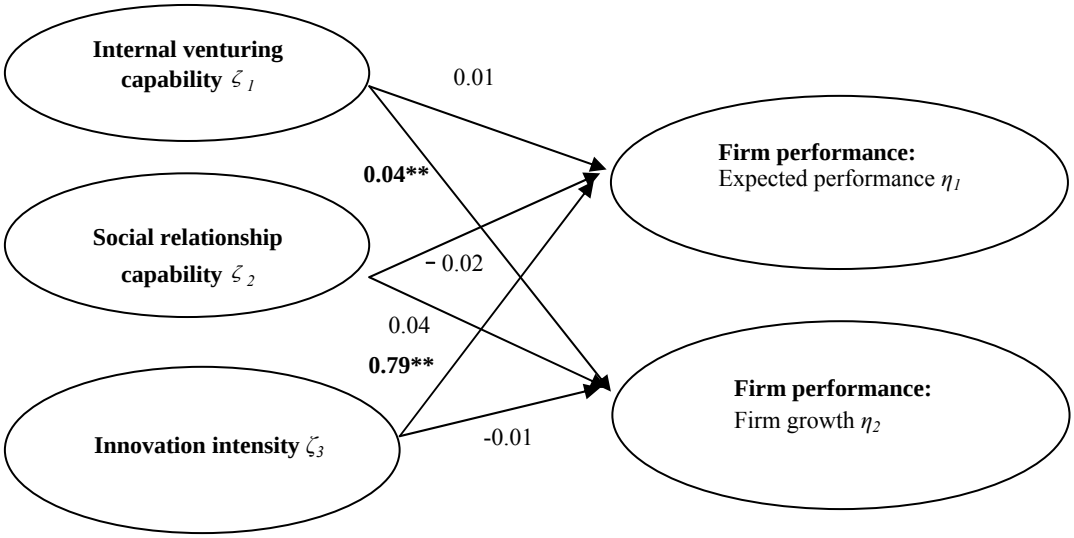


Figure 3. The rival model and the results (all predicted directions are positive as literatures suggested)

on the intensity of innovative activity (γ_{11}=-0.02, t-value=-0.86). Hypothesis 1 is thus not supported. With regard to Hypothesis 2, it is found that the external resources have a significantly positive effect on the intensity of innovative activity (γ_{12}=0.71, t-value=10.23). Thus Hypothesis 2 is supported. So is Hypothesis 3a, that the intensity of innovative activity is significant positively with the expected firm performance (β_{21}=0.78, t-value=8.37). Hypothesis 3b, that the effect of the intensity of innovative activity on the financial performance does not have a significant positively prediction, is not supported by this test (β_{31}=0.02, t-value=0.55). Results of the parameter estimate are summarized in Table 3 and showed in Figure 2.

Comparison with a rival model

We don't use the full model because this study argues that the intensity of innovative activities is a key mediate variable for firm performance[1]. Comparison with a rival model is an important way to assess the power of a specified model [7]. The rival model is showed in Figure 3. Our hypothesized model is based in an elaborate theory that hypothesizes a specific nomological inference of constructs. We thus compared our hypothesized model with the rival model using the following criteria: (1) overall fit, as measured by CFI; (2) percentage of the model's statistically significant parameters; (3) ability to explain the variance in the outcomes of interest, as measures by square multiple correla tions (SMC) of the focal and outcome variables; and (4) parsimony, as measured by the PNFI [44]. The direct paths of the rival model are supported in much of the literature [9] [16] [49] [63]. However, this study argues that the strategic operational process, which is the intensity of innovative activity, is the critical factor in a firm's performance. Thus, we do the comparison to ascertain the effectiveness of the hypothesized model.

The CFI for the rival model is a little higher than for the hypothesized model (CFI= 0.96 v.s 0.95). In our hypothesized model, 50% (or 2 of 4) aspects of the path are significant, whereas only 33% (or 2 of 6) aspects of the path are significant in the rival model. Moreover, little additional explanatory power is gained from the additional two paths in which the increment to SMC is 0.05 (for financial performance). In comparing the models, we see there is a difference in parsimony between the hypothesized and rival models (4 versus 6 paths). CFI is not an indicator that accounts for the parsimony difference, so we compare the two models using PNFI. The PNFI of the hypothesized model is 0.69, exceeding the rival's 0.63. Although there is no guideline to determine what the significant difference in PNFI values is, we note that a sacrifice of PNFI value is 91% (from 0.69 to 0.63). We accomplish a great improvement in parsimony without sacrificing too much CFI. Hence, a sacrifice is worthy for parsimony.

Based on these findings, we acknowledge that this comparison provided added confidence in the constructs of our hypothesized model. The intensity of innovative activity also represents a critical process of advantageous creation as a sufficient predictor for the firm.

5. Discussions

Overall, the results of this study provide support for the argument that social relationships (that is, external resources of a firm) facilitate the innovative activity of service firms, and that the innovative activity has a positive effect on the expected firm performance. This finding is robust at the dyadic level. The first finding is consistent with previous studies showing that interorganizational relationships are positively related to innovation intensity. The result also supports theorists who emphasized the

[1] We still tested the fit of the full model. According to the result, estimation of full model is no better than the hypothesized model. Hence, we argue that our hypothesized model is an effective model.

importance of acquiring external knowledge for product development [71].

It is clear that cooperation and innovation among service firms are important as operational measures. Here we have the implication that focus on the cooperation between firms will create innovation as well as market opportunity development. Cooperation creates interfirm benefits that are consistent with the literature on organizational advantage. The more interaction they have, the more business possibilities there will be; the more cooperation they have, the more their market opportunities are likely to be productive. The concept of social relationships, therefore, is central to the understanding of innovation and value creation [45]. An important point to note is that these productive possibilities need to be fully exploited through a firm's exchanges and cooperation [43]. On the basis of this argument, it seem reasonable to argue that innovation is better facilitated by interfirm cooperation, because the interaction will help a firm to consolidate existing markets and create new market share and market value.

However, our analysis found that the intensity of innovation is not influenced by internal resources and may even cause a negative effect, although it is non-significant. This is contrary to our argument and it is interesting to discuss. In theory, human resources will be positively related to firm innovation [63] [71]. But the relationship between the firm's human resources and innovation is diametrically opposed to what we predicted in Hypothesis 1.

We infer the reasons for this result as follows: (1) We didn't measure the organizational climate. Some people would be frustrated by a chaotic environment and seek some sort of stability to improve efficiency of the status quo. Others would be frustrated by the long list of unrealized opportunities for improvement. Service firms are located in a competitive market, which affects a firm's business orientation. (2) The codifiability of the knowledge assets is the other reason. Knowledge is embedded on the employee [49]. Service innovations sometimes are easier to create by codifiable knowledge than relatively tacit. All knowledge assets are codifiable to varying degrees [36]. All else being equal, the knowledge assets on the employee are codifiable mostly that make the insignificant effect. (3) RBV focuses on the firm-level analysis and our measures of the latent variables are also likely to represent firm-level information (archival data consists of firms' information).

Furthermore, innovation is needed to narrow down to the refined or specific level. A firm is composed of different kind of divisions, and innovation often rests with specific divisions or individuals [49]. Therefore, for example, we asked managers to respond regarding the innovative activity and that may cause bias between the two kinds of measures. (4) The type of innovation should be considered further. Innovation can be subdivided into radical and incremental. A large number of radical innovations spring from autonomous strategic behavior, while the greatest percentage of incremental innovations come from induced strategic behavior [59]. Therefore, employee is the reasons for why cause an unwilling result according to the different types of innovation [62]. Certainly, all these reasons may help explain the negative result of Hypothesis 1. The path from internal resources to the intensity of innovative activity should be addressed specifically in future study.

We also examined the relationship among the three latent variables regarding the intensity of innovative activity and the two kinds of firm performance. We showed how the mediate variable contributed to the firm's performance. While the intensity of innovative activity is related positively to the expected firm performance, contradicting to the prediction, there is a non-significant effect between innovation and financial performance. Internally developed innovations result from deliberate efforts. Most successful firms develop both radical and incremental innovations over time. Although critical to long-term competitiveness and performance, the outcomes of investments in innovative activities are uncertain and often not achieved in the short term, meaning that patience is required as firms evaluate the outcomes of their innovation efforts [6][48]. Therefore, we infer that innovation is an activity for future business and future growth but not on the instant. Thus Hypothesis 3a is supported. Financial performance shows that the last year's business outcome was not positively influenced by innovation, so Hypothesis 3b is not supported.

6. Conclusions and Future Research Direction

The view of strategic capabilities presented here includes proprietary resources that exist within a firm and social capital located among firms. To enhance efficiency and effectiveness of the intensity of innovative activities, cooperation between firms is the most important factor. The study results suggest that innovative activity should be integrated into managerial considerations, and that it is a critical process for firm performance. This study finds that the roots of innovation of a service firm are deeply embedded in social relationships. Second, this study also identifies that innovation is a business activity for the future. Service firms intend to innovate new products and processes, and create new market opportunities so as to sustain competitive advantage. Therefore, a firm's strategy underlies its theory of how to compete in the market successfully. Whether it is deliberate or emergent strategy, a firm generally needs to address in the best operational way what the critical economic processes in an industry or market are and how it can take advantage of these to generate competitive advantage for itself [9]. In conclusion, service firms experience a competitive advantage when their actions create economic value and when other competitors cannot pursue the same activity. This study

also found evidence for the suggestion that firms that invest more in cooperative relationships share value and common benefits. Moreover, they also need to encourage the development of strong personal and team relationships, a high level of trust and strong connections across porous boundaries [16][45][65]. The hypothesized framework this study develops could offer a useful ground for advanced tests of different phenomena.

There are several limitations in this study. First, the service firms' context of this study limits its potential generalizability for respective industries. Service firms include diverse businesses such as advertising, network information supply services, and so on. More studies focusing on different categories of service firms may shed light on the generalizability of the theoretical position developed here. Second, it is possible that the causality may flow in opposition to that proposed here. For example, perhaps the concern about the intensity of innovative activity promotes social interaction. Although we have built our hypotheses upon existing theories and past arguments, future research may show that reserve or interactive relationships exist. Third, the model was tested empirically in a Taiwanese sample. Therefore, future study in this area could replicate this study and extend it to other economic systems to see if the findings would be similar to those reported here. Finally, the interaction between internal resource advantage and social relationship advantage may constitute another direction for future study. This study offers interesting findings and contributes to the understanding of the strategic capabilities and innovation intensity of service firms.

REFERENCES

[1] P.S. Adler and S.-W. Kwon, "Social capital: prospects for a new concept", *Academy of Management Review* 27(1), 2002, pp.17-40.

[2] S. Agarwal and S.N. Ramaswami, "Choice of foreign market entry mode: impact of ownership, location and internalization factors", *Journal of International Business Studies* 23(1), 1992, pp.1-28.

[3] Alam, "An exploratory investigation of user involvement in new service development", *Journal of the Academy of Marketing Science* 30(3), 2002, pp.250-261.

[4] R. Amit and P.J.H. Schoemaker, "Strategic asset and organization rent", *Strategic Management Journal* 14(1),1993, pp.33-46.

[5] U. Andersson, M. Forsgren and U. Holm, "Subsidiary embeddedness and competence development in MNCs-A multi-level analysis", *Organization Studies* 22(6), 2001, pp.1013-1034.

[6] U. Andersson, M. Forsgren and U. Holm, "The strategic impact of external networks: subsidiary performance and competence development in the multinational corporation", *Strategic Management Journal* 23(11), 2002, pp.979-996.

[7] R.P. Bagozzi and Y. Yi, "On the evaluation of structural equation models", *Academy of Marketing Science* 16(1), 1998, pp.74-94.

[8] J.B. Barney, "Firm resources and sustained competitive advantage", *Journal of Management* 17(1), 1991, pp.99-120.

[9] Barney, J.B., *Gaining and Sustaining Competitive Advantage*, 3rd, Prentice Hall, 2006.

[10] Bell, D., *The Coming of Post-Industrial Society: A Venture in Social Forecasting*: Basic Books, Reissue edition, 2001.

[11] Burt, R.S., *Structural Holes: the Social Structure of Competition*, Cambridge. MA: Harvard University Press, 1992.

[12] H.-J. Cho and V. Pucik, "Relationship between innovation, quality, growth, profitability, and market value", *Strategic Management Journal* 26(6), 2005, pp.555-575.

[13] W.M. Cohen and D.A. Levinthal, "Absorptive capacity: a new perspective on learning and innovation", *Administrative Science Quarterly* 35(1), 1990, pp.128-152.

[14] I. Dierickx and K. Cool, "Asset stock accumulation and sustainability of competitive advantage", *Management Science* 35(12), 1989, pp.1504-1514.

[15] Doz , Y.L. and G. Hamel, *Alliance Advantage*. Boston, MA: Harvard Business School Press, 1998.

[16] Dyer, J.H., *Collaborative Advantage: Winning through Extended Enterprise Supplier Networks*. New York: Oxford University Press, 2000.

[17] I. Ekeledo and K. Sivakumar, "International market entry mode strategies of manufacturing firms and service firms: a resource-based perspective", *International Marketing Review* 21(1), 2004, pp.68-101.

[18] M.K. Erramilli, "Entry Mode Choice in Service Industries", *International Marketing Review* 7(5), 1990, pp.50-62.

[19] M.K. Erramilli and C.P. Rao, "Service firms' international entry-mode choice: a modified transaction-cost analysis approach", *Journal of Marketing* 57(3), 1993, pp.19-38.

[20] C. Fornell and D.F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research* 18(1), 1981, pp.39-50.

[21] C. Freeman and J. Hagedoorn, "Catching up or falling behind: patterns in international interfirm technology partnering", *World Development* 22(5), 1994, pp.771-780.

[22] M.S. Granovetter, "Economic action and social structure: the problem of embeddedness", *American Journal of Sociology* 91(3), 1985, pp.481-510.

[23] M.S. Granovetter, "The strength of weak ties", *American Journal of Sociology* 78(6), 1973, pp. 1360-1380.

[24] R.M. Grant, "The resource-based theory of competitive advantage: implications for strategy formulation", *California Management Review* 33(2), 1991, pp.114-135.

[25] R. Gulati, "Alliance and networks", *Strategic Management Journal* 19(4), 1998, pp.293-318.

[26] G. Hamel, "Competition for competence and inter partner learning within international strategic alliances", *Strategic Management Journal* 12(summer special issue), 1991, pp.83-103.

[27] J. Han, N. Kim and R. Srivastava, "Market orientation and organisational performance: is innovation a missing link?", *Journal of Marketing* 62(4), 1998, pp.30-46.

[28] T. Hennig-Thurau, K.P. Gwinner and D.D. Gremler, "Understanding relationship marketing outcomes: an integration of relational benefits and relationship quality", *Journal of Service Research* 4(3), 2002, pp.230-247.

[29] T. Hennig-Thurau, G. Walsh and O. Wruck, "An investigation into the factors determining the success of service innovations: the case of motion pictures", *Academy of Marketing Science Review* 6(1), 2001, pp.1-23.

[30] M.A. Hitt, R.E. Hoskisson and R.D. Ireland, "Mergers and acquisitions and managerial commitment to innovation in M-Form firms", *Strategic Management Journal* 11(special issue), 1990, pp.29-47.

[31] M.A. Huselid, "The impact of human resource management practices on turnover, productivity, and corporate financial performance", *Academy of Management Journal* 38(3), 1995, pp.635-672.

[32] A.C. Inkpen, "A note on the dynamics of learning alliances: competition, cooperation, and relative scope", *Strategic Management Journal* 21(7), 2000, pp.775-779.

[33] Jackson, S.E., A. DeNisi and M.A. Hitt, *Managing Knowledge for Sustained Competitive Advantage: Designing Strategies for Effective Human Resource Management*: Pfeiffer, 2003.

[34] P.V. Kale, H. Singh and H. Perlmutter, "Learning and protection of proprietary assets in strategic alliances: building relational capital", *Strategic Management Journal* 21(3), 2000, pp.217-237.

[35] B. Kogut, "Joint ventures: theoretical and empirical perspectives", *Strategic Management Journal* 9(4), 1988, pp.319-332.

[36] B. Kogut and U. Zander, "Knowledge of the firm, combinative capabilities, and the replication of technology", *Organization Science* 3(3), 1992, pp.383-397.

[37] P.J. Lane and M. Lubatkin, "Relative absorptive capacity and interorganizational learning", *Strategic Management Journal* 19(5), 1998, pp.461-477.

[38] J.N. Larsen, "Knowledge, human resources and social practice: the knowledge-intensive business service firm as a distributed knowledge system", *The Service Industries Journal* 21(1), 2001, pp.81-102.

[39] A. Leiponen,"Organization of knowledge and innovation: the case of Finnish business services", *Industry and Innovation* 12(2), 2005, pp.185-203.

[40] H. Littunen and T. Tohmo, "The high growth in new metal-based manufacturing and business service firms in Finland", *Small Business Economics* 21(2), 2003, pp.187-200.

[41] J.G. March, "Exploration and exploitation in organizational learning", *Organization Science* 2(1), 1991, pp.71-87.

[42] B. Menguc and S. Auh, "Creating a firm-level dynamic capability through capitalizing on market orientation and innovativeness", *Academy of Marketing Science* 34(1), 2006, pp.63-73.

[43] P. Moran and S. Ghoshal, "Markets, firms, and the process of economic development", *Academy of Management Review* 24(3), 1999, pp.390-412.

[44] R.M. Morgan and S.D. Hunt, "The commitment-trust theory of relationship marketing", *Journal of Marketing* 58(3), 1994, pp.20-38.

[45] J. Nahapiet and S. Ghoshal, "Social capital, intellectual capital, and the organizational advantage", *Academy of Management Review* 23(2), 1998, pp.242-266.

[46] J. Narver and S.F. Slater, "The effect of a market orientation on business profitability", *Journal of Marketing* 54(4), 1990, pp.20-35.

[47] I. Nonaka, "The knowledge-creating company", *Harvard Business Review* 69(6), 1991, pp.96-104.

[48] A. Ozsomer, R. Calantone and A.D. Benedetto, "What makes firms more innovative? a look at organizational and environmental factors", *Journal of Business and Industrial Marketing* 12(6), 1997, pp.400-416.

[49] E.T. Penrose, *The Theory of the Growth of the Firm*, 3rd, Oxford: Oxford University Press, 1995.

[50] P.D.S. Perez and J.M.G. Falcon, "The influence of human resource management in savings bank performance", *The Service Industries Journal* 24(2), 2004, pp.51-66.

[51] M.A. Peteraf, "The cornerstones of competitive advantage: a resource based view", *Strategic Management Journal* 14(3), 1993, pp.179-191.

[52] M.E. Porter, *Competitive Advantage: Creating and Sustaining Superior Performance*, New York: The Free Press,1985.

[53] C.K. Prahalad and G. Hamel, "The core competence of the corporation", *Harvard Business Review* 68(2), 1990, pp.79-91.

[54] C.K. Prahalad and V. Ramaswamy, "Co-opting customer competence", *Harvard Business Review* 78(1), 2000, pp.79-87.

[55] K. Provan, "Interorganizational cooperation and decision making autonomy in a consortium multihospital system", *Academy of management Review* 9(3), 1984, pp.494-504.

[56] L. Raymond and J. St-Pierre, "Customer dependency in manufacturing SMEs: implications for R&D and performance", *Journal of Small Business and Enterprise Development* 11(1), 2004, pp.23-33.

[57] C.A. Romano, "Identifying factors which influence product innovation: a case study approach", *Journal of Management Studies* 27(1), 1990, pp.75-95.

[58] R.W. Schmenner, "Service businesses and productivity", *Decision Sciences* 35(3), 2004, pp.335-347.

[59] Schumpeter, J.A., *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*, Cambridge, MA: Harvard University Press, 1934.

[60] C. Styles and T. Ambler, "The impact of relational variables on export performance: an empirical investigation in Australia and the UK", *Australian Journal of Management* 25(3), 2000, pp.261-281.

[61] M. Subramaniam and M.A. Youndt, "The influence of intellectual capital on the types of innovative capabilities", *Academy of Management Journal* 48(3), 2005, pp.450-463.

[62] D. A. Subramanian and S. Nilakanta, "Organizational

innovativeness: exploring the relationship between organizational determinants of innovation, types of innovations, and measures of organizational performance", *Omega* 24(6), 1996, pp.631-647.

[63] D.J. Teece, G. Pisano and A. Shuen, "Dynamic capabilities and strategic management", *Strategic Management Journal* 18(7), 1997, pp.509-533.

[64] Thompson, A.J., J.E. Gamble and A.J. Strickland, *Strategy: Winning in the Marketplace: Core Concepts, Analytical Tools, Cases*, 2nd, McGraw-Hill/Irwin, 2006.

[65] W. Tsai and S. Ghoshal, "Social capital and value creation: the role of intrafirm networks", *Academy of Management Journal* 41(4), 1998, pp.464-476.

[66] M. Tushman and P. Anderson, "Technological discontinuities and organizational environments", *Administrative Science Quarterly* 31(3), 1986, pp.439-465.

[67] Ulrich, D. and D. Lake, *Organisational Capability: Competing from the Inside Out*, New York: John Wiley & Sons, 1990.

[68] S.L. Vargo and R.F. Lusch, "Evolving to a new dominant logic for marketing", *Journal of Marketing* 68(1), 2004, pp.1-17.

[69] F.J.H.M. Verhees and M.T.G. Meulenberg, "Market orientation, innovativeness, product innovation, and performance in small firms", *Journal of Small Business Management* 42(2), 2004, pp.134-154.

[70] B. van Ark and M. Piatkowski, "Productivity, innovation and ICT in Old and New Europe", *International Economics and Economic Policy*, 1(2/3), 2004, pp.215-246

[71] von Hippel, E., *The Source of Innovation*, New York: Oxford University Press, 1988.

[72] B. Wernerfelt, "A resource-based view of the firm", *Strategic Management Journal* 5(2), 1984, pp.171-180.

[73] J.A. Wolff and T.L. Pett, "Small-firm performance: modeling the role of product and process improvements", *Journal of Small Business Management* 44(2), 2006, pp.268-284.

[74] T. Yoshikawa, A.A. Rasheed, D.K. Datta and J. Rosenstein, "Financial and product market integration: responses of Japanese firms", *Management International Review* 46(5), 2006, pp.529-555.

J. Serv. Sci. & Management, **2008, 1: 123-127**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

Impact of Supply Chain Coordination for Deteriorating Goods with Stock-Dependent Demand Rate

Chuanxu Wang

School of Economy and Management，Shanghai Maritime University Shanghai，China
Email: cxwang@shmtu.edu.cn

ABSTRACT

To analyze effects of supply chain coordination for deteriorating goods with stock-dependent demand rate, this paper presents decision models for order quantity and ordering cycle under two scenarios(decentralized supply chain, centralized supply chain). Numerical study is carried out to demonstrate the effectiveness of the proposed models, and to analyze the impact of supply chain coordination on supply chain profit. Sensitivity analysis is performed to study the impact of different parameters associated with the model, such as the rate of deterioration, the retailer's purchase cost, the manufacturer's production cost, the retailer's and manufacturer's holding cost on the supply chain profit increase percentages generated by the supply chain coordination.

Keywords: *supply chain coordination, deteriorating goods, ordering cycle, stock-dependent demand*

1. Introduction

In real life, many inventory goods, such as agricultural products, fashion goods, drugs and high-tech products, are subject to depletion through spoilage, shrinkage, decay and obsolescene [1]. The deterioration is quite prevalent and should not be disregarded. Inventory management for deterioration goods has received many attentions from researchers and practitioners. Most of the existing researches focus on the EOQ-based inventory decision models. Ghare and Schrader [2] presented the EOQ model by considering the combined effects of demand, usage and linear decay. Covert and Philip [3]used the variable deterioration rate of the two-parameter Weibull distribution, to formulate the inventory decision model under the assumptions of a constant demand rate, with no shortages allowed. Philip [4] modified this model by using the deterioration rate of the three-parameter Weibull distribution. Tadikamalla [5] adopted gamma distributed deterioration under constant demand over time, without shortages. Moon and Lee [6] presented the EOQ model with a normally distributed deterioration rate. Other deterioration inventory models have extended prior research by considering a time-varying demand function, with or without shortages. Dave and Patel [7] proposed an EOQ model under time-proportional demand, with no shortages allowed. Sachan [8] extended their model by considering shortages. Bahari-Kashani [9] generalized the problem by permitting variations in both replenishment cycle length and order quantity. Bose *et al* [10]developed an EOQ model for deterioration items incorporating the effects of inflation, time value of money, a linearly time-dependent demand rate and shortages. Replenishment decision models under time-proportional demand and exponentially decaying deterioration rate was developed in [1].

It is observed that large quantities of consumer goods displayed in a supermarket generate higher demands. Silver and Peterson [11] noted that the sales at the retail level tend to be proportional to the inventory displayed. Gupta and Vrat [12], Mandal and Phaujdar [13], Baker and Urban [14], Datta and Pal [15], etc developed the EOQ models with stock-dependent demand rate. Mandal and Phaujdar [16], Pal *et al*. [17] developed the inventory models for deteriorating items with stock-dependent demand rate. In this paper, we have extended these works, on deteriorating inventory research, by considering deteriorating goods with stock-dependent demand in a two-echelon supply chain consisting one manufacturer and one retailer, the objective is to investigate the effects of supply chain coordination on profit increase in the supply chain, and study the impact of different parameters associated with the model, such as the rate of deterioration, the retailer's purchase cost, the manufacturer's production cost, the retailer's and manufacturer's holding cost on the supply chain profit increase percentages generated by the supply chain coordination.

2. Assumptions and Notations

2.1. Assumptions

(1) The retailer replenishes the stocks from the exclusive source on an EOQ basis. Replenishments are instantaneous.

(2) Lead time is assumed to be zero for the sake of simplicity.

(3) No backorders are allowed.

(4) Demand rate is dependent on the instantaneous inventory level. The demand rate $d(I)$ of the item, when the inventory is I, is considered in the form $d(I) = \alpha I^{\beta}$, where $\alpha > 0$ and $0 < \beta < 1$ are scale and shape parameters (Baker and Urban 1988).

(5) The manufacturer's production rate is greater than or equal to the demand rate facing the retailer.

(6)The manufacturer is a make-to-order producer; it has a lot-for-lot production policy in response to the retailer's demand. In this particular case, the length of the manufacturer's production cycle is equal to the length of the retailer's replenishment cycle.

2.2. Notations

p the sale price for the retailer;
S the order cost per order for the retailer;
M the setup cost per lot for the manufacturer;
$c^r(c^m)$ procurement (manufacturing) cost per unit for retailer (manufacturer);
d the demand rate of the item in the marketplace;
q the production rate;
$h^r(h^m)$ the inventory holding cost as a fraction of the inventory cost for the retailer(manufacturer);
$\theta(t)$ the deterioration rate facing both the retailer and manufacturer，$0 \le \theta \le 1$；
$I^r(t), I^m(t)$ inventory level at time t for the retailer (manufacturer)；
Q the order quantity for the retailer;
T the replenishment cycle (or production cycle) for the retailer(manufacturer);

3. Basic Model

In this section, we first derive the profit model for the decentralized supply chain. Later, we present the profit model derived by considering the centralized supply chain.

3.1. The Decentralized Supply Chain

In the decentralized supply chain, each entity within the supply chain aims to maximize its own profit functions, with no consideration given to its counterpart's reaction or profit. The retailer makes a replenishment decision based on an EOQ policy that includes inventory holding cost and ordering cost.

During the replenishment cycle, the change in retailer's inventory level depends on demand and deterioration and is given by [1]:

$$\frac{dI^r(t)}{dt} + \theta I^r(t) = -\alpha[I^r(t)]^{\beta} \quad 0 \le t \le T, \tag{1}$$

As shown in Pal et al. (1993) [17], equation (1) can be rewritten as

$$\frac{1}{\theta}\left\{\frac{1}{I^r(t)} - \frac{\alpha[I^r(t)]^{\beta-2}}{\theta + \alpha[I^r(t)]^{\beta-1}}\right\} dI^r(t) = -dt \tag{2}$$

By integrating, we get

$$\ln\frac{[I^r(t)]^{\beta-1}}{\theta + \alpha[I^r(t)]^{\beta-1}} = \ln e^{(1-\beta)\theta t} + \ln C_1, \tag{3}$$

where C_1 is integration constant.

It can be rearranged as

$$\frac{1}{\alpha + \theta[I^r(t)]^{1-\beta}} = C_1 e^{(1-\beta)\theta t} \tag{4}$$

By using the boundary condition on inventory $I^r(0) = Q$， we can get

$$C_1 = \frac{1}{\alpha + \theta Q^{1-\beta}} \tag{5}$$

Substituting (5) in (4), the retailer's inventory level at time t ($0 \le t \le T$) can be expressed as

$$I^r(t) = [(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta t} - \alpha]^{1/(1-\beta)} / \theta^{1/(1-\beta)} \tag{6}$$

The inventory holding cost in a cycle for the retailer is

$$HC^r = h^r c^r \int_0^T I^r(t)dt \tag{7}$$

The retailer's total number of deteriorated goods in a cycle is given by

$$\begin{aligned} Q^{D,r} &= Q - I^r(T) - \int_0^T \alpha[I^r(t)]^{\beta} dt \\ &= Q - [(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta T} - \alpha]^{1/(1-\beta)} / \theta^{1/(1-\beta)} \\ &\quad - \frac{\int_0^T \alpha[(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta t} - \alpha]^{\beta/(1-\beta)} dt}{\theta^{\beta/(1-\beta)}} \end{aligned} \tag{8}$$

The profit per unit time for the retailer can be expressed as

$$\begin{aligned} \prod_1^r &= (p - c^r)\int_0^T \alpha[(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta t} - \alpha]^{\beta/(1-\beta)} dt / T\theta^{\beta/(1-\beta)} - \frac{S}{T} \\ &\quad - \frac{h^r c^r \int_0^T \alpha[(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta t} - \alpha]^{1/(1-\beta)} dt / \theta^{1/(1-\beta)}}{T} - \frac{c^r Q^{D,r}}{T} \end{aligned} \tag{9}$$

The profit function (9) is highly nonlinear and cannot be solved by analytical methods. We solve it by using optimization technology on computer and get the following optimal values:

Since the manufacturer is a make –to-order producer, it has a lot-for-lot production policy in response to the retailer's demand. During the production cycle, the change in manufacturer's inventory level is due to the combined effect of production and deterioration:

$$\frac{dI^m(t)}{dt} = q - \theta I^m(t) \quad t_m \le t \le T \ , \tag{10}$$

where t_m and T are the starting and stopping production times, respectively. With a make-to-order policy, the production quantity of the manufacturer is equal to the demand quantity of the retailer. Therefore, we get:

$$\int_0^T I^r(t)dt = \int_{t_m}^T I^m(t)dt \tag{11}$$

Solving the equation (6) by using the method proposed by Spiegel (1960)[18] , we can get:

$$I^m(t) = \frac{q}{\theta} + C_2 e^{-\theta t} , \tag{12}$$

where C_2 is constant. By using the boundary condition on inventory $I^m(t_m) = 0$， we can obtain:

$$C_2 = -\frac{q}{\theta} e^{\theta t_m}$$

Therefore, the manufacturer's inventory level at time t ($t_m \le t \le T$) can be expressed as

$$I^m(t) = e^{-\theta(t-t_m)} \int_{t_m}^t e^{\theta(t-t_m)} q \ dt = \frac{q}{\theta} - \frac{qe^{-\theta(t-t_m)}}{\theta} \tag{13}$$

The inventory holding cost in a cycle for the manufacturer is

$$HC^m = h^m c^m \int_{t_m}^T I^m(t)dt = h^m c^m \int_{t_m}^T \left(\frac{q}{\theta} - \frac{q}{\theta} e^{-\theta(t-t_m)} \right) dt$$

$$= h^m c^m \frac{q}{\theta^2} [-1 + e^{-\theta(T-t_m)} + (T-t_m)\theta] \tag{14}$$

The manufacturer's total number of deteriorated goods in a cycle is given by

$$Q^{D,m} = q(T-t_m) - I^m(T) - Q \tag{15}$$

The profit per unit time for the manufacturer is

$$\prod_1^m = \frac{(c^r - c^m)Q}{T} - \frac{M}{T} - \frac{HC^m}{T} - \frac{c^m Q^{D,m}}{T}$$

$$= \frac{c^r Q}{T} - \frac{M}{T} - \frac{h^m c^m q[-1 + e^{-\theta(T-t_m)} + (T-t_m)\theta]}{T\theta^2}$$

$$- \frac{c^m q(T-t_m)}{T} + \frac{c^m q}{T\theta} - \frac{c^m q e^{-\theta(T-t_m)}}{T\theta} \tag{16}$$

Based on the optimal order quantity and ordering cycle as well as equation (11), we can obtain optimal profit for the manufacturer using (16).

3.2. The Centralized Supply Chain

In the centralized supply chain, the order quantity and replenishment cycle are determined by considering the total profit incurred by both the retailer and the manufacturer, so that the overall profit is maximized. The centralized supply chain requires information sharing between manufacture and retailer. Sequential or concurrent engineering will be beneficial to the information sharing.

In this case, the total profit per unit time for the supply chain is the sum of Equation (9) and (16).

$$\prod_2^{sc} = \prod_1^r + \prod_1^m =$$

$$\frac{p\int_0^T \alpha[(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta t} - \alpha]^{\beta/(1-\beta)} dt}{T\theta^{\beta/(1-\beta)}} - \frac{S}{T}$$

$$- \frac{h^r c^r \int_0^T \alpha[(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta t} - \alpha]^{1/(1-\beta)} dt}{T\theta^{1/(1-\beta)}}$$

$$+ \frac{c^r[(\alpha + \theta Q^{1-\beta})e^{-(1-\beta)\theta T} - \alpha]^{1/(1-\beta)}}{T\theta^{1/(1-\beta)}}$$

$$- \frac{M}{T} - \frac{h^m c^m q[-1 + e^{-\theta(T-t_m)} + (T-t_m)\theta]}{T\theta^2}$$

$$- \frac{c^m q(T-t_m)}{T} + \frac{c^m q}{T\theta} - \frac{c^m q e^{-\theta(T-t_m)}}{T\theta} \tag{17}$$

Similarly, since supply chain profit function (17) is highly nonlinear, we solve it by using optimization technology on computer and get the optimal values.

4. Numerical Study

To illustrate the effect of the models, we give the following numerical example: $\alpha = 0.5$ units per time period, $\beta = 0.4$, $\theta = 0.1$, $S = \$10$, $M = \$20$, $p = \$20$ per unit, $q = 200$ units, $h^r = 0.35$, $h^m = 0.25$, $c^r = 3.5$, $c^m = 2.0$.

The models were implemented on a personal computer using Mathematica version 5.2. The results are outlined in Table 1, revealing that the profit for centralized supply chain is greater than that for decentralized supply chain.

Based on the numerical example considered above, we now perform the sensitivity analysis on the effects of changes in the model parameters such as the rate of deterioration, the retailer's purchase cost, the manufacturer's production cost, the retailer's and manufacturer's holding cost on the optimal order quantity for the retailer, the optimal replenishment cycle (or production cycle) for the retailer(manufacturer), the optimal supply chain profit, and supply chain profit increase percentages generated by the centralized policy. The effects of changes in the pa rameters values are shown in Table 2. The sensitivity analysis is performed by changing each of the parameters by –50%, -20%, +30% and +50% and keeping the other parameters unchanged. The results are demonstrated in

Table 1. The Sollution Results

Supply chain	Q^*	T^*	Π^r	Π^m	Π^{sc}
Decentralized	12.15	2.36	5.03	3.42	8.45
Centralized	20.63	2.75	4.76	4.56	9.32

Table 2. Effect of Changes in the Model Parameters

Parameters	Change (%)	Results of Optimization Procedure						
		Decentralized Supply Chain			Centralized Supply Chain			Profit Increase (%)
		Q^*	T^*	$\prod_1^{SC*}$	Q^*	T^*	$\prod_2^{SC*}$	
θ	+50	10.434	2.342	6.345	18.675	2.735	7.243	14.153
	+20	11.876	2.351	7.421	19.538	2.747	8.256	11.252
	-20	13.216	2.375	9.236	21.562	2.769	9.768	5.760
	-50	13.987	2.387	10.167	24.356	2.778	10.245	0.767
c^r	+50	11.579	1.987	17.948	17.917	2.294	18.831	4.920
	+20	11.682	2.173	14.656	19.267	2.473	15.857	8.195
	-20	15.876	2.563	7.063	30.013	2.931	7.771	10.024
	-50	16.786	3.126	4.754	38.681	3.477	5.335	12.221
c^m	+50	10.342	2.031	7.346	18.625	2.334	7.658	4.247
	+20	11.769	2.215	8.085	19.890	2.513	8.779	8.584
	-20	16.987	2.685	11.886	31.325	2.973	13.256	11.526
	-50	17.769	3.264	15.538	39.706	3.517	17.543	12.904
h^r	+50	10.876	2.046	7.258	17.018	2.305	7.589	4.560
	+20	11.986	2.263	7.982	18.845	2.433	8.648	8.344
	-20	16.765	2.765	12.189	32.098	3.384	13.767	12.946
	-50	17.875	3.078	15.876	40.022	3.441	18.089	13.939
h^m	+50	11.054	2.178	7.458	17.357	2.338	7.915	6.128
	+20	12.035	2.268	8.212	18.930	2.485	8.971	9.243
	-20	16.497	2.589	11.687	32.021	3.264	12.893	10.319
	-50	17.568	2.987	14.789	38.357	3.377	16.606	12.286

table 2. The following observations can be made from

(1) Whether the supply chain is centralized or decentralized, the optimal order quantity, the optimal replenishment cycle (or production cycle) , the optimal supply chain profit are decreasing in the deterioration rate. Meanwhile the percentages of supply chain profit increase realized by employing the centralized policy are increasing in the deterioration rates.

(2) Whether the supply chain is centralized or decentralized, the optimal order quantity and replenishment cycle (or production cycle) decrease, while the optimal supply chain profit increases in retailer's unit procurement cost.

(3) Whether the supply chain is centralized or decentralized, the optimal order quantity, the optimal replenishment cycle (or production cycle) and supply chain profit are decreasing in manufacturer's unit manufacturing cost.

(4) Whether the supply chain is centralized or decentralized, the optimal order quantity, the optimal replenishment cycle (or production cycle) and supply chain profit are decreasing in retailer's and manufacturer's inventory holding cost rate.

(5) The percentages of supply chain profit increase realized by employing the centralized policy are increasing in the deterioration rates, but are decreasing in the retailer's purchase cost, the manufacturer's production cost, the retailer's and manufacturer's holding cost.

5. Conclusions

This paper has investigated the effect of supply chain coordination for deteriorating goods with stock-dependent demand rate. Two profit models are developed with some assumptions based on exponentially decaying deterioration rates. The numerical study is conducted to demonstrate the effectiveness of the proposed models, and to analyze the impact of supply chain coordination on supply chain profit. Sensitivity analysis is performed to study the impact of different parameters associated with the model, such as the rate of deterioration, the retailer's purchase cost, the manufacturer's production cost, the retailer's and manufacturer's holding cost on the optimal order quantity for the retailer, the optimal replenishment cycle (or pro-

duction cycle) for the retailer(manufacturer), the optimal supply chain profit, and supply chain profit increase percentages generated by the supply chain coordination.

The following observations can be obtained from numerical analysis:

(1) A centralized policy is found to be always superior to a decentralized policy in terms of profit increase, especially when the deterioration rates are high.

(2)Whether the supply chain is centralized or decentralized, the optimal order quantity , the optimal replenishment cycle (or production cycle) , the optimal supply chain profit are decreasing in the deterioration rate, retailer's unit procurement cost, manufacturer's unit manufacturing cost, as well as retailer's and manufacturer's inventory holding cost rate. The optimal supply chain profit are decreasing in the deterioration rate, manufacturer's unit manufacturing cost, as well as retailer's and manufacturer's inventory holding cost rate. Meanwhile it is increasing in retailer's unit procurement cost.

(3)The percentages of supply chain profit increase realized by employing the centralized policy are increasing in the deterioration rates.

The proposed models can be used to analyze gricultural products, fashion goods, drugs and high-tech products supply chain. It is observed that large quantities of these deteriorationg goods displayed in a supermarket tend to generate higher demands. However, the models considered in this paper are somewhat idealized. In reality, when supply chain is coordinated, there are likely to be some costs incurred from information sharing scheme. The information sharing could also affect production, inventory, and other operations. The future research will further consider these factors. These models can be extended in the future research to consider more general deterioration rates. Another extension possibility would be to use other replenishment policies. In addition to profit models, the cost models can also be applied to evaluate the extent of coordination.

6. Acknowledgement

This paper is partially supported by Shanghai Shuguang Program under grant no. 06SG48.

REFERENCES

[1] Chen, J. M., Chen T. H. Effects of joint replenishment and channel coordination for managing multiple deteriorating products in a supply chain, *Journal of the operational research society*, 56，2005, pp.1224-1234.

[2] Ghare P. N.， Schrader GF. A model for exponentially decaying inventories, *Journal of Industry Engineering*, 15, 1963, pp.238-243.

[3] Covert R. P., Philip GC. An EOQ model for items with Weibull distribution deterioration, *AIIE Transactions*, 5, 1973, pp. 323-326.

[4] Philip G. C. A generalized EOQ model for itmes with Weibull distribution deterioration, *AIIE Transactions*, 6, 1974, pp.159-162.

[5] Tadikamalla，P. R. An EOQ inventory model for items with Gamma distributed deterioration. *AIIE Transactions*, 10, 1978, pp.78.

[6] Moon I., Lee S. The effects of inflation and time value of money on an economic order quantity model with a random product life cycle, *European Journal of Operational Research*, 125, 2000, pp.588-601.

[7] Dave, U., Patel L. K. (T, S_i)policy inventory model for deterioration items with time proportional demand. *Journal of the operational research society,* 32, 1981, pp.137-142.

[8] Sachan, R. S.On (T, S_i)policy inventory model deterioration items with time proportional demand. *Journal of the operational research society,* 35, 1984, pp.1013-1019.

[9] Bahari-Kashani H., Replenishment schedule for deteriorating items with time-proportional demand. *Journal of Operational Research Society*, 40, 1989, pp.75-81.

[10] Bose, S., Goswami, A., Chaudhuri,K.S., An EOQ model for deteriorating items with linear-dependent demand rate and shortages under inflation and time discounting. *Journal of the Operational Research Society*, 46, 1995, pp.771-782.

[11] Silver,E. A., Peterson,R., *Decision systems for inventory management and production planning*, 2nd edition, Wiley, New York, 1985.

[12] Gupta, R., Vrat, P., Inventory model for stock-dependent consumption rate. *Opsearch*, 23, 1986, pp.19-24.

[13] Mandal, B. N., Phaujdar, S., A note on an inventory model with stock-dependent consumption rate. *Opsearch*, 26, 1989a, pp.43-46.

[14] Baker, R. C. and Urban,T. L., A deterministic inventory system with an inventory-level-dependent demand rate. *Journal of Operational Research Society*, 39(9), 1988, pp.823-831.

[15] Datta, T. K., Pal, A. K., Effects of inflation and time-value of money on an inventory model with linear time-dependent demand rate and shortages. *European Journal of Operational Research*, 52, 1991, pp.1-8.

[16] Mandal, B. N., Phaujdar, S., An inventory model for deteriorating items and stock-dependent consumption rate. *Journal of Operational Research Society,* 40(5), 1989b., pp.483-488.

[17] Pal, S., Goswami, A., Chaudhuri, K.S., A deterministics inventory model for deteriorating items with stock-dependent demand rate. *International Journal of Production Economics*, 32, 1993, pp.291-299.

[18] Spiegel, M. R., *Applied Differential Equations*. Prentice-Hall: Englewood Cliffs, NJ, 1960.

J. Serv. Sci. & Management, **2008, 1: 128-134**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

Real-Time Resource Availability Signaling in IP Multimedia Subsystem Networks

Tanır Özçelebi[1], Igor Radovanović[1] & Debraj Sengupta[2]

[1]*Eindhoven University of Technology, PO Box 213, 5600 MB, Eindhoven, The Netherlands*
[2]*ICT Embedded BV. PO Box 6420, 5600 HK, Eindhoven, The Netherlands*
Email: [1]{t.ozcelebi, i.radovanovic}@tue.nl, [2]sengupta.debraj@gmail.com

ABSTRACT

IP Multimedia Subsystem (IMS) allows the use of unlicensed, non-dedicated and nondeterministic access networks for delivering IP multimedia services. Providing end-to-end Quality-of-Service (QoS) for resource demanding real-time services (e.g. real-time multimedia) over such networks is a challenging task due to varying resource availability of the network and the end-device[i] during a single session[ii]. A common solution to this problem is adapting service quality and type according to availability of system resources, which requires end-to-end monitoring and signaling of resource availability during a single session. This paper presents an extension of the IMS architecture for real-time resource availability monitoring and signaling. The novelty of the proposed extension is twofold. It takes into account uncontrolled access networks with no resource reservation on the one hand, and end devices on the other. Two system architecture components are introduced for monitoring and signaling of the real-time resource availability in both networks (e.g. bandwidth, buffer space), and end-devices (e.g. battery, CPU, memory, storage), namely a Resource Manager (RM) and a Resource Availability Server (RAS). Our solution does not require any modifications of the IMS architecture and can be implemented as a plug-in.

Keywords: *end-to-end quality of service, IP multimedia subsystem, resource management, session initiation protocol*

1. Introduction

Today's smart phones and personal digital assistants (PDA) with wireless local area network (WLAN) connectivity are capable of accessing voice, video and data services outdoors over the Internet. Unlike the general practice in traditional mobile telecom world, the quality of these services is not always guaranteed and users can only get best-effort service. Nevertheless, such services have recently attracted a lot of interest from the mobile telecom community and upgrading from plain voice services to data and real-time multimedia services was necessary. Such services are currently being enabled by the emerging 3rd Generation (3G) mobile phone technologies.

Mobile telecom operators used to own not only the usage license of their wireless frequency band but also the network infrastructure in which their mobile services are available, including access points. This has several advantages in terms of QoS. Licensing of the wireless frequency band assures that the wireless link is free of external signal pollution, i.e. non-telecom signals can not use the specific frequency band. Ownership of the network infrastructure allows the operator to monitor availability of network resources and enforce QoS mechanisms such as admission control (AC) [1], resource reservation (RSVP) [2] and traffic engineering [3] for their services. Using this approach, it is possible achieve the best service quality that meets the real-time system resource constraints.

Traditionally, all service types offered for certain devices were assumed to be available disregarding the availability of end-device and network resources. Moreover, mobile communications domain is currently migrating towards an All-IP network, which is accessible not only via mobile telecom access points but also via any IP access point. In the converged Next Generation Networking (NGN) framework [4] introduced recently, the functionality of services is independent from the underlying network, where managed and unmanaged IP networks are used together.

Note that QoS enforcement mechanisms used in mobile networks are no longer valid for NGN. For instance, applying AC and RSVP on NGN sessions in an unlicensed access network (e.g. a computer network) would not prevent QoS degradation at peek hours due to capacity overload and interference from non-NGN flows. The IMS [5] is an NGN architecture that was standardized by the 3rd Generation Partnership Project (3GPP) group [6], where the Session Initiation Protocol (SIP) [7] is employed at the application-layer as a control safeguard integrating

IMS with the Internet. According to IMS specifications, the users should be able to access services anytime, from anyplace and using any SIP-enabled device with IP version 6 (IPv6) [8].

Therefore, it is currently envisioned that the operators must provide services to users even if they are in access networks that are completely out of their control, hence the need for new end-to-end QoS mechanisms. This new approach is shown in Figrue. 1. Generally, the term "end-to-end" refers to the connection from the user-agent or the proxy to the server in telecom networks. According to this definition, the path from one end to the other lies within the core network owned and controlled by the operator. On the other hand, in this paper, the term "end-to-end" is used to refer to the connection from one user agent to the other, passing through the access networks and the core network. Thus, the access networks on the path constitute the variable-resource bottlenecks of the end-to-end network.

Guaranteeing network resource availability for resource demanding services is not feasible even if all parts of the communication path are operator-owned. This is because access networks are mostly nondeterministic, e.g. WLAN and Ethernet. The resource availability in the end-devices may also become the bottleneck in multimedia communication and it becomes impossible to *guarantee* end-to-end resource availability and consequently, service quality [5].

For maximizing user satisfaction, QoS needs to be adapted according to system resource availability. As resource availability in the system varies, both in networks and end devices, it has to be signaled to an end-device or an Application Server (AS) *during a single session*. The current IMS architecture does not provide specifics of such on-the-fly QoS negotiation. This work provides the details of such resource availability signaling for IMS, without altering the existing IMS architecture. The proposed system requires monitoring of available network and end-device resources, hence the need for Resource Managers (RM) at the communicating devices.

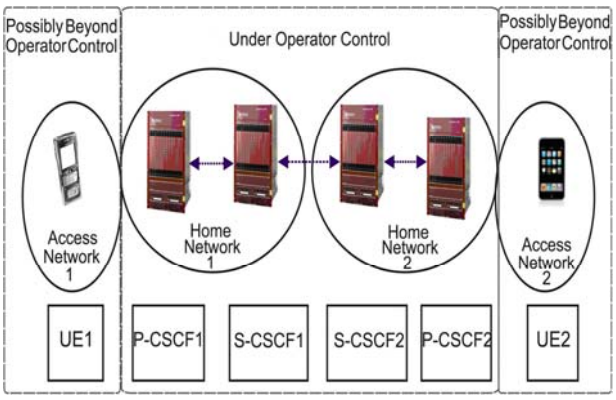


Figure 1. The end-to-end IMS network including access networks

The rest of this paper is organized as follows: The existing QoS mechanisms in IMS and related work in the literature are explained in Section 2. Our proposed solution with resource management and QoS adaptation is described in Section 3. The experimental results are given in Section 4. Finally, conclusions are drawn in Section 5.

2. Existing QoS Signaling Mechanisms for IMS

The IMS layered architecture is divided into three planes with different functionalities [5]. The transport and connectivity layer is separated from the application and service layer, by means of the signaling and control layer, whose job is to carry out call session control. In this plane, there is a standard set of control mechanisms valid for all services provisioned. Ideally, if there is a policy agreement among the networks that lie along the end-to-end path, the serving operator(s) can use these control mechanisms in order to improve the end-to-end QoS of the provisioned services. Due to such agreement, the operator would be aware of the available resources in these networks and the networks would apply the QoS decisions of the operator regarding call admission and resource reservation.

However, this scenario has a drawback. Even if such an agreement is present, a QoS policy translation problem across different networks still exists as the access networks outside the CN may have their own QoS models and semantics [9]. For example, there exist four QoS classes in UMTS framework (i.e. conversational, stream

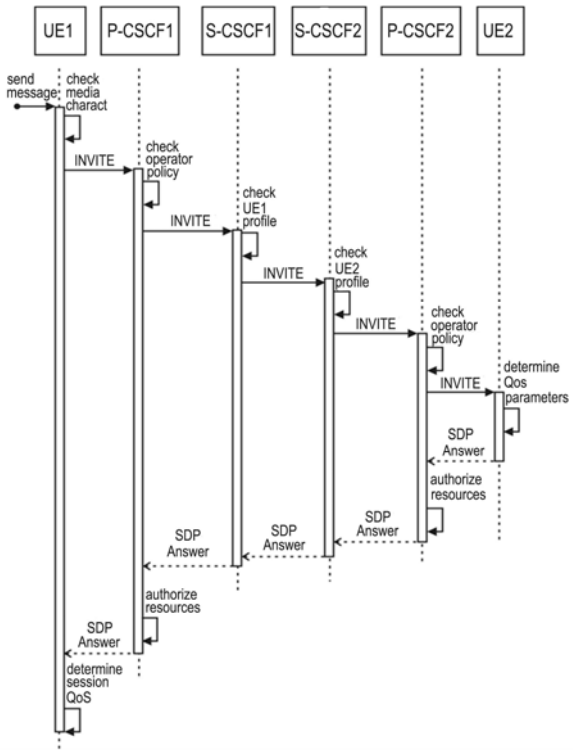


Figure 2. M**essage Sequence Diagram showing pre-session QoS negotiation in an IMS network**

ing, interactive, background) [10], two QoS classes in Differentiated Services (i.e. expedited forwarding, assured forwarding) [11], and three QoS classes in Integrated Services (i.e. guaranteed, controlled-load, best-effort) [12].

In the IMS standard, QoS parameters can be negotiated between two user equipments prior to the session establishment using SIP. Once the QoS parameters have been negotiated between user devices (checked against fixed device capabilities) and been approved/modified by the Call Session Control Functions (CSCF) (checked against user subscription credentials) associated with both users, the IMS network asks the CN and the access network to reserve resources for this session. The SIP INVITE and SIP UPDATE messages [7] are used for this purpose as shown in Figrue. 2. The first INVITE message from the caller user equipment (UE1) to the call receiver (UE2) carries the QoS proposal, which is checked against the subscription levels of users at the Serving-Call Session Control Functions (S-CSCF) in both of the home networks. The QoS parameters are modified at these locations if there is a mismatch.UE2 puts her own QoS proposal in the answer and this is again checked and modified at the associated S-CSCF's of the users. Finally, UE1 can accept this counter QoS proposal and start the session or try to renegotiate with a SIP UPDATE message. Within the body of these SIP messages the session data is passed using Session Description Protocol (SDP) [13].

Maniatis *et al*. [9] have tried to tackle the QoS model translation problem across different networks by an intelligent mapping algorithm for end-to-end QoS negotiation such that the best suiting QoS class is selected in each network along the session path. Similarly, the operator can employ AC by using the Policy Decision Function (PDF) at the access network border and by making the access network physical bearer enforce a Service Based Local Policy (SBLP), assuming the bearer is listening to the commands from the IMS signaling and control layer. Even inter-network policy agreement would not be enough for access networks with unlicensed frequency band and anonymous usage rights, since already existing IMS sessions can still be jeopardized by new non-IMS services initialized in the access network. Furthermore, the access network can be nondeterministic and non-dedicated with interferences from the outer world, which is the case for Ethernet and WiFi. In this case, IMS sessions are bound to suffer unless the session QoS parameters are modified to fit the resource availability when there is a shortage or a boost.

There are several RFCs published by the Internet Engineering Task Force (IETF) that propose resource/ capability signaling among end-devices. These signaling schemes mainly concentrate on multimedia services, whereas our method can be applied to any real-time service over IMS. In [14], a method that integrates resource management (specifically RSVP) and SIP signaling is introduced in order to make network resource reservation before the session is established, i.e. before the called end-device is alerted such that session establishment failures are avoided. However, this RFC proposes no signaling during a single session. In our proposed solution it is assumed that reservations may not be possible in the access networks, and therefore the relation with reservation protocols, like RSVP, is not considered in this paper. Furthermore, signaling is done during a single session. Another difference is that, we propose an architecture in which local resources of the end-devices can also be transmitted to other interested parties, whereas [14] gives information about network resources only.

In [15], Internet Media Guides (IMG), i.e. multimedia session descriptions that can use SDP format are introduced. However, it is denoted in [15] that SDP syntax causes a huge amount of overhead in delivering IMG metadata over the network and SDP can carry only a small subset of IMG metadata in practical cases (e.g. codec type).

The bandwidth modifier of [16] notifies the receiving end-device on the maximum media codec rate to be used and the communication bit-rate required for the bit stream. Thus, [16] aims to convey bit-rate information only, without conveying any information about end-device resource availability.

In [17], bandwidth modifiers for RTP Control Protocol (RTCP) are introduced to SDP such that the amount of bandwidth allocated to RTCP in an RTP session is adapted (typically kept below 5% of the overall data rate). We envision that SIP resource availability signaling is preferable for protecting the privacy of resource availability data compared to transport layer protocols (e.g. RTCP), which lack to provide means for authentication, encryption and billing.

An extended SDP protocol for capability declaration (e.g. codec) amongst end-devices to be used in multimedia sessions is introduced in [18]. It is declared that such capability declarations can be intended for session negotiation, but such session negotiation mechanisms are not described.

In the IMS, it is envisioned that the end-to-end QoS negotiation and resource allocation should be reevaluated during the session depending on requests from the application, network load and link quality [5]. On the other hand, the implementation specifics of such a *QoS renegotiation* mechanism are not provided.

3. Resource Availability Signaling During A Session

As explained in the previous section, IMS allows communication sessions for which session QoS guarantees cannot be given [5]. For example, if the user is connected through a non-dedicated access network, e.g. Ethernet or

WiFi, it may not be possible to guarantee the requested end-to-end QoS due to scarcity of both end-device and network resources deteriorating the user experience. The deterioration effect is especially apparent while using resource demanding services such as multimedia streaming. The best approach to solve this problem is to introduce service quality adaptation (SQA) according to system resource availability (local, remote and network). The design constraint for the solution is to avoid possible alterations to the IMS architecture.

In order to perform such adaptation, the decision units for service quality adaptation at the end-devices, i.e. Service Quality Management (SQM) modules, must be aware of the availability of local and remote resources (e.g. battery, memory, storage CPU etc.) and network resources (e.g. throughput, buffer space) during the session and adapt their service quality level/type accordingly, hence the need for *resource availability monitoring* and *signaling*. In this section, we introduce i) the Resource Management (RM) module, ii) a Resource Availability Server (RAS) as an application server, and iii) a resource availability signaling mechanism for real-time adaptation of multimedia communication streaming and data streaming (e.g. video-on-demand) services in the IMS network. The operation of the SQM modules is beyond the scope of this paper.

The RM module is a crucial part of the proposed resource availability signaling framework. It is responsible for tracking local and network resources available to the end-user device in real-time. At the receiving device, the RM module publishes this information to the RAS server to be accessed by the remote transmitting device. At the transmitting device, the RM is responsible for gathering the resource availability data of the remote receiving device from the RAS.

It is the proposed RAS server that is responsible for collecting resource availability information from the receiving end-devices and delivering it to the transmitting end-device. Note that in a multimedia communication scenario, e.g. video-conferencing, an end-device can be transmitting multimedia, receiving multimedia, or both. In order to be IMS compliant, it is appropriate to employ the existing SIP call session control protocol of IMS for resource availability signaling in the proposed architecture. In Figrue. 3, resource availability signaling is shown, where it is assumed that user equipments UE1 and UE2 have registered to each other's resource availability information at the RAS server.

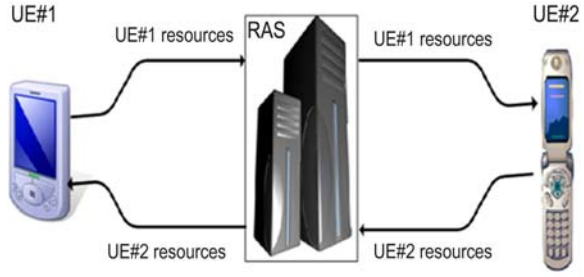


Figure 3. The proposed system architecture for mid-session SIP-based resource availability signaling

Resource availability data is carried from end-devices to RAS and back in SIP event notification messages, such that a resource update is signaled whenever the local resources (e.g. memory, CPU, storage etc.) or network resources (e.g. bandwidth, jitter etc.) at one end crosses a critical boundary threshold. It is beyond the scope of this paper to propose methods for determining those thresholds. For example, suppose that a given real-time service offers three ways in which it can operate: multimedia, audio only and text. The communicating parties can select either one of these modes based on their resource availability. For instance, they can choose text when 90% or more of the times the CPU is occupied with other tasks in one of the end-devices, they can chose multimedia when the CPU is occupancy is less than 10%, or they can choose audio only in between. Therefore, the amount of end-device and network resources spent on resource monitoring and signaling is negligible in the proposed framework. A worst case scenario is investigated in the next section.

The proposed message flow diagram from the end-device to RAS for resource availability signaling is depicted in Figure 4 and our additions to the SIP/SDP pa-

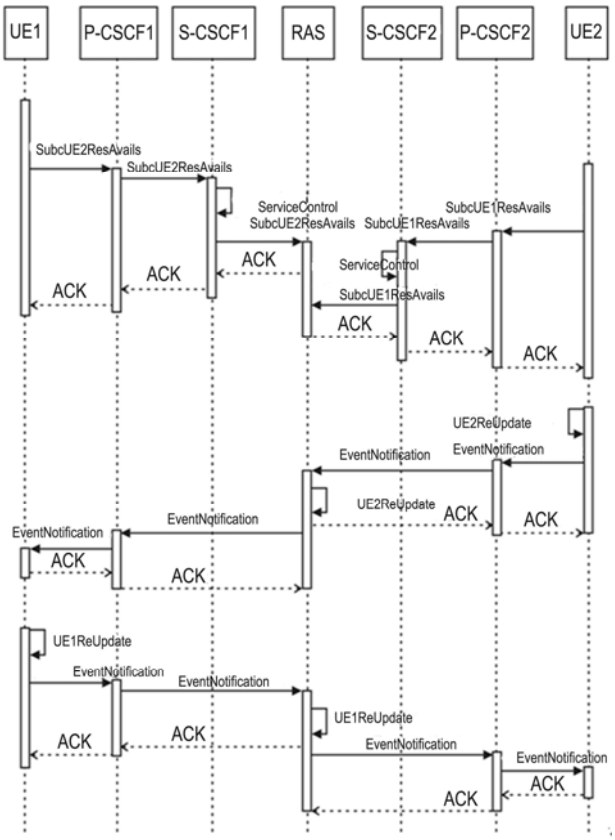


Figure 4. Message sequence diagram showing SIP based resource availability update signaling

Table 1. Proposed additional resource data in SDP

r ::= “memory”	“CPU”	“storage”	“throughput”	“battery”
t ::= “Mbytes”	“Kbytes”	“kbps”	“seconds”	“percentage”
a ::= <resource availability measure>

Table 2. Example SDP with resource data

```
NOTIFY sip:abc.somename.com SIP/2.0
Via: SIP/2.0/UDP abc.tue.nl:5060;branch=z9hG4bKnashds7
Max-Forwards: 70
To: Bob <sip:server.somename.com>
From: Bob <sip:abc.somename.com>;tag=456248
Call-ID: 843817637684230@998sdasdh09
CSeq: 1826 NOTIFY
Contact: <sip:abc@192.0.2.4>
Expires: 7200
Content-Type: application/sdp
Content-Length: 131

v= RM/1.45
o= abc 5876768686868 7698798797979 IP 1.2.3.4
s= 123456789
i= Resource update to presence server
c= IN IP4 1.2.3.4
b= 100 kbps
k= none

r= memory
i= free memory status
t= kbytes
a= 12450

r= battery
i= battery charge remaining
t= percentage
a= 86
   --msg ends--
```

rameters as resource indicators are shown in Table 1.

After the addition of the proposed resource availability parameters, an example resource update SDP message looks as shown in Table 2. Here, both users will be aware of each other's local and network resources and SQM can use this information to perform its functionality.

4. Experimental Results

In our experiments, the RAS server (i.e. a PC) and the RM modules at both end-devices (i.e. two PDA's with WinCE operating system) communicate with each other during multimedia streaming between the two PDA's. The PDA applications start communicating using SIP INVITE. The media flow is started after the ACK is received from the caller. The session ends with a SIP BYE message which terminates the media session. Resource availability data from each PDA is reported to the RAS module using our SIP NOTIFY messages with the new header fields introduced in Table 1.

The multimedia session between PDA's is carried over Real-time Transport Protocol (RTP) and the PDA's are

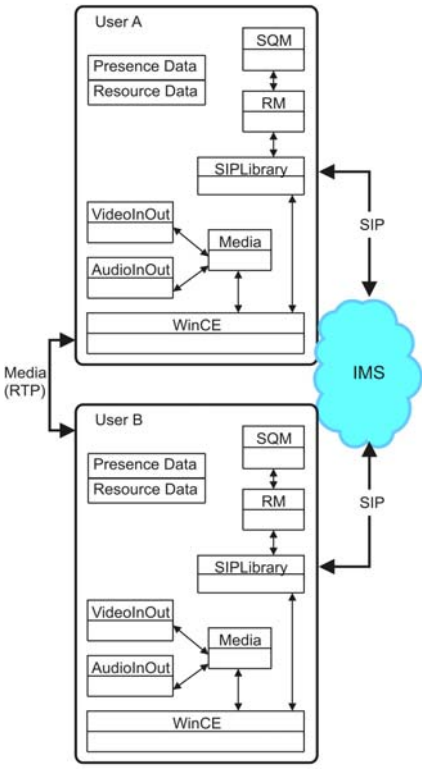


Figure 5. Deployment view of the user devices and IMS core

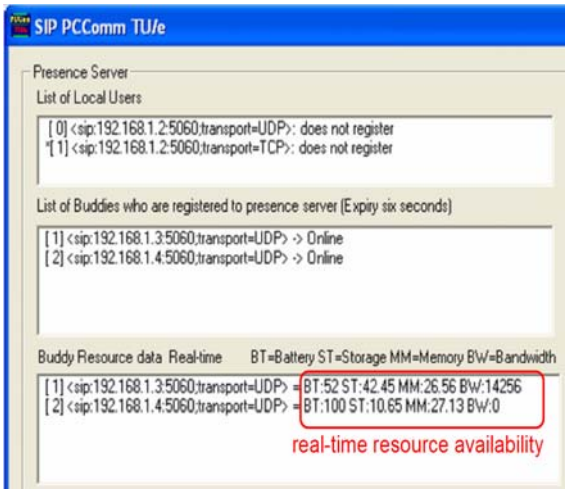


Figure 6. A snapshot of the RAS log interface

subscribed to each other's resource information at the

RAS server when a call takes place. The real-time resource availability signaling is done in parallel with the session using SIP/SDP as shown in Figure 5.
Figure 7 shows the PDA User Interface (UI) of the client test application. The top left menu is used to make a call or to exit the application. The first line shows the local IP address and port. The local and the remote resource availability data is displayed in the second line and the last line respectively.

In our experiments, we assumed that the real-time service that is going to make use of the proposed architect-

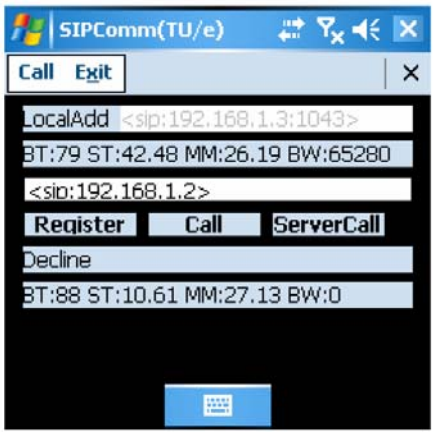


Figure7. A snapshot of a user interface of the PDA

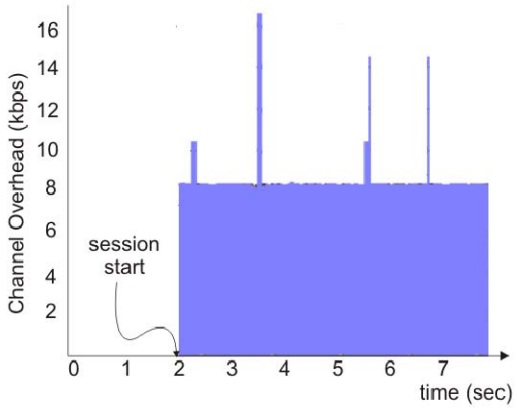


Figure 8. Resource availability signaling channel overhead in the worst case scenario (signaling period: 1 sec.)

ture is a multimedia service. Multimedia services are the driving services that are offered by the IMS. In order to enable those services, availability of considerable amount of resources in the IMS network and end-devices is needed during a single session. Since the amount of resources available varies, quality of the provided services drops. One solution to this problem is introducing quality adaptation based on real-time resource availability, which can be achieved by multimedia bit rate adaptation.

However, such adaptation cannot be done arbitrarily, since multimedia codecs have their own limitations in changing the encoding bit rate on-the-fly even in the case of scalable codecs or bit stream switching. Therefore, in multimedia service quality adaptation, adaptation speed should not be higher than that of the multimedia codec. In [19], it is argued that the video adaptation algorithms in the literature need up to 3 groups of pictures (GOP) in order to converge to a target bit rate every time the video is adapted. Here a GOP is defined as a frame sequence of a given structure in a video stream, whose first frame is an intra-coded (I) frame. Furthermore, it is also denoted in [19] that the size of a GOP has to be kept large in an encoded video bit stream in order to attain reasonable compression efficiency and 1 GOP per second is taken as a rule of thumb, which would allow 1 adaptation in every 3 seconds for the other rate controllers in the literature and 1 adaptation per second for the advanced rate controller of [19]. Therefore, we assume that the maximum video adaptation frequency is 1 adaptation per second for typical videos. Updating resource availability at the speed higher than the adaptation speed would result in no quality improvements. Moreover, resource consumption in the network and end-devices will be higher.

The maximum overhead caused by the proposed architecture will be in case the resource availability is done at the maximum adaptation speed of the multimedia codec used (i.e. 1 adaptation per second). To analyze this, we used periodic resource availability update with a period of 1-second, for which the resource availability signaling overhead is measured to be 8 kbps as shown in Figrue. 8. The signaling overhead is zero before the session starts, and it increases to 8 kbps on average per session after the session is initiated.

In [20], the design and implementation of an Open IMS core has been done and the load on the IMS core network and the proxies due to SIP message flow is investigated. In an IMS network with 100,000 subscribers, 1/3 of the subscribers are assumed to be online and registered simultaneously at a typical instant and 1/16 of these subscribers are assumed to be engaged in a multimedia session with an average duration of 180 seconds. In this case, the system would have to support 11,57 calls per second and the Open IMS core would have to process around 81 SIP messages per second (7 SIP messages for each multimedia call setup). In their experimental results, it was shown that a simple Intel Pentium 4 processor running at 3GHz (HyperThreaded) is enough to do the tasks of all IMS core components at once (i.e. I-CSCF, S-CSCF, P-CSCF and HSS) and still handle 120 SIP messages per second (around 17 calls per second). Considering the above data, in a worst case scenario of the proposed architecture, i.e. when each and every one of the active users has to adapt their multimedia within a given second, around 2000 SIP messaging events would need to be handled by the IMS core. This is quite realizable in a real-life deployment of the IMS core network since i) all components (CSCF's and HSS's) of the system normally reside on different hardware nodes in a deployed IMS core, ii) using multiple instances of the same component (e.g. multiple S-CSCF's) is very common for load-balancing, and iii) the state-of-the-art processors of today (e.g. multi-core processors) are much more powerful than a 3GHz Intel Pentium 4 processor. Clearly, in a more realistic case, the resource availability signaling overhead decreases even further when the resource signaling is done based on critical thresholds as described in the previous section.

The RAS is an additional server unit that can be implemented as an AS and it is independent of the IMS CSCF. Therefore, the existence of RAS does not put any

computational overhead on the CSCF's.

5. Conclusion

The emerging IMS framework allows mobile users to benefit from highly resource consuming services (e.g. multimedia streaming) at anytime and anyplace. In contrary to the traditional practice in telecom networks, the IMS access networks can be non-dedicated and/or nondeterministic, and the availability of the network and the end-device resources may show severe oscillations within a single session. The currently available mechanisms in the IMS standard are insufficient to provide guaranteed QoS. Service quality adaptation is needed in order to enhance user experience, which can be realized using resource availability monitoring and SIP-based end-to-end resource availability signaling as proposed in this paper.

The main contribution of this paper is an extended IMS architecture for real-time resource availability monitoring and signaling. The novelty of the proposed architecture comes from the fact that it takes into consideration uncontrolled IMS access networks with no resource reservation on the one hand, and end-devices with limited resources on the other. The available resources of a user's end-device are collected at the RM and advertised to the users that are subscribed to her resource status at the RAS server through SIP signaling in real-time.

Experiments show that resource availability signaling introduces negligible overhead to the overall network traffic even in a worst case scenario for real-time multimedia services.

6. Acknowledgement

This work has been supported by Converged Messaging Technology (COMET) consortium funded by the European Commission 6th Framework (FP6) Programme.

REFERENCES

[1] R. Yavatkar, D. Pendarakis and R. Guerin, "A Framework for Policy-based Admission Control," *RFC* 2753, IETF Network Working Group, http://www.ietf.org/rfc/rfc2753.txt, January 2000.

[2] R. Braden, L. Zhang, S. Berson, S. Herzog and S. Jamin, "Resource ReSerVation Protocol (RSVP), " *RFC* 2205, IETF Network Working Group, http://www.ietf.org/rfc/rfc2205.txt, September 1997.

[3] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja and X. Xiao, "Overview and Principles of Internet Traffic Engineering," *RFC* 3272, IETF Network Working Group, http://www.ietf.org /rfc/rfc3272.txt, May 2002.

[4] http://www.itu.int/ITU-T/ngn/index.phtml.

[5] Digital Cellular Telecommunications System (Phase 2+), *Universal Mobile Telecommunications System (UMTS), IP Multimedia Subsystem (IMS)*, Stage 2, V7.6.0, TS 23.228, 3GPP, December 2006.

[6] http://www.3gpp.org.

[7] J. Rosenberg, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley and E. Schooler, "SIP: Session Initiation Protocol," *RFC* 3261, IETF Network Working Group, http://www.ietf.org/rfc/rfc3261.txt, June 2002.

[8] http://www.comet-consortium.org.

[9] S.I. Maniatis, E.G. Nikolouzou and I.S. Venieris, "End-to-End QoS Specification Issues in the Converged All-IP Wired and Wireless Environment, " *IEEE Communications Magazine*, vol. 42, no. 6, 2004, pp. 80-86.

[10] S. Chia,"The Universal Mobile Telecommunication System, " *IEEE Communications Magazine*, vol. 30, no. 12, 1992,pp. 54-62.

[11] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, "An Architecture for Differentiated Services, " RFC 2475, IETF DiffServ Working Group, http://www.ietf.org /rfc/rfc2475.txt, December 1998.

[12] R. Braden, D. Clark and S. Shenker, "Integrated Services in the Internet Architecture: an Overview, "RFC 1633, IETF Network Working Group, http://www.ietf.org/rfc/rfc1633.txt, June 1994.

[13] M. Handley and V. Jacobson, "SDP: Session Description Protocol," *RFC* 2327, IETF Network Working Group, http://www.ietf.org/rfc/rfc2327.txt, April 1998.

[14] G. Camarillo, W. Marshall and J. Rosenberg, "Integration of Resource Management and Session Initiation Protocol (SIP)", *RFC* 3312, October 2002.

[15] Y. Nomura, R. Walsh, J-P. Luoma, H. Asaeda and H. Schulzrinne, "A Framework for the Usage of Internet Media Guides (IMGs)", *RFC* 4435, April 2006.

[16] M.Westerlund, "A Transport Independent Bandwidth Modifier for the Session Description Protocol (SDP)", *RFC* 3890, September 2004.

[17] S.Casner, "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", *RFC* 3556, July 2003.

[18] F. Andreasen, "Session Description Protocol (SDP) Simple Capability Declaration", *RFC* 3407, October 2002.

[19] F. De Vito, T. Ozcelebi, O. Sunay, M. Tekalp, R. Civanlar and J. C. De Martin, "Per-GOP Bitrate Adaptation for H.264 Compressed Video Sequences," in L. Atzori *et al*. (Eds.): *LNCS* 3893, pp. 198-206, Springer-Verlag Berlin Heidelberg 2006.

[20] D. Vingarzan, P. Weik, T. Magedanz, "Design and Implementation of an Open IMS Core," in T. Magedanz *et al*. (Eds.) *LNCS* 3744, pp. 284-293, Springer-Verlag Berlin Heidelberg 2005.

[i] In this paper, the terms "end-device", "user-device" and "user-equipment" are used interchangeably.
[ii] A session is a durable connection between two user devices or a user device and a server.

J. Serv. Sci. & Management, **2008, 1: 135-142**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

Empirical Study on the Performance of Initial Public Offerings in China

Liang Peng

Graduate School of Economics and Business Administration, Hokkaido University, Japan, 060-0809
School of Management, Central South University, Changsha, P.R.China, 410083
E-mail:liang261@pop.econ.hokudai.ac.jp

ABSTRACT

We study the long-run performance of 166 IPOs listed on China's Shanghai Stock Exchanges from 2000 to 2002. We find that the average market-adjusted cumulative return and buy-and-hold return over the three years after listing are -32.02% and -20.88%, which are both significantly negative. What's more, as an additional robustness check, we calculate wealth relatives. One year after listing, WR less than 1 and we obtain a three-year wealth relative of 0.6826, consistent with the CAR and BHAR estimates. We then use a cross-sectional analysis to explain the long-run underperformance of Chinese IPOs. The results show that the aftermarket performance is positive after listing (6 months) but thereafter returns decline. Buying A-share IPOs immediately after listing and holding the investment for three years results in negative returns and wealth relatives less than one.

Keywords: *Initial Public Offering, Long-run Performance, Market Efficiency*

1. Introduction

Since Ritter's [19] has presented convincing empirical evidence that IPOs underperform in the long-run, a number of corresponding studies have sought to reveal the existence the sever aftermarket underperformance for issuers in different countries. Aggarwal [1] and Loughran *et al.* [11] examine the returns on IPOs during the three years after going public for a number of countries. They equally-weight the IPOs in their respective samples and both find underperformance. Comparable results have also been found for IPOs in economies under transition such as China. From table 1, the aftermarket underperformance is not unique to the US IPO market, it also exists in a number of other countries, like Canada, Japan and so on.

Chinese stock markets expanded rapidly following the opening of securities markets in Shanghai and Shenzhen in the early 1990s. Although until now they are only 18 years old, they enjoy very high growth. The importance and newness of the markets and unique institutional features make China a special environment to conduct research on IPOs and findings from studies in other markets cannot be extrapolated to China. Some of the previous studies have noted that Chinese IPOs enjoy the world's highest initial returns at around 200-300% [18] [4]. However, there is little research on the long-run performance of Chinese IPOs due to the data shortage and the current results on the long-run performance are mixed. Sun and Tong [20] look briefly at the long-run share returns (raw returns and Hong Kong Hang Seng Index adjusted returns) of IPOs and find stock returns show some mild improvements up to five years after share issue privatizations. Chan *et al.* [4] study 570 A-share IPOs and 39 B-share IPOs from 1993-98 and 1995-98 respectively and find that in the long-run (within three-year after listing) A-share IPOs slightly underperform the size- and/or book/market-matched portfolios while B-shares outperform the benchmark portfolios.

The contribution of this paper is to use more updated data and to present a deeper understanding of the special features in the Chinese IPO market to study the performance, especially the long-run performance computing by both the abnormal three-year cumulative returns and buy-and-hold returns of IPOs and use a cross-sectional analysis to explain the long-run underperformance of Chinese IPOs.

The rest of the paper is organized as follows: Section 2 introduces the features of China's stock markets; Section 3 presents the data and methodology for calculating the long-run returns; Section 4 provides the analysis of empirical results on the long-run returns. Section 5 examines several different hypotheses of the cross-sectional variance in abnormal returns. The summary and conclusion appear in Section 6.

2. Features of China's Stock Market

Following the economic reforms that began in 1978, the Chinese stock market was finally established in the early 1990s. The Shanghai Securities Exchange was opened in 1990, followed by the establishment of the Shenzhen

Table 1. International evidence on the aftermarket performance of IPOs

Country	Author(s)	Number of IPOs	Issuing Years	Aftermarket performance
Australia	Lee,Taylor& Walter(1996)	266	1976-89	–46.5%
Brazil	Aggarwal, Leal&Hernandez(1993)	62	1980-90	–47%
Canada	Kooli and Suret(2003)	445	1991-98	–16.86%
Chile	Aggarwal, Leal&Hernandez(1993	28	1982-90	–23.7%
Finland	Keloharju(1993)	79	1984-89	–21.1%
Germany	Ljungqvist(1997)	145	1970-90	–12.1%
Hong Kong	M cguiness (1993)	72	1980-90	–18.3%
Japan	Cai&Wei(1997)	172	1971-90	–27%
Korea	Kim, Krinsky&Lee(1995)	99	1985-88	2.00%
New Zealand	Firth(1997)	143	1979-87	–10%
Sweden	Loughran,Ritter&Rydqvist(1994)	162	1980-90	1.20%
United Kindom	Levis(1993)	712	1980-88	–8.1%
United States	Loughran&Ritter(1995)	4753	1970-90	–20%
United States	Simon (1989)	35	1926-33	–39%
United States	Ritter (1991)	1526	1975-84	–29.1%

Stock Exchange in 1991.Chinese government tries to use the market to develop the economy, while still keep some socialist characteristics, which means the government plays a crucial role in monitoring and regulating the stock markets.

In Chinese stock market, there are several classes of share, A-shares, B-shares, H-shares, N-shares and non-tradable shares. A-shares were tradable only by domestic investors (and quoted in Yuan) whilst B-shares were tradable only by foreign investors (and quoted in US Dollars), although recently domestic investors have also been able to trade B-shares. Hand N-shares are listed on the Hong Kong and New York Stock exchanges, respectively. Non-tradable shares include those owned by the state and other state owned enterprises (legal person shares), although China started to merge this dual shareholding system in May 2002.When going public, shares not retained by the government, other enterprises or employees are sold to outside investors. There are five types of shares in China: government shares, which are held by the State Assets Management Bureau (SAMB); legal entity shares (or C shares), which are held by other state-owned enterprises; employee shares, which are held by managers and employees; ordinary domestic individual shares (or A-shares), which can be purchased only by Chinese citizens of the PRC on the Shanghai Securities or the Shenzhen Stock Exchange; and foreign shares, which can be purchased only by foreign investors in Mainland China (B-share), in Hong Kong (H-share), or on the NYSE (N-share). Only the A-shares and B-shares are listed on the Shanghai Securities and Shenzhen Stock Exchanges. The first three types of shares are not tradable in the official exchanges, although employee shares are allowed to be listed three years after the IPO.

3. Data and Methodology

We study the long-run performance of 166 IPOs listed on China's Shanghai Stock exchanges from 2000 to 2002. The data mainly from the China Centre for Economic Research and www.sse.com.cn, while the financial data and transactional data of shares (mainly daily closing prices) come from the Wind Information database.

The following criteria are used in selecting the final sample: (1) issuing firms are listed in Shanghai Stock Exchange (2) stock price data for issuers, market capitalization are available on the Wind Information database (3) common-share IPOs are selected, and exclude units, close-end funds, and real estate investment trust offerings.

To analyze the after performance of Chinese IPOs, we apply the standard event study methodology. Thus abnormal returns of our sample are computed using the Cumulative Abnormal Returns (CAR) and Buy-and-Hold Abnormal Returns (BHAR). There continues to be disagreement regarding the measurement of long-run abnormal return performance. Barber and Lyon [2], Lyon *et al*. [17], Kothari and Warner [10] and Fama [5] analyze the difference of these two methods, and find that there is no consensus on the preferred one. Lyon *et al*. [17] document that BHARs should be used if the research question is whether or not investors earn abnormal stock returns by holding stocks over a particular time horizon. While the CAR approach should be employed to answer the following question: do sample firms persistently earn abnormal monthly returns? Though CARs implicitly assume frequent portfolio rebalancing, Fama [5] justifies its use since it would produce fewer spurious rejections of market efficiency than would the use of BHARs. There also exists a greater knowledge of the distribution properties and the statistical tests for CARs. Since in China, the ma-

jority of investors are individual investors, and they trade much more frequently than those in other markets, to guarantee the robustness of our results, we estimate the three-year abnormal return after IPO using both measures.

Research in the US has frequently used a matching firm approach to measure long-run abnormal returns. However, the small number of companies available means there would be a bias caused by the repeated use of matching companies. We therefore both measure three-year post-IPO abnormal returns relative to the Shanghai Stock Exchange A-share index, a capitalization weighted index of 166 companies and matching firms. So we build our conclusions on different methodologies, limited by the paper, we use the event time approach and use two methodologies within each approach: we first examine the cumulative abnormal returns (CARs) and Buy-and-hold returns (BHARs) with the shanghai stock exchange Index as the benchmark and then examine the cumulative abnormal returns (CARs) and Buy-and-hold returns (BHARs) with matching firms.[1]

We measure the monthly return to both the company r_{it} and the market over the three year period after the IPO. The market benchmark return r_{mt} is the return to the Shanghai Stock Exchange A-share index or the return of the matching firms. We follow Ritter (1991) [19] and exclude the initial return at the time of the IPO. The CAR is obtained from the individual firm abnormal returns. The CAR over the 36 months from listing is the sum of the average monthly market-adjusted returns. The BHAR is the difference between the holding-period return of stock 'i' and the market return:

The first measure we use is the three-year buy-and-hold market-adjusted returns (BHAR), defined as:

$$\mathrm{BHR}_{\mathrm{iT}}=\prod_{\mathrm{t}=1}^{\mathrm{T}}\left(1+r_{it}\right) \tag{1}$$

$$\mathrm{BHAR}_{\mathrm{iT}}=\prod_{\mathrm{t}=1}^{\mathrm{T}}\left(1+r_{it}\right)-\prod_{\mathrm{t}=1}^{\mathrm{T}}\left(1+r_{mt}\right) \tag{2}$$

The mean BHAR over a period T is:

$$\mathrm{BHAR}_{\mathrm{T}}=\frac{1}{\mathrm{n}}\sum_{i=1}^{n}\mathrm{BHAR}_{\mathrm{iT}} \tag{3}$$

A simple *t*-test is employed to test the null hypothesis of zero mean market-adjusted buy-and-hold return:

$$t=\frac{BHAR_t}{\sigma(BHAR_{i,T})/\sqrt{n}} \tag{4}$$

Where $\sigma(BHAR_{i,T})$ is the standard deviation of the buy-and-hold market-adjusted returns, and *n* is the sample size.

The second measure we use is the three-year cumulative market-adjusted long-run performance (CAR), defined as:

$$CAR_{iT}=\sum_{t=1}^{T}\left(r_{it}-r_{mt}\right) \tag{5}$$

$$CAR_T=\frac{1}{n}\sum_{i=1}^{n}CAR_{iT} \tag{6}$$

The statistical significance of cumulative abnormal returns is tested by:

$$t=\frac{CAR_{it}}{\sigma(CAR_{it})/\sqrt{n_t}} \tag{7}$$

Where $\sigma(CAR_{it})$ is the standard deviation of cumulative abnormal returns for the sample of *n* firms and n_t is the number of IPOs on the t^{th} month.

What's more, as an additional robustness check, we calculate wealth relatives. Wealth relatives (WR) are the ratio of the end-of-period wealth from holding a portfolio of issuers to the end-of period wealth from holding the market benchmark. A wealth relative less than 1 indicates underperformance relative to the benchmark portfolio.

4. Results on the long-run performance

4.1. Results on the Long-run Performance by Using Shanghai Stock Exchange Index as the Benchmark

Tables 2 and 3 present the abnormal returns calculated using Shanghai Stock Exchange Index as the benchmark.This two tables present the mean percentage CAR and BHAR within 36 months after listing for 166 IPOs listed on the Shanghai A-share market between 2000 and 2002. The initial return at listing is excluded from the estimation of the CAR and BHAR.

From the result, IPO performance has been poor whichever method is used. For example, From Table 1 and Table 2, either method shows the IPO performance-begins to underperformed than that of the market index after 6 months of listing and the mean BHAR and CAR estimated over one year period after issue is -3.2%(t-statistic=-2.314) -2.05%(t-statistic=-1.259), over a two year period after issue is-11.37%(t-statistic=-5.285) and -14.07%(t-statistic=-5.427), and a three-year period after issue is -20.88%(t-statistic=-9.806) and -32.02%(t-statistic=-9.37). One year after listing, WR less than 1 and we obtain a three-year wealth relative of 0.6826.

Figure1 shows the plot of CAR, BHAR and the monthly AR for the first 36 months after the listing. It shows the IPO performance within the first months after listing slightly outperforms the market index. The continuation of this trend varies with the measurement. By CAR, the IPO performance outperforms the market index for the first nine months after listing; while with BHAR,

[1] In China, the total number of market days of each year, less the long holidays, is about 240. However, in consistent with the study with Ritter (1991) and Loughran and Ritter (1995), we assumed that the number of market days per year is 252 and each month has 21 market days. On and from the first day of listing to the 21st market day is the first event month and then the second, the third and so on. So, the word "month" herein refers to a market month, the word "year" to the market year and the express "three years" to 756 market days.

Table 2. Mean of CAR and BHAR of a shares on Shanghai Stock Exchange (RA:A-Share Index of Shanghai Stock Exchange; BHRA: A- Share Index of Shanghai Stock Exchange)

Months/Return	RIPO	RA(2)	CAR (3)=(1) – (2)	BHRIPO (4)	BHRA (5)	BHAR (6)=(4) – (5)	WR
0 –1	– 0.9996	–0.9962	– 0.0034	1.0002	1.0042	– 0.004	0.9963
0 –3	– 0.9817	–0.9841	0.01245	1.0174	1.0042	0.0132	1.0118
0 –6	– 1.0033	–1.0168	0.0135	0.9874	0.9798	0.0076	1.0048
0 –12	– 1.1367	–1.1162	– 0.0205	0.8471	0.8791	– 0.032	0.9581
0 –18	– 1.2489	–1.181	– 0.0679	0.736	0.8069	– 0.0709	0.896
0 –24	– 1.3223	–1.1816	– 0.1407	0.6837	0.8022	– 0.1137	0.8323
0 –36	– 1.5468	–1.2242	– 0.3202	0.5407	0.7495	– 0.2088	0.6826

Table 3. Mean and t-Statistic test of CAR and BHAR. ** significant at the 0.05, * significant at the 0.01 (significance is given by a skewness-adjusted t-statistic)**

Return	Months	Mean	Media	Std. Dev	Skewness	Kurtosis	Min.	Max.	t-Statistic	Sig
	0 –1	– 0.004	– 0.0143	0.0778	0.5281	– 0.1944	– 0.159	0.2138	0.661	0.51
	0 – 3	0.0132	0.0058	0.1278	1.2024	2.7328	– 0.2188	0.5064	1.335	0.184
	0 – 6	0.0076	0.0014	0.1397	0.6874	1.0327	– 0.2541	0.5632	0.699	0.485
BHAR	0 – 12	– 0.032	– 0.0318	0.1783	– 0.0367	0.8188	– 0.6113	0.4725	– 2.314**	0.022
	0 – 18	– 0.0709	– 0.0751	0.2018	0.1141	1.8557	– 0.7564	0.7431	– 4.524***	0
	0 – 24	– 0.1137	– 0.1404	0.2772	1.5804	5.5915	– 0.76	1.2381	– 5.285***	0
	0 – 36	– 0.2088	– 0.2835	0.2743	1.7754	4.7055	– 0.6167	1.1777	– 9.806***	0
	0 – 1	–0.0034	– 0.0126	0.0773	0.3886	– 0.4122	– 0.1674	0.1962	– 0.569	0.57
	0 –3	0.0125	0.0098	0.1212	0.6578	1.0641	– 0.2583	0.4168	1.324	0.187
	0 – 6	0.0135	0.0187	0.1407	0.1368	– 0.1955	–0.2829	0.4385	1.237	0.218
CAR	0 – 12	– 0.0205	– 0.0043	0.2103	– 0.5158	0.9987	– 0.7396	0.4309	– 1.259	0.21
	0 – 18	–0.0679	– 0.058	0.2637	– 0.6502	1.822	–1.1305	0.6636	– 3.319***	0.001
	0 – 24	– 0.1407	– 0.1436	0.334	0.15	1.2852	– 2.2523	0.9658	– 5.427***	0
	0– 36	–0.32.2	– 0.3499	0.4403	0.3586	0.1121	– 1.5934	1.0116	– 9.37***	0

IPO performance outperforms the market index for the first 6months after listing. Although there is a little difference between the long-run abnormal returns by these two measurements, the overall trends are consistent: the IPO performance outperforms the market index within a short period after listing, but it is poor than the market index in the long run. In addition, from the AR of each individual month, the IPO performance is almost equal to the market index within the first nine months after listing, but it drops significantly in the 10^{th} month and then continues to keep equal to the market index. However, its performance for every month after the 16^{th} month is significantly poor than the market index.

4.2. Results on the Long-run Performance by Using Matching Firms

Research in the US has frequently used a matching firm approach to measure long-run abnormal returns. So we also compute the long-term abnormal return with the matching firms. Our reference portfolios are formed continually on the basis of firm size. To construct the size control portfolio, all Chinese stocks are ranked each month according to their market capitalization, and four quartile portfolios are formed.

From Table 4, within one-year, results of abnormal return of an IPO computing by CAR and BHAR begin to show big deviations. The one-year period CAR is -10.89%, but the one-year period BHAR is only -6.14%.

In addition, the differences between the 0-18 month, two-year and three-year period CAR and BHAR are even bigger, especially the two-year and the three-year abnormal returns. The deviations in these two methods results in the matching firm will be adjusted at the end of each year. For the same sample firm, the matching firm is different and there are big differences between the market prices of different matching firms. When different matching firms are linked at the end of each year, big differences may appear between the closing price of the last market day of the previous year and the closing price of the first market day of another matching firm for the current year, thus making the return of the first market day of the current year to be very small or very big. BHAR is obtained using the continuous multiplication, so this makes the results unreliable. On basis of this, when using the return of the matching firm to calculate the

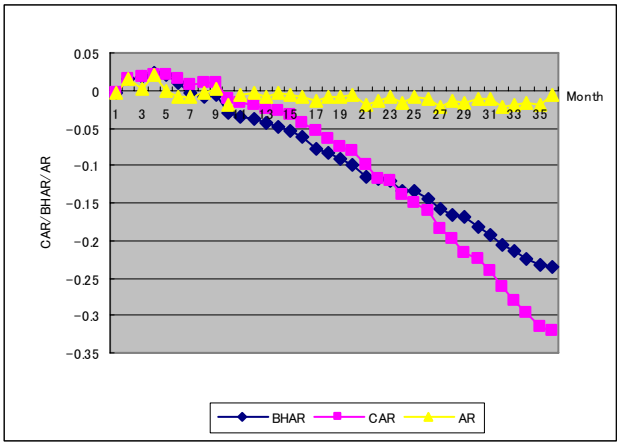


Figure 1. Three year AR, CAR and BHAR for IPOs listed in 2000-2002

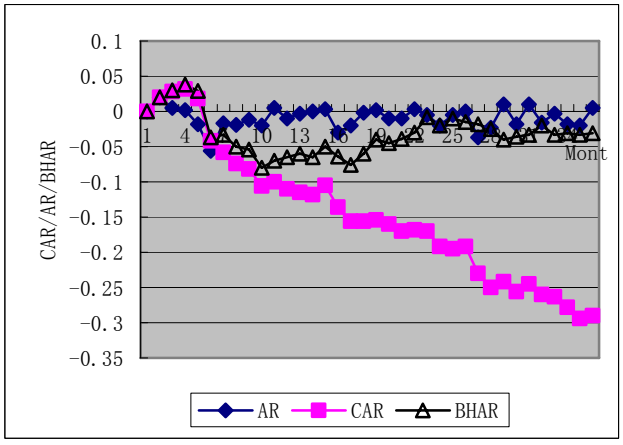


Figure 2 Three year AR , CAR and BHAR for IPOs listed in 2000-2002

Table 4. Mean of CAR and BHAR of A shares on Shanghai Stock Exchange (BHRm: m-matching firm)

Months/Return	IPO	Matching firm	CAR	BHRIPO	BHRm	BHAR	WR
0–1	–0.9996	–0.9976	–0.002	1.0002	1.0027	–0.0025	0.9975
0–3	–0.9817	–1.0048	0.02307	1.0175	0.9905	0.0269	1.0273
0–6	–1.0033	–0.9608	–0.0426	0.9868	1.0175	–0.0307	0.9698
0–12	–1.1367	–1.0278	–0.1089	0.8421	0.9035	–0.0614	0.932
0–18	–1.2489	–1.0938	–0.1552	0.7226	0.7783	–0.0558	0.9284
0–24	–1.3223	–1.13	0.1923	0.669	0.6849	–0.016	0.9768
0–36	–1.5444	–1.2558	–0.2886	0.5147	0.5305	–0.0158	0.9702

Table 5. Mean and t- Statistics of CAR and BHAR. *, ** and * mean that it is significant at 0.1, 0.05 and 0.01**

Return	Months	Mean	Median	Std.Dev.	Skewness	Kurtosis	Min.	Max.	t-statistic	Sig.
	0–1	–0.0025	–0.0055	0.2131	–1.515	10.5136	–1.3176	0.6509	–0.151	0.88
	0–3	0.0269	0.0405	0.2782	–0.6819	3.5738	–1.2764	0.8272	1.247	0.214
	0–6	–0.0307	0.0111	0.4516	–2.8202	15.2875	–3.157	0.8763	–0.875	0.383
BHAR	0–12	–0.0614	0.0549	0.47	–1.885	5.3669	–2.476	0.5415	–1.683*	0.094
	0–18	–0.0558	0.0251	0.4454	–1.8782	6.323	–2.1279	0.7636	–1.613	0.109
	0–24	–0.016	0.0043	0.3925	–1.3557	8.2853	–2.3123	1.2539	0.542	0.601
	0–36	–0.0158	–0.142	0.3567	–0.0927	3.6216	–1.4209	1.3853	0.57	0.57
	0–1	–0.002	–0.0043	0.2228	–1.5615	12.0763	–1.4026	0.7531	–0.114	0.909
	0–3	0.0231	0.043	0.2711	–0.8661	4.4508	–1.3544	0.8025	1.096	0.274
	0–6	–0.0426	0.0107	0.4551	–2.7806	13.7181	–3.0421	0.7497	1.205	0.23
CAR	0–12	–0.1089	0.0076	0.5624	–1.523	3.9671	–2.8921	0.8262	–2.495**	0.014
	0–18	–0.1552	–0.0521	0.5794	–1.3606	3.3573	–2.8201	0.8532	–3.451***	0.001
	0–24	–0.1923	–0.0891	0.6053	–1.1468	2.986	–2.818	1.1703	–4.094***	0
	0–36	–0.2886	–0.245	0.7067	–0.4151	0.802	–2.4443	1.7588	–5.261***	0

long-term performance of IPO, it is better for us to use the CAR to explain the degree of long-term underperformance of IPO.

From Table 4 and 5, the CAR over one year after listing is found to be –10.89%, with a t-statistic of –2.495%, two years after listing is found to be –19.23% with a t-statistic of –4.094%, three years after listing is found to be –28.86% with a t-statistic of –5.261%. The BHAR is significant at the 0.1 level only for the first year and the long-run return is –6.14%. It shows that IPO performance has been poor whichever method is used.Fig.2 shows the plot of AR, CAR and BHAR for each month within three years. From Fig.2, for the first six months after the listing, IPO performance is slightly outperform a matching firm,

but it is underperformed after six months. Although it rebounds after that period, but the degree is small. In the long run, the IPO performance has been poor.

5. Explaining for the Long-run Underperformance

5.1. Some Determinants with Long-run Performance

Previous research has identified several factors important in explaining post-IPO performance. These incorporate the consistent evidence of investor over optimism at issue, together with the perceived riskiness of the firm and the degree to which there is information asymmetry between issuer and investor. Different from the previous studies, in this section, we conduct multivariate regression analyses to examine cross-sectional determinants of the aftermarket performance of China's IPOs from the aspect of the characteristic of a firm. For this purpose, we posit the following regression model:

$$MAR3 = \beta_0 + \beta_1 \text{ LnTA} + \beta_2 RSC + \beta_3 IND + \beta_4 AGE \quad (8)$$

MAR3: the aftermarket performance of IPOs measured by the market adjusted buy-and-hold return for the three year aftermarket period excluding the initial-return period; LnTA: the natural log of total assets before going public; RSC: the percentage of trade shares; IND: industry dummy to show the industry is high-tech or traditional; 1-traditional, 0-high-tech; AGE: the natural log of a firm's incorporation age in years at the time of the IPO.

The first fact is the size of the company. Large companies have less ex ante risk than small companies, respectively, because there is more information about them and because they are likely to be more closely monitored by government and regulatory agencies. We therefore hypothesize negative signs on the LnTA. The second fact is the percentage of tradable shares. The majority of Chinese domestic shares are nontraded shares owned by the government or by other companies, so fewer than half of all shares are tradable. We assume that smaller government holdings result in better restructuring, so the better long-run performance a listed company would expect to have. We hypothesize the positive signs on the tradable shares (RSC). The third factor is the industry effect. Like Ritter [19], we examine the industry effect on the IPO long-run returns. The variable we use is IND: a dummy variable with the value o assigned to high-tech, and traditional is 1. If a company belongs to the high-tech industry, it is expected to enjoy high growth in the near future. However at the same time, companies with high-tech features will face more risk, which could be rewarded by better long-run performance. Therefore, we expect a positive coefficient for the high-tech dummy. The fourth factor is the company's incorporation age in years at the time of the IPO. The old companies tend to be more mature and have stable performances than the younger companies [19]. They expect to have better long-run performance. So we hypothesizes it should be positively related to long-run performance.

5.2. Cross-sectional Analysis

In the regression analysis, the results of overall parameter of the model are presented in Table 6. The variance results of regression are presented in Table 7. the results of multivariate regressions are presented in Table8.

From Table 6, the adjusted R^2 is to increase with the improvement of model. In model 3, it is 12.6%. From the regression result of Table 7, sum of squares are from 1.173 in model l to 1.763 in model 3. the significant of F is lower than 1, it shows that the effect of regression is significant at 0.01 level.

From Table 8, three variables are in the model, and they are all at the significant of 0.05 level. As for the relationship between the company size and IPO long-run returns, the estimation results shows that it have no relationship with the long-run returns. The variable that measures the percentage of tradable shares is negative. In China, after listing, the government is still the biggest shareholder of listed companies. Unlike the developed markets, particularly the US, they suggest that the proportion retained by current shareholders reflects the demand for an IPO [6]. In China, the listing mechanism does not allow demand to feed back to either the size of issue or the issue price. Our result is consistence with Tian [21] he concludes that firm performance is positively related to government ownership when the level of owner ship is greater than approximately30%. In China, government ownership is beneficial to firms.

Turing to the relationship between the industry effects, the coefficient is 0.154, the traditional industry have a better long-run performance. Incorporation age and long-run performance are negatively related to each other at the 0.05 significance level. This is similar to the findings in previous research (Chen&Meng2000). Incorporation age we calculated is after the shareholding reform.

6. Conclusions

This paper attempts to satisfy the great interest in Chinese evidence on the long-run performance of 166 A-share IPOs on Shanghai Stock Exchange between January 2001 and 2002. The main results are the following:

①. When we use shanghai stock exchange Index as the benchmark to calculate the CAR and BHAR of our sample within 3 years after listing, either method shows the IPO performance begins to underperformed than that of the market index after 6 months of listing and the mean BHAR and CAR estimated over one year period after issue is –3.2%(t-statistic=–2.314) –2.05% (t-statistic= –1.259), over a two year period after issue is –11.37%(t-statistic=–5.285) and –14.07% (t-statistic= –5.427), and a three-year period after issue is –20.88% (t-statistic=–9.806) and –32.02% (t-statistic=–9.37).

Table 6. Results of parameter of the model

Modle	R	R Square	Adj.R Square	Std.Error of the Estimate	Durbin-Watson
1	0.307(a)	0.094	0.089	0.2618	
2	0.344(b)	0.118	0.108	0.2591	
3	0.377(c)	0.142	0.126	0.2564	1.838

a Predictors:(constant), SC

b Predictors:(constant), RSC, AGE

c Predictors:(constant), RSC,AGE, IND

Table 7. Variance results of regression

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.173	1	1.173	17.108	0.000(a)
	Residual	11.214	164	0.069		
	Total	12.414	165			
2	Regression	1.47	2	0.735	10.951	0.000(b)
	Residual	10.944	163	0.067		
	Total	12.414	165			
3	Regression	1.763	3	0.588	8.937	0.000(c)
	Residual	10.651	162	0.066		
	Total	12.414	165			

a Predictors:(constant), RSC

b Predictors:(constant), RSC, AGE

c Predictors:(constant), RSC,AGE, IND

d Dependent Variable: BHAR3

Table 8. Results of multivariate regressions: determinants of long-run performance of China's IPOs

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std.Error.	Beta		
1	contant	0.057	0.067		0.84	0.402
	RSC	–0.009	0.002	–0.307	–4.136	0
2	contant	0.197	0.094		2.091	0.038
	RSC	–0.009	0.002	–0.310	–4.22	0
	AGE	–0.018	0.008	–0.155	–.106	0.037
3	contant	0.163	0.095		1.721	0.087
	RSC	–0.009	0.002	–0.315	–4.325	0
	AGE	–0.018	0.008	–0.157	–2.150	0.033
	IND	0.084	0.04	0.154	2.109	0.037

One year after listing, WR less than 1 and we obtain a three-year wealth relative of 0.6826.

②. In order to reduce the influences of incorrect model setup errors on the research results, we also use matching firm approach to measure long-run abnormal returns, we also found the CAR over one year after listing is found to be –10.89%, with a t-statistic of –2.495%, two years afterlisting is found to be –19.23% with a t-statistic of –4.094%, three years after listing is found to be –28.86% with a t-statistic of –5.261%. The BHAR is significant at the 0.1 level only for the first year and the long-run return is –6.14%.

③. By looking at the market adjusted buy-and-hold return for the three year aftermarket period, and using a cross-sectional analysis, we have a better understanding of the long-run performance of Chinese IPOs. From the

characteristic of a firm, the ones with more government ownership, traditional industry features and the shareholding reform made later are better performs.

The results show that China's IPO performed better than market index returns and the returns of matching firms within a short period after listing (6 months) but exhibited significant underperformance in the long run. The results obtained from this study provide important information for prospective investors in new issuers to understand better the Chinese IPO markets. As we mentioned before, the aftermarket underperformance is appeared not unique to the developed market, like US. It also exists in a number of countries, like Canada, Japan and so on. Chinese IPO market is still a emerging market and is small relative to the overall economy and many structural and institutional problems remain, but it appears that the development of the market is approaching maturity, holding out promising long-term prospects.

7. Acknowledgement

I would like to thank Professor Hamada Yasuyuki for a number of helpful comments on my paper. I also thank my research lab colleagues for their friendship. All errors are the author's responsibility.

REFERENCES

[1] Aggarwal, R., R. Leal and L. Hernandez, The After-market Performance of Initial Public Offerings in Latin America, *Financial Management*, Vol,22, 1993, pp.42-53.

[2] Barber B.，Lyon，J. Detecting long-run abnormal stock returns: the empirical power and Specification of test statistics, *Journal of Financial Economics*,Vol.43, 1997,pp.341-372.

[3] Cai, J., and K.Wei.. The investment and operating performance of Japanese IPO, *Pacific Basin Finance Journal*, Vol.5, 1997, pp.389-417.

[4] Chan, K., J.B. Wang and K.C. Wei, Underpricing and Long-term Performance of IPOs in China, *Journal of Corporate Finance*, Vol.10, 2004, pp.409-430.

[5] Fama, E.F., Market Efficiency, Long-term Returns, and Behavioral Finance, *Journal of Financial Economics*, Vol .49, 1998, pp.283-306.

[6] Hanley, K., The Underpricing of initial public offerings and the partial adjustment phenomenon. *Journal of Financial Econ,* Vol. 34, 1993, pp.213–250.

[7] Keloharju, M.. The winner's curse, legal liability, and the long-run price performance of initial public offerings in the Finland. *Journal of Financial Economics*, Vol.34, 1993, pp.251-277.

[8] Kim, J., I, Krinsky, and J.Lee. The aftermarket performance of initial public offerings in Korea, *Pacific Basin Finance Journal*,Vol.3, 1995, pp.429-448.

[9] Kooli,M., and J.M.Suret. The aftermarket performance of Canadian IPOs. *Journal of Multinational Financial Management*: Forthcoing, 2003.

[10] Kothari, S.P. and J.B. Warner, Measuring Long-horizon Security Price Performance, *Journal of Financial Economics* Vol.43, 1997, pp.301-340.

[11] Loughran,Tim, Jay R.Ritter, Kristian Rydqvist, "Initial Public offerings: International Insights", *Pacific Basin Finance Journal*,Vol.2, 1994, pp.165-199.

[12] Levis, M., The Long-run Performance of Initial Public Offerings: The UK Experience 1980-1988, *Financial Management*, Vol, 22, 1993, pp.28-41.

[13] Ljungqvist, A.. Pricing initial public offerings: Futher evidence from Germany. *European Economic Review*, Vol.41, 1997, pp.1309-1320.

[14] Loughran, Tim, Ritter J. "the new issues puzzle", *Journal of Finance*, Vol.50, 1995, pp.23-51.

[15] Loughran, T., J.R. Ritter and K. Rydqvist, International Public Offerings: *International Insights, Pacific-Basin Financial Journal*, Vol. 2, 1994, pp.165-199.

[16] Michael Firth an Analysis of the stock market performance of new issues in New Zealand，Pacific *Basin Finance Journal*,Vol.5, 1997, pp.63-85.

[17] Lyon, J., B. Barber and C.L. Tsai, Improved Methods for Tests of Long-run Abnormal Stock Returns, *Journal of Finance*, Vol. 54, 1999, pp.165-201.

[18] Mok, H.M.K. and Y.V. Hui, Underpricing and After Market Performance of IPOs in Shanghai, China, *Pacific-Basin Finance Journal*, Vol. 6, 1998, pp.453-474.

[19] Ritter，Jay R. "The long-run performance of initial public offerings", *Journal of Finance*, Vol.46, 1991, pp.3-27.

[20] Sun, Q. and H.S. Tong, China Share Issue Privatization: The Extent of its Success, *Journal of Financial Economics*, Vol .70, 2003, pp.183-222.

[21] Tian, G., State shareholding and the value of Chinese Firms, Working Paper, London Business School Wong K, Tan R, Liu W.2006, The cross-section of stock returns on the Shanghai stock exchange. *Rev Quant Financial Account*, Vol, 26, 2000, pp.23–39.

J. Serv. Sci. & Management, **2008, 1: 143-152**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

Systems Plan for Combating Identity Theft – A Theoretical Framework*

Shao-Bo Ji[1, 2], Shawn Smith-Chao[1] & Qing-Fei Min[2]

[1]*Sprott School of Business, Carleton University, Ottawa, ON, Canada K1S 5B6*
[2]*School of Management, Dalian University of Technology, Dalian 116024, Liaoning, China*
E-mail: sji@business.carleton.ca

ABSTRACT

The Internet has made it easier for individuals and organizations to communicate and conduct business online. At the same time, personal, commercial, and government information has become a target for identity theft. The incidences of identity theft have increased substantially in the Internet age. Increasing news reports of bank/credit cards theft, assumed identity for economical and criminal activities has created a growing concern for individuals, businesses, and governments. As a result, it's become an important and urgent task for us to find managerial and technical solutions to combat identity fraud and theft. Solutions to identity theft problem must deal with multiple parties and coordinated efforts must be made among concerned parties. This paper is to provide a comprehensive view of identity theft issue from system planner's perspective. The roles of identity owner, issuer, checker, and protector, are examined to provide a starting point for organizational and information systems design.

Keywords: *identity theft, information system plan, inter-organizational coordination*

1. Introduction

Identify theft problem has been one of the major concerns for individuals, businesses, and governments in the Internet age. Businesses and governments have been trying to find solutions through legislations, law enforcement, and new technologies. In academic community, the research into the issue of identity management is generally lacking. Most researches have been focusing on technological and/or legal issues. Those researches have dealt with the identity problem from an operational and narrowed perspective, rather than conceptual and systematic viewpoint [1, 2].

Identity theft is "…the misuse of another individual's personal information to commit fraud" [3]. Within research community, there seems to be no consensus of what identity theft is. For example, Sproule and Archer [4] gave one of the definitions for identity theft as "crimes involving use of a real person's identity". Others define identity theft as "false identifiers, false or fraudulent documents, or a stolen identity in the commission of a crime" [1] and as "the unauthorized collection and fraudulent use of someone else's personal information" [2]. For the purpose of this paper, the definition suggested in [3] is adopted.

The seriousness of identity theft is difficult to determine, since the crime may be undetected for a long period of time, and there is no one centralized database for identity theft. Anecdotal evidences and statistics drawn from a number of sources do indicate, however, that the problem is growing in many parts of the world, particularly in North American, i.e., United States and Canada, and the problem has quickly become a global phenomenon (see for example reports at www.ftc.gov and www.rcmp.ca as well as the recently released The President's Identity Theft Task Force's Report [3]). Consequently, tremendous efforts have been made over the past decade by governments, businesses, and academic research community of various disciplines to understand the issues and to find solutions for combating the problems from the aspects of social and technological, legislative and law enforcement, and business and management. As a result, a number of comprehensive research framework, model, and practical solution have been proposed and the results of the research initiatives are emerging [3, 4, 5, 6, 7]. In developing countries such as China, although there is a lack of official statistics, over the past decade, the amount of identity theft related reports, such as cell-phone short messages fraud and fake identity cards and seals, has been on the rise significantly due to the wide spread use of the Internet and mobile devices such as the cell-phone [8]. A comprehensive literature review in identity theft research can be found in Newman and McNally [9].

Identity thieves use a number of techniques to acquire data from individuals. Unsophisticated, but effective tech-

*This study is funded by Ontario Research Network for Electronic Commerce (ORNEC), Ontario, Canada.

niques include the theft of wallets, cell-phone, and laptops. Credit card statements or other documents can be taken from trash or intercepted in the mail. One of the most effective means of theft is called "Social Engineering", which involves the thief contacting the victim directly, and convincing them to disclose passwords or other information by posing as agents of ID issuing agencies. Victims are also targeted over the Internet, through social engineering by email message or by intercepting and capturing financial or identity information while it is being transferred online. The information of a large number of individuals can be captured at one time by hacking into and stealing data from financial or government databases, a crime which has drawn significant media attention. In China, due to the lack of individual's privacy concerns and the lack of mechanisms of public protection, the situations seem to be more serious. For instance, many such individual's information as name, date of birth, identity card number, and cell-phone number are openly published in the public domain such as the Internet, posters, and internal documentations.

Combating identity theft requires actions from multiple parties. It requires technological, legal and law enforcement, economical, and managerial solutions. Solving identity theft problem must deal with multiple parties and coordination among the parties since the confirmation of individual's identity often involves multiple steps, multiple methods, and with multiple parties. This is especially important in the virtual environment where the verification, validation, and authentication of identity are conducted online and virtually no face-to-face identity confirmation. To this end, the US government has taken a major step in dealing with the issue. The President's Identity Theft Task Force's Report entitled "Combating Identity Theft: A Strategic Plan" is an important blueprint that guides government agencies, businesses, and individuals to combating identity theft [3]. Released in April 2007, the report not only describes and summarizes the identity theft stages and sequences but also, more importantly, recommends and proposes an important strategy from the perspective of identity theft prevention, consumer protection, and law enforcement. As outlined in the report, the occurrences of identity theft generally take place in three stages: 1) the acquisition of a victim's personal information by the identity thief; 2) the attempt of misuse of stolen information by the identity thief; and 3) the identity thief committed crime and the victims suffer the loss (emotional and/or financial). The reports made a number of recommendations which focus on the four areas: 1) identity theft prevention by keeping individual consumer data safe, in public and private sectors as well as individual consumers, through better data security measure and awareness education; 2) preventing identity thieves from using consumer data to steal individual's identity through authentication and comprehensive record keeping mechanism of private sector use of social security numbers; 3) the creation of assistance program of identity theft; 4) stronger law enforcement through National Identity Theft Law Enforcement Center, better coordination and information intelligence sharing between law enforcement and private sector, and better coordination with foreign law enforcement. The report provides an excellent starting point for government agencies, private businesses, and consumers in combating identity theft. It is particularly useful for public policy formulation. Similarly, a number of studies have been conducted to address the issue from consumer's perspective in the financial service sector [5], the policy options in coping with identity theft [6], and the general description and trends in different regions of the US [7]. For example, Mayer investigated "the state of consumer self-protection with respect to financial privacy" [5]. Using content analysis, field experiment, and telephone survey, he identified various advice offered from government, business (nonprofit and for profit organizations), and financial journalists and concluded that consumers are general aware of the problems and they are facing the trade-off between keeping their privacy and spending money in protecting their identity. The study suggested that more efforts are needed for governments and businesses to create an environment that provides consumers efficient methods of protecting their identity. Other studies explored the potential solution for combating identity theft that goes beyond existing recommendations and practices. For example, Eisenstein suggested that "the existing approach to combating identity theft will not work" [6]. Applying system dynamic modeling technique, Eisenstein proposed and tested a model that explained the motivations and actions of various players of identity theft. He concluded and suggested that an "inexpensive security freezes" can be an effective way of reducing identity theft.

While aforementioned US government strategic plan and recommendations suggested in various studies are useful in dealing with public policies, understanding the nature of identity theft, and finding solutions in combating identity theft, they are limited in the sense of covering the scope of the issues. More works have yet to be done to implement the recommendations included in the US Strategic Plan. As stated earlier, the nature of the identity theft is identity thieves' misuse of another individual's personal data. As a result, information technology and information systems play a special role in the process. As a powerful tool, information system and associated technologies can be designed in such a way that it can facilitate the process of combating identity theft. This study complements the US Strategic Plan by providing operational level recommendations from the viewpoint of systems analysis and design. It pays special attention to multi-party coordination in combating identity theft which is strongly recommended by the report.

From system analysis and design perspective, understanding identity theft problem and designing managerial and technological systems to combat identity theft re-

quires a holistic and comprehensive framework. Specifically, identifying key players and their associated roles and relationships is an important first step. In this aspect, two frameworks are suited for such purpose. The first framework is Wang *et al*.'s [10] contextual framework for combating identity theft. The second is a framework proposed by Zachman [11]. Wang *et al*.'s framework identified five main roles involved in the ownership, issuance, use, protection, and abuse of identity information. These roles are identity owner, identity issuer, identity checker, identity protector, and identity thief. They are linked by a web of relationships which identify the principle responsibilities and interactions of the members of the identity verification, validation, and authentication chain. Zachman's framework defines the interfaces and the integration of all of the components of a system by perspective and aspect. The framework has been applied in many applications and enterprises and has been proposed for part of the US government's internal enterprise architecture [12].

The purpose of this paper is to demonstrate how the Zachman framework can be applied to the conceptual model by Wang *et al*. [10] to provide a comprehensive and complete view of the identity theft problem.

2. Identity Theft Conceptual Framework

Preventing identity theft are the responsibilities of many individuals in many organizations. As outlined in Wang *et al*.'s [10], there are primarily four stakeholders and one aggressor. The focus of identity theft issue is the identity owner, the person the data actually identifies and to whom the identity information and documents are possessed or assigned. The owner's physical, biographical, psychological, and/or financial data is created and stored on the identity document or database. The identity issuer, normally a government institution or financial institution, acquires, creates, and produces these information and documents. When the identity owner wishes to make transactions using their identity information or documents, they are checked by an identity checker, which is a person or device responsible for determining the validity of his/her identity information and documents. From information processing viewpoint, this process typically go through three stages: verification, validation, and authentication. The checker works with the identity issuer, and often these two roles are accomplished in the same organization. For example, banks have their own identity checker to validate the owners of accounts. The identity protector is responsible for determining if the security of the identity chain has been breached, and for prosecuting and punishing identity thieves. The protector works with issuers and the checkers to maintain a vigilant guard, but has a largely post-incident rapport with owners. Finally, the identity thief/abuser is, of course, any party attempting to misuse, copy, or steal the identity owner's information or documents.

PERSPECTIVE

	Things	Processes	Connectivity	People	Timing	Motivation
Planner	List of important things to the business	List of processes business performs	List of locations in which the business operates	List of business responsibilities	List of events significant to the business	List of business goals/strategy
Owner	Semantic model	Business process model	Logistics network	Work flow model	Master schedule	Business plan
Designer	Logical data model	Application architecture	Distributed system architecture	Human interface architecture	Processing structure	Business rule model
Builder	Physical data model	System design	Configuratio n design	Presentation architecture	Control structure	Rule design
Subcontractor	Data definition	Program	Network architecture	Security architecture	Timing definition	Rule specification
Functioning Enterprise	Data	Process	Network	Organization	Schedule	Strategy

ASPECT

Figure 1. Zachman framework of information system architecture

The importance of such framework is that it identifies the main parties and the interconnections between the various parties. This is critical since many identity theft experts have listed a lack of coordination of relevant parties as a key factor in the proliferation of identity theft crime [1]. Coordinated solutions include the use of shared identity databases [1, 15, 16], international laws for identity theft and fraud [1], cohesive national law enforcement programs [1, 2], and trusted authenticators [17].

3. Zachman's System Framework

First published in IBM Systems Journal in 1987, the framework is considered as a "logical construct (or architecture) for defining and controlling the interfaces and the integration of all of the components of a system." [11] Shown in Figure 1, the framework is composed of a matrix of six columns and six rows. The rows represent the main perspectives (of different roles) for viewing and framing the system: planner, owner, designer, builder, subcontractor, and functioning enterprise. The columns represent aspects (of different views) of the system: things, processes, connectivity, people, timing, and motivation.

For any system design and implementation, various parties are involved. First, the system planner initiates the conceptualization of the system, and then determines its purpose and scope based on system owner's requests. Then, the system designer, the architect of the system, takes the owner's requirements for the system and design the system according to the technical and economical feasibility set out by the planner. The builder is responsible for the actual realization and implementation of the system. The builder may employ subcontractors to do the actual construction and assembly of the system. The final result of the process is the functioning enterprise. Each perspective in the framework (other than the functioning enterprise) can be embodied by a person, a group, an organization, or another system.

The rows of the system correspond to what (things), how (processes), where (connectivity), who (people), when (timing), and why (motivation), and. Things represent the materials or structures that the system is composed of. Processes represent system's functional specifications. Connectivity shows where the linkages exist in the system and the locations of process flows. People refer to individuals who perform their tasks and assume their responsibility. Timing indicates when events are expected to occur. Motivation refers to the motivating elements of the system such as the purposes and intentions of the system owner. The framework provides a tool for us to address main system questions and main perspectives of all relevant parties. The intersections of each row and column are called artifact and they are filled with the relevant diagrams, lists, or structures. Examples of artifacts used when designing systems architecture are shown in Figure 1. An artifact can be modeled in a variety of means, using one or more tools, as long as the contents match the purpose of the aspect, and the relevant perspective.

4. A Framework for Identity Theft Issue

A complete Zachman's framework for each identity role and their interactions will require the knowledge of identity management, resources, as well as system and organizational system requirements. Detailed analysis of each cells and intersections (6 rows and 6 columns and their intersections) will require a much more lengthy discussion. For the purpose of this paper, we focus on the role of system planner. Reviewing the planner's role for each member of the identity chain addresses many of the key aspects of the identity theft problem, and can be helpful for organizational and system designers to understand system requirements, organization or information, and to manage the identity and to prevent identity theft from occurring. In this section, each role will be explored using system planner's role.

4.1. Identity Owner

The identity owner is the originating source of the identity chain. Though the issuer is responsible for the determination and creation of all identity information and documents, the process requires the existence of an owner to initiate the identity definition process. In a special situation, identity thieves can create the identity information and documents based on a fictitious person or a dead person. The source data for the identity information and document often describes the attributes possessed by the identity owner such as age, date of birth, addresses and additional data that are assigned by the identity issuer such as a bank account or identification numbers. Usability, security, and privacy are typically the main concerns of the owners. These artifacts are somewhat paradoxical, as ease of use is not generally associated with increased security and privacy. Tradeoffs are often required. This however represents a key need for owners, and the responsibility for meeting this need is shared with the issuer. Documents which are cumbersome to hold and use are not attractive. Similarly, documents which are not secure are not attractive. Input from both the owner and issuer is necessary to balance these two concerns.

Privacy is a real concern to the owners. Although related to security, privacy indicates the desire for the owner to minimize the exposure of their information to the issuer, and to other members of the identity chain. Issuers have to consider the privacy concerns of owners, they want to maximize the amount of data they collect from owners, to support the needs of their identity information and documents. Privacy is a key artifact for the owner, and may be a weakest point in the identity chain. In most western countries, some owners' desire to maintain privacy will prevent them from entering the chain at all. Addressing the needs of these owners is difficult,

since suspicion and a lack of trust often prevent the issuer or other identity chain member from making meaningful contact with the owner. The identity protector and information security technology plays a key role here.

The identity owner establishes the identity that will be used and stored in identity database and shown on documents. The owners collaborate with the issuers to create data and documents, and also to update or change data as their personal information change. The creation and updating processes are important, as the validity of identity documents rest on their accuracy of the data within. Examining the intersection of things and processes, we see that these processes must be easy, secure, and private for them to fulfill the owners' needs.

To meet these requirements, owners need to use their identity data and documents responsibly and to protect their own information and to report any abuses. Irresponsible document use can obviate any security procedures put into place, and place the owner beyond the reach of identity protectors, and even possibly in opposition to them or other members of the identity chain. Unfortunately, the fraud detection and reporting process is one of the weakest and least developed processes in the identity chain [9, 13, 14]. Although identity issuers, checkers, and protectors all play roles in this process, the owner can usually detect abuses first. Proper security procedures by the checker and issuer can prevent transgression, however, once they have occurred, they will remain hidden unless until discovered by the owner, often more than a year later [2].

The owner needs to use their identity information and identity documents at a number of locations. Transaction points represent any places (physical or virtual) where an identity owner and identity checker interact. These include retail stores, issuer checkpoints and offices, government offices, online buying sites, and even at home. Home as a location has special significance, since this is where identity information and documents are stored. This is a common breach point for criminals who acquire information about the owner and their identity from trash, or theft of mails, or the physical documents [2]. Identity information and documents are also required for work, for verification of identity, or for hiring or human resource purposes. Work locations and databases are potential areas of threats and point of penetration since these systems are maintained or monitored by fewer and less stringent security systems than identity issuer sites [18]. Employee records are in fact the number one source for identity data for thieves [19]. Increasingly employers have become the targets for legal action, due to improper storage and handling of employee data [20].

Maintaining the security of identity information and documents is one of the two responsibilities (of people) assigned to the owners. This responsibility is shared with identity checkers and identity issuers. However, the primary target for most identity theft is the owner [13]. According to one report, over 50% of identity victims know the identity of the person who defrauded them [13]. Owners should ensure that they maintain physical control of their identity documents as have a strong knowledge of where their identity data is being used and stored, but for most people this is at best a difficult task. Due to large number of transaction points for identity information and documents used, often people are not aware of the spread of their identity data.

The integrity of the identity networks that are built around individuals rests also on the owners presenting truthful and accurate identity data about themselves. When owners lie or falsely present their data, they move beyond the reach of the security procedures of checkers and issuers, and the legal reach of identity protectors. Fraudulent behavior on the part of owners is another area that needs attention for research [21, 22]. Though some identity checks are done by inspection (looking at signatures and photo identification), many are done electronically. Electronic checks rely on having consistent data throughout. As discussed later, data dispersion is a key concern for the identity chain process, as this increases the chances of inaccuracies, and improper updates. Centralization of data simplifies both the checking process, and decreases the number of inaccuracies, leading to drastic improvements in the costs of managing customer data [23, 24].

The motivations of the identity owner are straightforward. Owners want to be able to use their transaction data, while maintaining the security of their identity information. An identity control and use process that can address and improve both concerns would be optimal. Developments in biometrics technologies may facilitate the ability for owners to safely use their identity data and documents. Work in this field has encountered a great deal of resistance due to privacy concerns by groups and individuals [25, 26]. Proposed solutions must allay these fears without compromising the distinct advantages of biometric identification. It must be realized, however, that even biometric solutions are not foolproof, unless care is taken to maintain security of biometric data, and cooperation among the identity chain members [27, 28, 29].

4.2. Identity Issuer

Identity issuer is the force behind the identity process. Government agencies, financial institutions, employers, retailers, and professional organizations create identity information and documents. The information and documents are often critical or even central to their businesses. Credit card companies for example thrive on the transactions made with their cards. Banks maintain and operate financial accounts for owners, drawing income from fees and investments. Governments are mandated with providing services that require identification, and track individuals for the purposes of security and taxation.

The identity issuer must guard the interests of owners. They need to coordinate with identity protectors to ensure the security of the data. Although the identity issuers have a number of unique artifacts, they share a substantial number with other members of the identity chain.

Issuers hold the repositories for identity data and create identity documents. They support their businesses by ensuring that transactions are secure and that any losses to themselves or other members are minimized. Credibility is a key artifact with respect to issuers concerns. Owners that do not have faith in the security of their identity data will abandon the use of the issuer's documents [30, 31]. In addition, identity checkers that find the verification process cumbersome or insecure may not recognize or accept the identity documents.

Identity issuers handle a number of processes related to their documents and the owners' data. They produce the documents themselves and often create the identifiers used in transactions with the documents. Maintaining, updating, and altering identity data is also done by the issuers. Issuers are also a likely party to find abusers, by examining data looking for suspect transactions, and by tracking bad verification attempts. Issuers are also partners with identity protection services, working to protect other identity chain members.

An important but overlooked process for issuers is the responsible use of identity documents. Particularly for governments and financial bureaus, a call has been made to clarify what these organizations can and are allowed to do with the data they control [32, 33, 34, 35]. Examining the overlap between things and processes, it can be seen that the desire to increase revenue, minimize loss, and exposure must be balanced against credibility issues. Revenues can often be increased by issuers by distributing the information they hold to third parties. They can also attempt to minimize loss by coordinating and distributing data with checkers and protectors. For example, verification through biometric identifiers or cross validation through multiple identifiers can improve the chances that the issuer or checker cannot be the victim of fraud, but this puts the owner at risk, which is exposing greater and more sensitive data in order to verify their identity.

Though owners may use identity documents at a number of transaction points, the creation, verification, and update of these documents occurs at branches of the issuing organization. Though the contact may be physical or virtual, all transaction will flow through these nodes. Even identity checking is done by contacting the onsite databases of the issuing organization. This allows the issuer to ensure that validation procedures are consistent, and that validation data is centralized.

Centralization and consistency fit with the issuer's role of safeguarding identity data. Although owners are the custodians of the actual identity documents, identity data security is shared by both owners and issuers. Other document processes, such as creation, maintenance, and destruction are handled exclusively by issuers.

Maintenance tasks include updating identity data, replacing damaged, lost, or stolen identity documents. Document destruction may be in response to theft attempts, or if the owner no longer wishes to interact with a particular user. Institutional and government issuers may also require document destruction if the personal leaves their jurisdictional realm or due to death. The topic of document maintenance and destruction and its impact on the identity cycle has not been extensively researched, however the media has addressed the issue [36].

The issuer is the most frequent contact with identity data and documents, due to their control of the actual data repositories, and the necessity for use of this data for update, verification, and creation events. Issuers are also the first contacts with owners who are seeking identity documents. This particular event is relevant, as the actual validation and assertion of the identity of an owner occurs at this time. Errors in collection of identity data, or successful fraud on the part of the owner can create an inaccurate identity, which can have a long-term impact on all further transactions. Often identity documents are validated based on other identity documents are data. Errors can compound, and allow identity thieves to breach security processes. With the creation of a single identity document (called a breeder document, an identity thief can spawn numerous other documents, solidifying their fraudulent identity. The longer the chain that is spawned, the more difficult it may be to detect the original error or fraud [1]. The motivations of identity issuers vary based on the nature of the organization.

Institutional and governmental bodies require identity documents to identify and track individuals. Documents of this type include birth certificates, driving licenses, health service cards, and passports. These documents may or may not have attached services or privileges. These documents are the most common breeder documents as the institutional weight adds credibility to the identity document. This is also why these are the documents of choice for identity thieves, as a successful intrusion in this early stage enables a large number of other documents to be obtained [1]. Fraudulent documents that are not institutional are more easily detected, as they lack the necessary and often required supporting documentation. This is why the Internet crime has proliferated, as it enables criminals to secure non-institutional documents and use them without supporting institutional data. Validation using this data can help curb this type of fraud and theft. Biometric identification can also act as a strong deterrent to online crime, but this has not been implemented extensively online at this time.

Non-institutional issuers generally operate to support their businesses, and often for financial gain. Identity documents of this type include credit cards, bank transac-

tion documents (cards and books), membership cards, and frequent transaction cards. Documents of this type usually have an attached service or privilege, and often incur fees for use. Though these documents offer benefits for users to attract adoption, the primary purpose of these documents is for the financial gain of the issuer. As such, the credibility and security of these documents is a key concern for the issuer. The issuer often suffers monetary or credibility loss if a breach is successful. For this reason the issuer will often employ identity protectors themselves rather than relying on institutional protectors.

It's worthwhile to mention here is the fact that many western countries such as Canada and UK, as a result of 9/11, have started examining the issue of creating national identity card. For example, the British government passed the British Identity Cards Act in 2006 [39]. It will help government to better manage the identity data. At the same time, the implementation of the Act raised many concerns due to potential misuse and privacy breaches. Other countries such as Canada abandoned the suggestion of creating national identity database due to its citizen's privacy concerns.

4.3. Identity Checker

The identity checker locates at the central point of the identity chain. They process transactions using the identity owner's data and documents, and verifying their authenticity with the identity issuer. The identity checker, however, is the party with low incentives to work with the other identity chain members. In most cases, they have little to gain by accepting greater responsibility. However, identity checker is actually the party best suited for preventing fraud and theft.

Identity checker is mainly motivated by financial gain in case of commercial organizations, or protection of financial loss and prevention of crimes in case of government agencies. Owners use their identity information and documents to obtain goods and services from the identity checker, who uses information to establish identity, and to process the payments. It is certainly in checker's best interest to ensure the credibility of its transaction process, and to guarantee the security of the owner's identity. This conflict of motivation leads to a central conflict in the identity chain. For example, databases are critical to allow retailers to respond to customer needs. These databases contain identity data about their customers, some of which is collected without the knowledge or explicit consent of owners.

The data is a prime target for thieves, as these databases are a far more efficient means of collecting identity data than targeting owners directly. The identities of the customers can be accessed from a single database. Identity thieves often have access identity information within checker's organizations since they know the existing security protocols or lack thereof. Until recently, incentives do not exist for identity checker organizations to monitor their employees. However, the introduction of the recent legislation makes businesses liable for losses from employee theft or careless safeguarding of identity data [20]. It is therefore critical that checkers consider their role as a guardian. Checkers who are focused solely on material gains will overlook the necessity of this role, and may even actively sell identity data to the third parties. They may also be less careful in their processing procedures of identify data.

Verification procedures often require confirmation from secondary identity sources, a process that is often ignored by checkers who are not motivated to follow these procedures, with a few exceptions such as credit card transaction. Additionally, identity documents often can only verify that the person who is holding the identity document is who they say they are. It cannot be used to validate the actual rights of the holder with respect to the issuer [1]. Possessing a forged document can allow an identity thief to get approval with respect to identity checkers, since the biometric data will seem to be correct.

Checkers and issuers must still be careful to ensure that the identity document holder can actually be validated as a recognized recipient of the benefits conferred by possessing the document [37]. This is a key issue with respect to future identity theft proposals.

The early detection of a thief can prevent loss to the owner or issuer. Issuers often require patterns of transactions to identify thieves, or the commission of a starkly inappropriate transaction. Smart thieves however can elude these means of detection. The issuer relies heavily on owners and checkers to report crimes, but both parties have low reporting rates [13, 38].

4.4. Identity Protector

The identity protector is an important member, typically with certain authority and power, of the identity chain. The protectors may often interact with issuers and with owners, and even checkers, their responsibilities and processes are unique. The motivations of the protectors are not directly related to the transaction processes, and the existence of the protector role relies on the existence of identity fraud and theft. Protectors are generally the last party to have a presence in a particular transaction chain, or identity issue. They maintain lists of offenders and complaints, and use these to investigate abuses. They also compile statistics which can be used to gauge which transaction points, methods, and documents are most at risk for intrusion by thieves. Where possible, protectors will prosecute offenders, or will at least create incident reports that can be tracked for future use.

In addition to the reactive processes, they also develop methods to detect identity fraud and theft and they create legislation which can be enforced. Although identity theft is not a new crime, its proliferation on the Internet is a recent development. The protector role is evolving, and

has been criticized as being indifferent in the past. Identity checkers were also ambivalent since the costs were passed on to the issuer or the owner. There was very little protection afforded to identity owners, since the identity issuer would often make it difficult for the owner to recoup their losses. The introduction of Identity Theft Federal Law in the United States and subsequent other laws have rectified many of these issues, giving owners more protection, issuers more responsibility, and protectors the legal basis to prosecute thieves [9, 35]. Still, many crimes are not reported. This makes protectors at a difficult situation. Issuers and checkers often do not report crimes and losses, fearing it may tarnish their credibility. Instead they deal with the theft as a "cost of doing business" [14]. Owners themselves are often to blame, as only 26% of owners report incidents to the police [13].

One of the greatest challenges in the legal arena is that the Internet is global, and therefore requires global laws and global protectors. At this time, such agencies and laws are not in place. A comprehensive system of global laws and a policing agency with global jurisdiction will be required to successfully curb and control identity fraud and theft in the future [1].

Protectors enter the transaction chain usually when the owner or issuer contacts them. It is possible that protectors may be employed by issuers, and also be located at issuer sites. The protectors are also likely to be monitoring actual transaction points, looking for abuses. In general however, owners or issuers bring their problems to protectors who then act upon that information.

The identity protector has a number of important roles. Most protectors use passive protection. Though the protector will produce materials that attempt to assist the owner in the protection of the identity, they will not take action with regards to a particular owner's plight until they are notified that something has gone wrong. Active protectors will take actions to prevent identity theft or fraud from occurring, and may actually review and observe transaction points to determine if underlying transaction patterns are suspicious. Active protectors will often contact the owner first, querying their transaction activities, and warning owners of potential identity security breaches.

The passive type of protector is associated with governmental and judicial agencies which have the potential to access data within large jurisdictions, but do not have a business or legal framework with which they can actively pursue issues. Privacy concerns of individuals are main concern in this respect. Although government agencies may have nearly unlimited access to owner identity data, access to that data must be justified within a strict legal framework. Passive identity protectors are opposed by other types of protectors including privacy protectors, and government watchdog groups. Finally, as governmental or judicial agencies supported by taxpayer funds or donations, passive protectors do not have established business models that bring returns based on the quality or vigilance of their protection.

Active protectors however, are usually in the employ of the identity issuer or possibly identity checker. These groups work to prevent breaches in the security measures of the established identity chain, and to seal these breaches as quickly as possible. They will monitor previous and possible even live transactions, looking for patterns that might indicate fraud or theft. They will also contact owners directly, when anomalies are detected, or if certain transaction thresholds are reached.

Active protectors have a very different perspective and orientation, as the success of the endeavors will often save their employers and benefactors from financial or reputation loss. An important difference between passive and active protectors is those active protectors usually work in the best interests of the issuer rather than passive protectors who work in the best interests of the owner. Active protectors will therefore often investigate owner issues, including potential misuse or error of identity data or documents. Passive protectors will only become involved in these cases based on legal infraction.

Active and passive protectors do work together however, to fulfill additional protector roles. Passive protectors are responsible for law development and enforcement. They make laws to protect other members of the identity chain. They also determine punishments for offenders and carry out the enforcement process. Passive protectors rely heavily on active protectors in the execution of these responsibilities, and active protectors provide evidence used in the prosecution of offenders. Active protectors also provide information and advice to passive protectors who then covert this into policy and law. The role is not one-sided however as passive protectors ensure that there is a balance between the interests of issuers, checkers, and owners. They also provide a voice for owners, when the protection system fails them, and owners find themselves in opposition to active protectors.

5. Conclusions

The importance of Wang *et al.*'s and Zachman's frameworks for system planning to combat identity theft is that it provides a comprehensive view of various roles and their relationships in the identity chain. We believe that combating identity theft will require coordination of identity owner, issuer, checker, and protector. We hope our work will provide a starting point for organization and system designers to include and consider issues relating to identity theft when designing its systems. Specifically, we believe that it's necessary for owners, issuers, checkers, and protectors to collaborate. The system must be designed to facilitate the collaborations. With the development of information and communication technologies, many of the collaborative tasks can be automated. Some tasks, however, must rely on human interventions and

manual processes. System planners (either designing an information system or designing organizational rules, policies, and procedure) can benefit from the framework and associated issues discussed in this paper. For organizational designers of identity issuers, checkers and even protectors, the issues discussed in this paper can be used as a discussion base when formulating rules with regard to identity information management. For information system designers, the issues discussed here can serve as a tool for system planning and system requirements. It may be helpful for them to find technical solutions based on the issues addressed.

This paper discussed identity theft issues only from the system planner's perspective. Research scope should be expanded by including owner, designer, and builder's perspectives in the future. Detailed analysis of designer's perspective will be useful for information systems design for the purpose of identity information protection.

REFERENCES

[1] G. R. Gordon, N.A. Willox Jr., D.J. Rebovich, T.M. Regan, and J.B. Gordon. "Identity Fraud: A Critical National and Global Threat", *Journal of Economic Crime Management*, 2 (1), 2004, pp. 1-48.

[2] P. Lawson and J. Lawford, "Identity Theft: The Need for Consumer Protection." Ottawa, Canada: *The Public Interest Advocacy Center Working Paper*, 2003, ISBN 1-895060-59-1.

[3] A. R. Gonzales and D.P. Majoras, "Combating Identity Theft: A Strategic Plan", 2007, Office of the President: *U.S. Department of Justice*.

[4] S. Sproule and N. Archer, "Defining Identity Theft – A Discussion Paper", Prepared for the Ontario Research Network in Electronic Commerce (ORNEC) I*dentity Theft Research Program*, April 6, 2006, pp. 1-37.

[5] R. N. Mayer, "Defending Your Financial Privacy: The Benefits and Limits of Self-Help", American *Association for the Advancement of Retirees* (Ed.), February 2006.

[6] E. M. Eisenstein, "Identity Theft: An Exploratory Study with Implications for Marketers", *Journal of Business Research, forthcoming*, 2008.

[7] K. S. Koong, L.C. Liu, S. Bai, and Lin, "Identity Theft in the USA: Evidence from 2002 to 2006", *International Journal of Mobile Communications*, 6(2), 2008, pp.199 - 216.

[8] Q. Min, S. Ji, and Q. Zhong, "An Overview of Identity Theft: Governance and Research Trend", *Journal of Public Management*, 4 (1), 2007, pp.49 – 55 (in Chinese).

[9] G. R. Newman and M.M. McNally, "Identity Theft Literature Review", U.S. Department of Justice, July 2005, http://www.ncjrs.gov/pdffiles1/nij/grants/210459.pdf.

[10] W. J. Wang, Y.F. Yuan, and N. Archer, "A Theoretical Framework for Combating Identity Theft", *IEEE Privacy & Security*, 2006, pp. 30-38.

[11] J. Zachman, "A Framework for Information Systems Architecture", *IBM Systems Journal*, 26 (3), 1987, pp.454-471.

[12] O'Rourke, C., N. Fishman, and W. Selkow, Enterprise Architecture Using the Zachman Framework, *Thompson Course Technology*, 2003, Boston.

[13] Synovate, "Federal Trade Commission – 2006 Identity Theft Survey Report", 2006, McLean, *VA*, pp.1-108.

[14] R. V. Clarke, "Shoplifting", U.S. Department of Justice, COPS, and Center for Problem Oriented Policing, 2001, Retrieved from http://www.popcenter.org/Problems/ problem-shoplifting.htm.

[15] K. Dolan, "Internet Auction Fraud: The Silent Victims", *Journal of Economic Crime Management*, 2 (1), 2004, pp.1-22.

[16] E. J. Potter, "Customer Authentication: The Evolution of Signature Verification in Financial Institutions", *Journal of Economic Crime Management*, 1 (1), 2002, pp.1-19.

[17] R. Pinheiro, "Preventing Identity Theft Using Trusted Authenticators", *Journal of Economic Crime Management*, 2 (1), 2004, pp. 1-16.

[18] M. Shreve, "Employers Slow To Recognize Identity Theft", *Business Insurance*, 38 (36), 2004, pp. 4.

[19] S.J. Wells, "Stolen Identity", *HR Magazine*, 47 (12), 2002, pp.30-38.

[20] D. Cadrain, "Liability for Employee Identity Theft Is Growing", *HR Magazine*, 50 (6), 2005, pp.35-40.

[21] N. B. Zabriskie, "Fraud by Consumers", *Journal of Retailing*, 48 (4), 1973, pp.22-27.

[22] R. E. Wilkes, "Fraudulent Behavior by Consumers", *Journal of Marketing*, 42 (4), 1978, pp.67-75.

[23] J. R. Shah and M.B. Murtaza, "Effective Customer Relationship Management through Web Services", *Journal of Computer Information Systems*, 46 (1), 2005, pp.98-109

[24] R. J. Canter, "Successful CRM Hinges on Data Quality", DestinationCRM.com, April 30, 2002, retrieved from http://www.destinationcrm.com/articles/default.asp?articleid=2177.

[25] G. Tomko, "Biometrics As A Privacy-Enhancing Technology: Friend Or Foe Of Privacy?", Paper presented at the Privacy Laws & Business, the 9th Privacy Commissioners'/Data Protection Authorities Workshop, 15 September, 1998. Santiago de Compostela, Spain. Retrieved from http://www.ct.gov/dss/cwp/view.asp?a=2349&q=304844.

[26] J. D. Woodward, "Biometrics: Privacy's Foe or Privacy's Friend?" *Proceedings of the IEEE*, 85 (9), 1997, pp.1480-1492.

[27] J. Grijpink, "Biometrics And Identity Fraud Protection - Two Barriers To Realizing The Benefits Of Biometrics - A Chain Perspective On Biometrics, And Identity Fraud", *Computer Law & Security Report*, 21 (2), 2005a, pp.138-145.

[28] J. Grijpink, "Biometrics And Identity Fraud Protection - Two Barriers To Realizing The Benefits Of Biometrics - A Chain Perspective On Biometrics, And Identity Fraud - Part II", *Computer Law & Security Report*, 21 (3), 2005b, pp.249-256.

[29] Computer Weekly, "Biometrics Accuracy Will Depend on Registration", *Computer Weekly*, 2/10/2004, pp.26.

[30] Verisign, "White Paper: The Merchant Supply Chain – Dangers, Challenges, and Solutions", 05/20/2005, http://www.verisign.com/static/005292.pdf.

[31] R. Behar, "Never Heard Of Acxiom? Chances Are It's Heard Of You", *Fortune*, 2/23/2004, 149 (4), pp. 140-148.

[32] S.H. Wildstrom, "Your Data, Naked on the Net", *Business Week Online*, 2/9/2006, pp.20.

[33] D. Gillin, "The Federal Trade Commission and Internet Privacy", *Marketing Research*, 12 (3), 2000, pp.39-41.

[34] G. Hulme, "Lack of trust hampering online direct marketing", *B to B*, 90 (12), 10/10/2005, pp.4.

[35] The New York Times, To Fight Identity Theft, a Call for Banks to Disclose All Incidents, March 21, 2007, *Business/Financial Desk*; SECTC.

[36] Financial Executive, "Computer Disposal Raises Legal Issues", *Financial Executive*, 19 (7), 2003, pp.11-12.

[37] G.R. Gordon and N.A. Willox Jr., "Using Identity Authentication and Eligibility Assessment to Mitigate the Risk of Improper Payments", *Journal of Economic Crime Management*, 3 (1), 2005, pp.1-24.

[38] Federal Trade Commission (2006), "Consumer Fraud and Identity Theft Complaint Data, January – December 2005", http://www.consumer.gov/sentinel/pubs/Top10Fraud2005.pdf.

[39] London School of Economics and Political Science, "The Identity Project: An Assessment of the UK Identity Cards Bill and Its Implication", June 2005, pp. 1 – 318, retrieved from http://is2.lse.ac.uk/idcard/identityreport.pdf.

J. Serv. Sci. & Management, **2008, 1: 153-158**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

Reconfigurable Logistics Information System Based on Soft Components Technology

Feng Liang

Dept. of Industrial Engineering, Nankai University.Tianjin 300457, China
liangfeng79@nankai.edu.cn

ABSTRACT

Logistics information system should be reconfigurable, open and extensible to support the total supply chain. Soft component technology is introduced to design and build the logistics information system oriented to the business demand of kinds of supply chains. When the supply chain is changed, logistics information system organizes all kinds of required logistics resources, by recalling predefined soft components to dynamically configure logistics information system, which can ensure quick respond to the changes of supply chain demand and support dynamic reconfiguration of the logistics information system. In this paper, a method of reconfiguration of logistics information system by using soft component technology is proposed. The key problems to the subject domain are investigated and an illustrative case study is presented.

Keywords: *supply chain management, logistics information system, soft component technology, logistics function componen*

1. Introduction

Logistics information system (LIS) is becoming important as it provides efficient and effective logistics management that aims to reduce cost and cycle time for its customers on the supply chain. LIS are flexible tools for collecting, aggregating and analyzing data from the operative applications (Purchasing, Sales, Logistics, Inventory Controlling, Plant Maintenance, Quality Management/Inspection Processing), which enable users to continually control target criteria and to react in time to exceptional situations.

Logistics management typically deals with various inbound and outbound logistics activities that involve all levels of planning and execution [1, 2, 3, 4, 5]. However, many small and medium-sized third party logistics (3PL) providers still focus on internal operation performance but lack the vision of collaborating with other supply chain participants on improving overall supply chain performance. Moreover, these 3PL providers are unable to synchronize information with trading partners in real-time for making timely decision or providing responsive services. It is unresponsive in today's digital era. New method is needed to support the business growth of 3PL providers in the coming future, which should be reconfigurable to support the different supply chain and to response different customer demands.

As a new concept, soft components technology accelerates the development of the imperative reconfiguration form [6]. The soft components technology has many advantages, such as reuse, simplified development process, and reduced development cost, improved quality of developing logistics and so on. Consequently, the logistics information system based on soft components technology has inherent capabilities for capacity change, customer variety and shorter changeover time and cost. The soft components, applied to the reconfigurable logistics information system, can link the different component by the well-defined (friendly and standardized) interface to improve the system flexibility and to accomplish the response to the different customer requirement. To support the different kinds of supply chain, logistics information systems are organized rapidly to adapt the requirement of the response in the logistics enterprise according to the current customer requirements. This approach can ensure efficient and predictable uses of the logistics resources.

The rest of this paper is organized as follows. The concepts and development of soft component technology will be presented in Section 2. The architecture and implementation of reconfigurable logistics information system is described in Section 3 and 4. A case study is presented in Section 5. Finally, conclusions are made in Section 6.

2. Soft Components Technology

The "component'' paradigm is emerging in the more specific domain of software components. [7]. From the software engineering point of view, a software component is only a unit of composition with contractually

specified interfaces and explicit context dependencies, which can be deployed independently and is subject to composition by third parties [10]. The primary goal of component concept is reuse, which presents some important advantages. For example, if 50% of the components in new software have been already efficiently used elsewhere, it means that the software developers need to develop only 50% of the new software [8]. Consequently, the developer can produce this software more rapidly compared with entirely writing it.

At present, there are trends that those theories and methods are applied to the other systems. When applied to the common systems, a component is defined as a part of an application that is developed and tested independently and integrated later into the application through simple communication [9]. Broadly defined, soft components technology focuses on the conception of plug and play soft components which can be used immediately after embedded as the center, and thus the system can be built by combining components. It is unnecessary to construct entirely the new software system, which is very different from others. Reversely, using the existing soft components directly, the new system can be assembled (or be modified within reason moderately). So the soft components technology has two aspects as follows:

- Software Components: Corresponding with the component conception of software system, the technology of software components such as JavaBeans and COM etc. is adopted at the area of application development. Because those technologies are mature and commercial, they can rapidly accomplish the construction of a new system. The benefits of legacy software adaptation come from simplified testing and more controlled software modifications with a view to future software system changes.

- Flexible Components: From the system structure point of view, the whole system is composed of flexible components dynamically. The structure of each component can be changed and thus flexible. A component can both embed other components and also be embedded into other components. The granularity of component is relevant to the actual instance specification.

Unlike other systems, systems based on soft component are designed to be installed with the exact logistics capacity and functionality needed, and to be upgraded (in terms of capacity and functionality) in the future. Expanded capacity and functionality enables the logistics of more complex part types and logistics a variety of part types on the same system. Compared with the traditional reconfiguration technology, soft components technology has the following characteristics:

- Plug and play. One can integrate components into the frame conveniently, without either modifying code or compiling again. But he must verify its specifications' coherence with the new environment

- Independent. The interface of component is separated from the accomplishment of the component. Thus, the system designer only needs to take care of the interface, without knowing the implementation details. It also provides an encapsulation and execution environment for logistics systems. The well-defined interface provides encapsulation and a uniform view to the overall software design.

- Customization. The developers design the system capability and flexibility (hardware and controls) to exactly match the application, which represent the ability to adapt the customized (non-general) flexibility of logistics information systems to meet new requirements with orders.

Considering the above advantages of soft component technology, we introduce it to the reconfiguration of logistics information system. From a system-engineering point of view, the key issue with component-based logistics information system architecture is the efficient utilization of the logistics resources in data-intensive applications.

3. Reconfigurable Logistics Information System

The reconfigurable logistics information system adopt the method of "building blocks" to organize logistics information system, that is, a new system is developed by using existing components as building blocks. Firstly, the logistics resources are designed to be individual toy brick. Then, when the order arrives, user can organize "building block" to form the logistics information system in terms of the logistics order plan. When the logistics task plan changes, one can manage and control logistics resources with the certain rules, and replace some portion of "building block" or add other "building block" to adjust the system structure to respond to the change of logistics dynamically. From the modeling technology perspective, those "building block" are soft components presenting logistics resources, which can constitute many logistics information systems. For different applications, different constriction conditions are established according to the existing logistics resources, and different resources can be optimized and reconfigured to form the optimum logistics information system. When there may be several feasible candidates, one needs to choose a most appropriate one.

Soft components technology applied to construct the logistics information system has the following characteristics:

- Separates the system designer from the model builder of logistics resources, thus the designer do not need to consider the specific logistics resource model. The system designer and soft component developer may work together in concurrent way.

- Enhance the system flexibility. Given logistics tasks, the required system can be constituted by the corresponding soft components stored in the components library. And the system does not need to be designed entirely, which ensure that the similar function parts are not repeatedly written. As thus, it not only enhances system flexibility but also improves the planners' efficiency.

- Elevate the system openness. By hiding and encapsulating the structural details of data and algorithms that underpin the logistics system and logistics resources, and by supplying a specification interface to the user, the soft component technology offers a high openness. But in integrating a component into a new environment, one must verify its specifications' coherence with the new environment, which ensures that the function is not disrupted by its use mode. [10]

- Improve the system agility. When the condition changes or the resources conflicts, one can schedule the logistics resources effectively, and reconfigure the logistics system rapidly.

Therefore, using the soft components technology, the logistics information system can realize the dynamic change of system structure to respond the change of supply chain environment and logistics task, which makes it suitable to the different type of logistics applications. Furthermore, along with the above procedure, the domain-public soft components library will be formed gradually, which can be applied to the different applications to amplify the system generalization.

4. System Implementation

It is very complicated that how to design reconfigurable logistics information system. Figure 1 shows that the development procedure of soft component-based system is a concurrent and two-dimensional procedure, where the system planner is separated from the model developer of logistics resources. From the system planner perspective, it follows the common development procedure, namely requirement analyzing, system simple design, system detailed design, assembling & testing, and logistics delivery, etc. In this step, close-loop control is introduced to feedback the required information, which contributes to the designers to solve the problems arisen in practice. On the other hand, from the soft component developer perspective, it represents soft component development procedure, which consists of group classifying, characteristic distilling, and component test. In this mode, soft component technology encapsulates machines as virtual devices, making the architecture more understandable for software developers. In the following section, we will discuss the two procedures in detail.

4.1. System Planning

To develop soft component-based system, it is necessary to divide the application into components. So the first step to be considered is to have an investigation on domain analysis, which helps to the understanding of a problem and its solution components. This analysis allows for identifying and classifying by types of themes treated, which plays a key role in the identification of soft components. From the analysis, a list of domains that characterize the activities of logistics task planning is established. At his stage, the whole system is divided into three parts: common part, reusable part (which is relevant to some domain) and special part. As a consequence, the components can be classified into the following three classes based on functionality: Generic Component (GC), Domain-Generic Component (DGC) and Domain-Specific Component (DSC). The design steps are discussed as follows:

Build genetic component.
It is obvious that there are some basic parts in a common system, which are to provide some genetic service, such as database management part, network connect part, etc. These generic components could be used as basic service to build other systems, which are genetic elements independent of the considered problem and specific elements dedicated to the developed system [7]. These generic models are obtained after an analysis of a domain, which make the architecture adaptable to various kinds of applications. It is horizontal reuse, which consists of using generic components in different applications.

Design domain-generic component
Domain-generic component is a special component that contributes some genetic services to a special domain, which is only applied in some special applications. The group technology (GT) is introduced to classify the logistics resources frequently, which is based on the principle of grouping similar parts into families. Given different orders, the different components are assembled to construct the oriented- order system. During this period, the particular information flow of a special order need not be considered, because the following domain-special components should be assembled to meet the special demands. This is a vertical reuse which means that components are specific to a domain of expertise or implementation. Its objective is to derive generic models from a family of systems in some domain.

Develop domain-special component
The goal of domain-special component is to provide some special services for a special order in a special domain. After investigating the enterprise and building the above basic module, the developer can design the domain-special components to construct the oriented-order logistics system.

The steps of system design flow are illustrated as follows: when the logistics task arrives, job shop controller assign the tasks to the corresponding system controller, which will form the oriented- order logistics system in terms of adding component directly, selecting appropriate

component and recompiling component. Using the special communication channels, the system controllers can send massage to the workstations, which directly control the corresponding equipments.

Appling this construction mode, the logistics information system can be capable of handling a variety of logistics tasks efficiently. Ideally, if the components constituting the logistics information system have form a certain scale, which mean that a large structural frame is considered as the invariant around which small variant functions can be added (i.e. modifying code lines), the components in this system can be regarded as a total collection to be stored in component library. In addition, the relation among the components is dynamically allied. When the order arrives, the required components are selected and locked to construct logistics information system with other components; after the tasks fulfilled, the relations are broken automatically, and the components are unlocked and free to be applied to other systems.

4.2. Soft Component Design

The soft component design is a key problem of soft component-based system development. In this paper, the components in logistics systems are classified two classes as follows:

Figure 1 illustrates the structure of logistics function component, which has three layers. Information layer includes the basic characteristics of logistics resource; application layer represents the behavior of component to finish the actual process function; the information integration technology is applied to the design of interface layer to ensure seamless conjunction among the components, which support the plug and play of component, and realize the reconfigurable logistics information system.

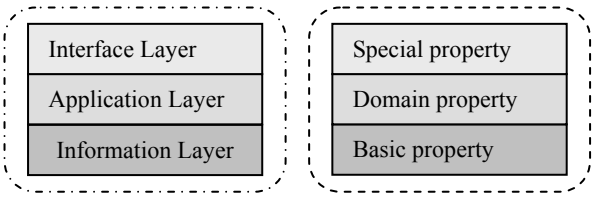


Figure 1. Logistics function component design

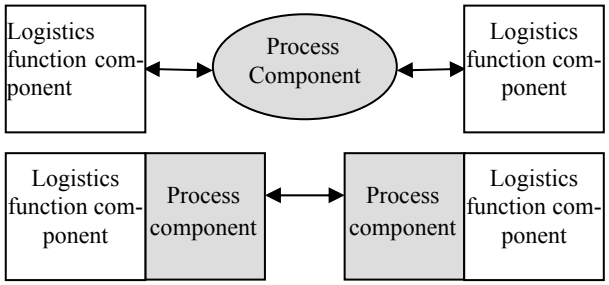


Figure 2. The relation between logistics function component & process component

4.2.1. Logistics Function Component

To design these soft components, we can draw up and sum up existing logistics resources, abstract and mimic some of the traits of logistics parameter, capsule by object-oriented and aspect-oriented concept to construct soft components, etc. Different components in the same domain can be derived or inherited from the domain-generic components. When the structures of existing components need to be modified, we should inherit from the former components as far as possible to keep consistency of the components structure.

4.2.2. Process Component

It is different between process component and logistics function component, which is that the farmer's operating object is not the detailed data or information but some logistics function components. The process component represents the changeable relation between the logistics function components, such as the coordination and the connection among the logistics resources. Process component not only can be regarded as a single component in the component library, but also can form another component by combining with other logistics function components. Figure 3 shows the relation between the logistics function component and the process component.

In a systemic point of view, the whole logistics information system is looked as a network by dividing the logistics information into the logistics function component and the process component, where the logistics information component is the "node" and the process component is the "chain" to connect the nodes, which deal with the total logic process procedure jointly. So the lo-

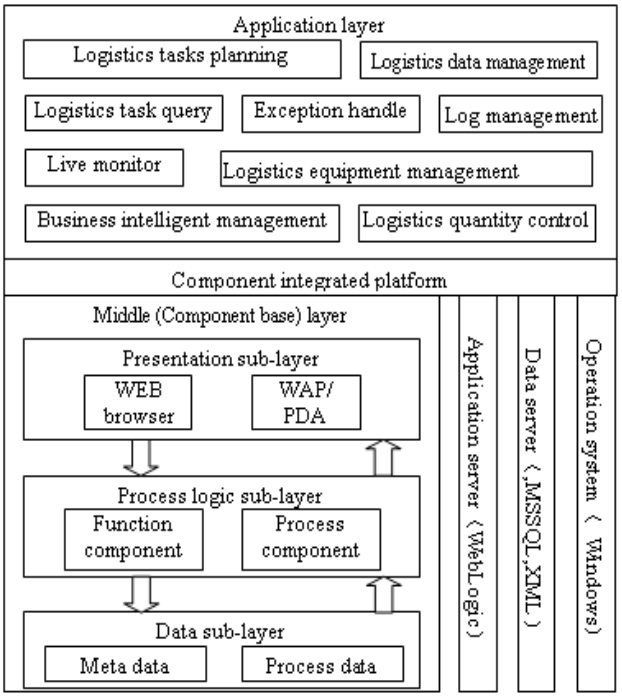


Figure 3. The structure of reconfigurable logistics information system

gistics information system oriented order is formed due to the different combination of logistics function component and process component.

5. Software System

The goal of this work is to build a reconfigurable logistics information system to improve the system's capability to respond to changing customer requirements. This model is based on the soft components technology previously described and interactions between components allowing their integration. The system based on the structure will support the reconfiguration among various heterogeneous components and remote logistics control. Logistics information system accomplishes the integration of each process layer based on this structure, and is reconfigured according to the process flow. The system is programmed in Java of J2EE mode, in which the components are expressed in the form of JavaBean.

The model is built with three layers including information layer, middle layer and application layer to improve system capabilities as shown in Figure 3.

(1) The first one is responsible for the relational information in the system, which includes the management of user information, order information and logistics resource information, etc.

(2) The second layer builds the object model from a collection of relational data. This object model constitutes the basic platform on which different soft components build their own specialized object models for their specific business processes. Middle layer is composed of three sub-layers: presentation sub-layer, process logic sub-layer, and data sub-layer, which communicate with application layer by component integrated platform.

(3) The third layer presents information to the users and interacts with them. The application layer offers the following functions: Logistics tasks planning, Logistics data management, Logistics task query, Exception handle, Log management, Live monitor, Logistics equipment management, Business intelligent management, Logistics quantity control etc.

6. Case Study & Discussion

This case study illustrates the detailed procedure for constructing a logistics information system and completing the logistics task.

(1) When a logistics task named "Task1" arrives, the task controller is created to manage the complete procedure until the task is completed.

(2) With the function of "Component Organization" provided by the system, different components are contained in the task controller. In the instance, several logistics function components and process component are selected to construct the logistics information system to deal with the logistics task. Furthermore, it is determined that the task is to be completed by the logistics term named"lt_1".

(3) When the task controllers in ""lt_1", which includes 10 logistics tasks completed in 8 transportation vehicles, have constructed the corresponding logistics information systems, there should be a schedule on the transportation sequences of all the logistics tasks. In this case study, a heuristic algorithm is applied to schedule task-vehicle at the phase of component organization, because the scheduling algorithms are encapsulated as process components. However if there are different algorithms selected by different tasks in the same logistics term, some indices are introduced to compare the difference between these algorithms, such as transportation distance, due date etc. Finally, the logistics term will adopt the optimal algorithm to solve the task-vehicle scheduling problem.

(4) During the period of transportation, with the help of the RFID technology user can scan the procedure of tasks in transportation in the module of "live monitor". Furthermore, the task state will change according to the feedback information from the vehicles. When the task state is "Done With", which means the logistics task has been completed, the task controller should be destroyed automatically and the relevant components should be released. All the procedure will be recorded in details in the system log.

To construct the reconfigurable logistics information system, the significant obstacle is to abstract and build the component model from the actual logistics activities and logistics enterprises. Considering every enterprise has the own types of logistics operation and activity, If the component model can not express the actual logistics activities or the granularity of component is not enough fine, the system cannot to organizes the predefined soft components to support the rapid response when the supply chain changed. Furthermore there are also risks involved in integration of the components. This is due to the fact that the current logistics systems became complex increasingly. Therefore, this is also a new method that requires further research and development in certain key issues.

7. Conclusion

In this paper, we have proposed a methodology for reconfigurable logistics information system, which make an important investigation to build the logistics information system to support kinds of supply chain. Having briefly defined the main concepts of soft component technology, we have clarified two aspects: software component and flexible component. Then, we have proposed the architecture for logistics information system based on soft component technology, which use three layers affect the application capabilities. In some sense, it is specific to the reuse of soft components and its interactions, which are intensively tested in a wide variety of situations. With an

example, the rapid reconfiguration of logistics information system based on soft components technology is certified to be a useful method to reconfigure new systems rapidly and flexibly, which decreases the development cost and time, and consequently improves the planners' efficiency.

8. Acknowledgment

The work described in this paper was substantially supported by a Humanities Science Research Youth Foundation from Nankai University (NkQ07055)

REFERENCES

[1] Amami Mokhtar, Brimberg Jack, Web-based interorganizational information systems for logistics outsourcing, *Annals of Telecommunications*, 58(1-2), 2003, pp.266-296.

[2] Helo Petri, Szekely Bulcsu, Logistics information systems: An analysis of software solutions for supply chain coordination, *Industrial Management and Data Systems*, 105(1), 2005, pp.5-18.

[3] Jiao Jianxin, Lee Alan Chee Guan, Chin Kee Min, Chang Chia Shian, Ma Qinhai, BPR-based information systems planning for e-logistics: A case study, *International Journal of Services, Technology and Management*, 4(4-6), 2003, pp.384-407.

[4] Dong Qianli,Research into the Strategy of Third Party Logistics Based on Supply Chain Management,China Soft Science, 10, 2000, pp.34-37.

[5] Deng guiying, Solution of the third-party logistics application based on Web Services,*Journal of University of Shanghai for Science and Technology*, 28(3), 2006, pp.290-298.

[6] A. L. Azevedo, J.P. Sousa, A component-based approach to support order planning in a distributed manufacturing enterprise, *Journal of Materials Processing Technology*, 107, 2000, pp.431-438.

[7] Philippe Aniorte, A distributed adaptable software architecture derived from a component model, *Computer Standards & Interfaces,*25, 2003, pp.275–282.

[8] C. Szyperski, Component Software, Addison-Wesley, Reading, MA, 1997.

[9] M. Jenko, N. Medjeral, P.Butala, Component-based software as a framework for concurrent design of programs and platforms, *Microprocessors and Microsystems* 25, 2001, pp.287-296.

[10] S. Pierrea, N. Nouisser, Reusing software components in telecommunications network engineering, A*dvances in Engineering Software*, 31, 2000, pp.159–172.

J. Serv. Sci. & Management, **2008, 1: 159-164**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

Technology Innovation Stratagem and Role Played by MNCs: Evidence from Chinese Pharmaceutical Sector

Li-Li Cao[1] & Xiang-Dong Chen[1]

[1]*The School of Economics & Management. Bei Hang University, Beijing, 100083, P. R. China*
Email:caolili31083125@sina.com

ABSTRACT

It is widely regarded that the pharmaceutical sector is quite different from other manufactures for its character in innovation activities. This paper analyzed the nature of innovation activities by multinational companies (MNCs), which stand for foreign capital, focusing on two parameters named: size and scope of innovation through investigating patent data from 1985 to 2005. The paper reveals the role of the foreign capital and identifies their technology stratagem by developing a tool named Double- S Matrix. Research finding indicates that foreign capitals play an important role on overall national innovative competitiveness in this sector. Implications from this research can be used as reference to policy makers and business companies in China.

Keywords: *size of innovation, scope of innovation, Double- S Matrix*

1. Introduction

It is widely believed in international literature that innovation movement in pharmaceutical field is unique if one compares with other manufacture sectors. The major reasons for this are frequently listed as typical R&D intensive sectors, more explicit knowledge based technologies, and one of the industries highly interdependent among larger MNCs and local SMEs in high tech sectors, especially, due to discontinuity nature in innovation process accompanied with emergence of bio-pharmaceutical technology in 1970's [7].

Technology development route in pharmaceutical industries thus can become more versatile or diversified, as Bogner and Thomas's work [4] indicated (as four different technology trajectories based on historical facts in pharmaceutical sectors in the US). Bogner and Thomas [4] therefore considered that strategic innovation and its classification were the major issues for firms in the industries, especially for larger MNCs. The sophisticated nature in innovation strategies in firms as well as in countries in the whole in pharmaceutical sectors can be best described by Hara's hypothesis [8]. He emphasized four basic forces regarding to input of R&D activities and outcome of the R&D process, namely, molecule-biological (technological) learning, interact factors during application process, market forces, and strategic roles of organizations.

The objective of this paper is to produce new empirical evidence pertaining to the characters of technology competition by MNCs. Based on patent data; this article holds several original features. Firstly, investigate the characters of innovator from a macro-level, data based on countries level was employed to reflect the innovation by MNCs. Secondly, developed a tool to analysis the role of the foreign capital and reveal their stratagem.

2. Literature Review and Research Framework

2.1. Innovation in the Pharmaceutical Sector

Due to its exceptionally high R&D to sales, the pharmaceutical industry is often characterized as a technology- and science-driven sector. The radical innovation in drug industry is a function of the new technological and scientific knowledge embedded in the drug development [1]. The current research on innovation in the sector falls into four major aspects. Firstly, regional character is more significant in drug industry. Previous research shows that there is a intensive demand of regional proximity due to the science-based nature [12]. Secondly, it is a complex progress of innovation due to a long industry chain and diversiform actors. Most of scholars considered that innovation in the sector reflects breakthrough of innovation chain, which is also decided by the science-based character [10]. Thirdly, patents are a good measure of inventive activity [6]. The pharmaceutical sector has a high propensity to patent relative to other industries [17], also correlated with ratings of basic research excellence [18].

2.2. Technology Competition Based Diversification

There are many literatures to reveal the character of technology competition. Many of them focus on the two parameters: size and diversification, as well as their relationship [3]. It wildly believed that larger countries have a

tendency to spread their research activities across many technological fields while small countries tend to concentrate on narrow niches. Later, Cantwell [7] used the data of 24 famous MNCs, focused on the size-diversification relationship has been stable over time or has changed in different historical periods, shows how the nature of this relationship has changed historically. Castaldi [6] analysis diversification patterns of Italian manufacturing firms, offered empirical evidence on a robust relationship between firm size and diversification. Breschi [5] analysis the diversification of innovation activates by identified two type of innovators: diversified innovators and specialized innovators . *H*owever, The theoretical and empirical literature has paid extensive attention to large firms' technological diversification of technology competition [11]; the equivalent phenomenon has been largely neglected and there are paucity of research at a country level.

3. Empirical Methodology and the Data

3.1. Hypothesis

In China, it is necessary for local firms to innovation to increase added value of products. It is important to study the innovation disciplinarian of the sector could promote industry's development and renew the structure of industry. However, the innovation disciplinarian must be influenced by foreign innovation activities because of the nature of FDI in China's pharmaceutical. This paper provides a research on the foreign patenting activities based on three main hypothesizes.

Hypothesis 1: Technology resource competition is determined at least by two parameters: size of technology and scope of technology.

Hypothesis 2: In Chinese pharmaceutical sector, there are two kinds of innovators, one is local firms and the other is MNCs invested by overseas capital. The two kinds of innovators interact with each other to control more technological resource. One side, they pursue the quantity of patent to ensure they have the expert advantage in certain field; the other side, they explore the technological scope to achieve the diversification effect.

Hypothesis 3: The diversification effect mainly comes from two aspects. One is from the diversification among the different product portfolio like the diversification between chemical preparations and biochemical medicines. The other is from the different process technology in producing a certain product. For example: the diversification between C07K and C12P which both belong to the technology for biochemical medicines.

3.2. The Empirical Methodology

Size of innovation

It is not our objective to examine the overall effect of region size upon specialization in general (e.g. in new production and trade) , but rather to concentrate on the innovation effort. Therefore, size here is related only to the technological activities carried out by regions in China's pharmaceutical sector. This paper measured innovation through the number of patent. However, patent data is a kind of count variable that obey the Poisson distribution. So we defined it in $size_i = \log TP_i$ where TP_i represents the accumulative patent in i th period.

Scope of innovation

It is demonstrated that the entropy value is an effective measurement of the degree of diversification. This measure takes into account both the number of technologies in which a region might be active, as well as the relative distribution of technological activity across the technologies. Thus, this study will also utilize the entropy index to analysis the technological diversification among the sample in order to analysis the scope of technology resource. Therefore, we defined: $entropy_j = \sum_{i=1}^{n} p_{ij} \ln \frac{1}{p_{ij}}$, where p_{ij} in the present context represents the share of patents in China's pharmaceutical sector accounted for by the i th technology. The value of the entropy measure ranges between zero and $\ln n$, where a value of zero represents a region concentrating on one technology only and a value approaching $\ln n$ represents a region with an even distribution of patents.

One distinct advantage of using the entropy measure is its addictive properties. Specifically, the measure can be decomposed into elements, which define the contribution of technological diversification at various levels of aggregation to the total. The analysis in the present study makes use of the additive by examining technological diversification at two related level of aggregation.

The first level of aggregation follows the classification used by the China Patent Office, which make a distinction between 15 fields we selected such as A61K and C07D and so on. it is possible to distinguish between rather narrowly defined technological capability and diversification. It is a reflection of the total diversification both from the product various and process discrepancy. The second level aggregates the classes of technology into the 6 subsectors. In this classification, patents might fall into fields like biochemical medicines, Chemical preparations and so on. it could provide information about distribution of products among different regions. It is a proxy to reflect the scope of product technologies. Considering the addictive properties of entropy, the diversification of process technologies could be obtained as follow: $entropy_{i2} = entropy_{iT} - entropy_{i1}$, where the $entropy_{iT}$ Stand for the entropy of total 15IPC classes, $entropy_{i1}$ Stand for the entropy of 6 sub-sectors which represent diversification of product, $entropy_{i2}$ Stand for the diversification of process technologies.

And the main body of diversification will be defined as

the scope of technology:

$$scope_i = \max\{entropy_{i1}, entropy_{i2}\}$$

In order to investigation the role of foreign capital in Chinese pharmaceutical sector, a hierarchical cluster was employed. In a hierarchical cluster problem, a set of sample ($S = \{S_i, i = 1,2,...,n\}$) is to be clustered with respect to a set of criteria ($C = \{C_i, i = 1,2,\ldots,n\}$). Therefore, a $n \times m$ decision matrix can be obtained as

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

In this article m is 2, and n could be 10. Matrix X needs to be normalized as $Z_{ij} = \dfrac{X_{ij} - \overline{X_j}}{S_j}$

$i = 1,2....10;\ j = 1,2 \qquad i = 1,2....10;\ j = 1,2$

As a consequence, a normalized decision matrix is obtained. Then the distance was defined by the Ward method. The finally step is take use of the SPSS software to obtain the consequents. It could cluster the sample into fixed clusters, which represent the different role of the foreign capital; it will correspond to the analysis by our tool named Size-Scope Matrix (Double-S Matrix). It is showed in Figure 1.

Tech-leaders usually distinguish themselves not only by possessing the highest size of patenting but also the broadest scope of patenting. According to this, they are companies with high potentials to dig the technology capabilities in a broad product scope. Tech-widen types are characterized by a comparative small size and broad scope. They represent that kind of samples that aim at expanding the width of the product technology scope and allocated their technology capabilities average among the scope.

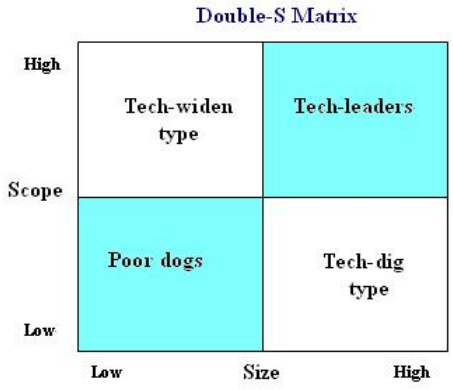


Figure 1. Double-S Matrix

Tech-dig types have a contrary nature compared to Tech-widen types. Samples belong to this type have a high enthusiasm in patenting in relative smaller scope. They aimed at digging a few kinds of technology lengthways to attain the specialized advantages. Poor dogs represent the samples that play a comparatively piddling role in Chinese pharmaceutical sector. They take a conservative attitude in patenting in China and invest inadequately.

3.3. The Sample and Data

Data based on countries level was employed to reflect the innovation by MNCs. Considered the character of the pharmaceutical sector which depend tightly on the government regulatory, historical period from 1985-2005 is then divided into four sub-periods which shows in Table 1.

15 IPC fields corresponding to pharmaceutical was selected which account for 82% of the total patent in the sector. The explanation of this technology field is list in Table 2; furthermore, the IPC classes were sorted related to industry classes, which are a sub-sector of pharmaceutical.

4. The Analysis of the Results

4.1. Study on Size of Innovation

Pharmaceutical sector is always thought to be a filed that attracts most FDI in china. Chen Xiangdong [19] investigated that the patent activities performed by FDI companies are most concentrated in Pharmaceutical sector. The patenting by foreign patenting is increasing over time as showed in Figure 2.

Thus, there are two main innovators in pharmaceutical

Table 1. The division of time periods

period	Divided criterion
1985-1991	*Patent Law Of The People's Republic Of China* entered into force in 1985
1992-1997	signed *Memorandum of Understanding on the Protection of Intellectual Property*, modified *Patent Law* in 1992
1998-2001	21 of the top 25 international companies have set up FDI companies in China
2002-2005	China joined WTO formally on Dec. 11, the second revision of patent law was put into practice.

Table 2. The technology fields

industry classes	IPC Classifications
chemical preparations	A61K
medicinal chemicals	C07D, C07C, C07H
biochemical medicines	C12N,C07K,C12P
drug-like compounds	C08L,C08G,C01B,C08F
pharmacy and Therapy	C12Q,A61B
medical apparatus	A61F,A61M

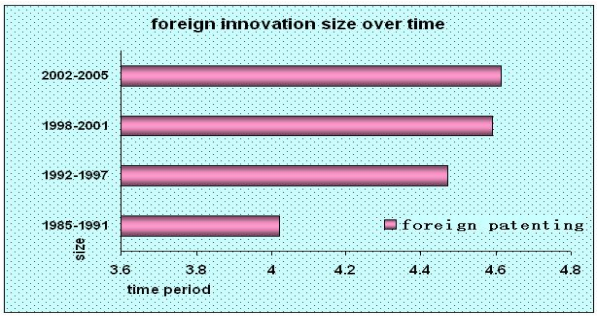


Figure 2. Foreign innovation sizes over time

sector, MNCs and local companies. It is essential to distinguish foreign patenting by FDI companies and local patenting by local companies, because it could not only indicate that the structure of innovation, but also identify major force of innovation in Chinese pharmaceutical sector. Evidence from data suggested that foreign patenting played an important part (percentage is 71%, 64%, 49%, 42% over time), however, there is a decreasing tendency of foreign percentage over time. This tendency could not influence us to draw a judgment that it is lack of indigenous innovation capacity since critical technologies in half is controlled by foreign capital. The statistic evidence could not support the point that China has an advantage as Host County.

4.2. Evolution Character of Scope of Innovation

For the foreign capital, Fig 3.shows that there is a continuous increasing in diversification over time, while the counterpart shows an obvious fluctuation of indigenous patenting resource in pharmaceutical sector.

Evidence from the Wilcoxon Signed Ranks shows that there are positive changes among the four phases, however, the changes among the first three phases are positive but not significant (Sig. =0.114), only the technology scope of 2002-2005 period is significant changed to that of 1998-2001(Sig. =0.007). Investigate from the whole period; the increasing in scope is significant statistically (Sig. =0.013).

From the patent law passed to protection intelligence property right in 1985, foreign capital turned eyes to Chinese pharmaceutical market and take limited use of the law to protect their innovation for the shortage of the law. Therefore, the size and scope of foreign capital is smallest among the 4 periods. When China revised the patent law and implemented Regulation on Protection of Medicine, the Chinese medical market became regular gradually. Until 1998, twenty-one MNCs among the top 25 pharmaceutical ones had invested to China, mainly through founding overseas-invested enterprises to seize market; however, considerable part of their patenting is not produced in China. When China entered into WTO, it is more attractive for overseas capital because of China's signature on some related medical terms. Moreover, since the Danish corporation NovoNordisk set up R & D center

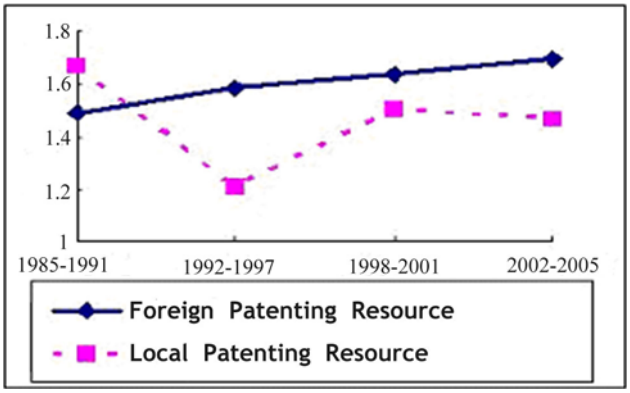


Figure 3. Scope of innovation over time

Test Statistics[b]

	entropy2-entropy1	entropy3-entropy2	entropy4-entropy3	entropy4-entropy1
Z	-1.580[a]	-1.580[a]	-2.701[a]	-2.497[a]
Asymp.Sig.(2-tailed)	.114	.114	.007	.013

Figure 4. Results of Wilcoxon Signed Ranks. a: Based on negative ranks; b: Wilcoxon Signed Ranks Test

in BeiJing in 2002, it was followed by many other MNCs. It is obvious that MNCs turned investment on R&D activities instead of operation activities. These R&D centers have the liabilities to develop new drugs directly, which could enter more sub-sectors and seize patenting resource conveniently.

In the first three phases (1985-1991, 1992-1997, 1998-2001), foreign capital aimed mainly at market, therefore, they patent to keep their market shares. With China regulate the related laws step by step, foreign capital patent their technologies dramatically to pursue the advantage along the route of their existent products, In the latest phase (2002-2005), foreign capital aimed at acquisition of integrated advantage including not only about market shares, but also critical technologies. Changes of concept prompted MNCs entered more sub-sectors and enlarged their technology scope.

4.3. Innovation Stratagem and Role Played by Foreign Capitals

The further research is investigation about the role of overseas capital to identify their technology stratagem. According to the Double-S Matrix (Figure.1), the role of overseas capital could be analyzed in size and scope.

A Cluster method was employed to divide the samples into four clusters. If the center of gravity of each cluster falls into the four quadrants defined in the Double-S Matrix, it could confirm the validity of the Double-S Matrix. Employed the software (spss) to perform the hierarchical cluster, take the number of Cluster as four, then compute the four cluster's center of gravity to scatter them on the scatter plot of clusters, compare the result to the scatter plot of samples.

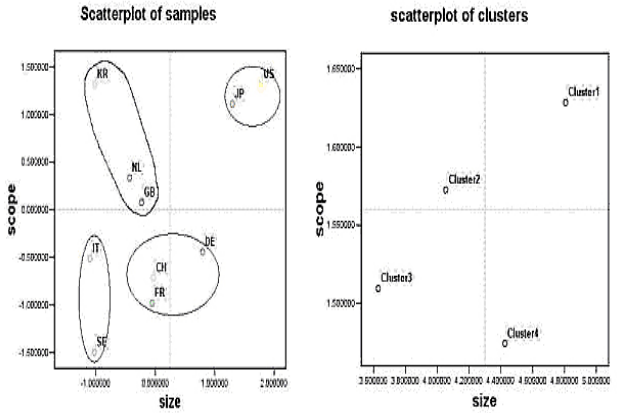


Figure 5. Comparative scatters

Figure 5 shows the result from the Cluster method is accordant with the four types defined by the Double –S

Matrix. It is apparently that four clusters belong to four quadrants to confirm the validity of the Matrix. Hierarchical cluster offers a class statistically corresponding to the four roles defined by the Matrix. It is list in table 3.

In Chinese pharmaceutical sector, America and Japan process the plentiful technology resource in most of the sub-sectors. They play a significant role as the tech-leader in patenting activities in Chinese pharmaceutical sector. Usually, they have absolutely advantage in most of the sub-sectors, while they have strength to carry out integrated innovation strategy to expand the width and depth of the technology. Tech – dig types with Represent of German, mainly concern about a few sub-sectors, which is important or familiar to them. They dig the potential technology capability with motivation to obtain the expert status and specialized technologies in some special sub-sectors which they have related competition advantage; therefore, they could capture monopoly profit by leading the dominant paradigm or establishing the technology criterion of these sub-sectors. Korea is the typical country, which belongs to the tech-width type, which has characters of using a relative smaller size to cover a relative broad scope of technology. They dabble in a broad filed and seize technology resource as broad as possible. They stimulate by diversified technology in order to achieve the scope economy, through integrating the operating and marketing of their diversified technologies and products. Refer to the poor dogs, they have a narrow scope and small size compare to the other three types. It could suggest that they hold a conservative attitude to the Chi nese pharmaceutical sector. However, they only keep the appropriate size and scope of technology to ensure their qualification to take part in this sector. If they have the expectation to seize the dominant technology in some sub-sector, they could increase their innovation activities to be tech-dig type. If there are business opportunities in many sub-sectors, they could implement a diversified stratagem to obtain profit.

Table 3. Role of foreign capital

Cluster	Countries	Type
Cluster 1	America, Japan	Tech-leaders
Cluster 2	England, Holland, Korea	Tech-widen types
Cluster 3	Sweden, Italy	Poor dogs
Cluster 4	Switzerland, France, German	Tech-dig type

5. Main Founding and Conclusion

This paper investigated in the size and scope of technology as well as originally developing the double-S Matrix to distinguish the role of foreign countries, based on the patent data from 1985 to 2005. There are two main founding this paper achieves:

Firstly, foreign capital played a critical role in pharmaceutical sector in China, with respect to the size and scope over time. Secondly, there are four types of foreign capital defined by the double-S matrix. Accordingly, the largest countries, which called Tech-leaders, were always digging the technology capabilities in a broad product scope. While, smaller countries face two stratagems to select: technologically specialized in existed sub-sector with related comparative advantage, or expanding the product technology scope in a low technologically potential.

These findings carry some suggestions for public policy. Governments should understand that the role of the foreign capital. Public policy should be, therefore, tailored according to the condition. One side, the government should attract foreign capital to invest in China to obtain the technology transfer and other profit. According to the previous literature, the effect of technology transfer is depend on the absorb capabilities to a large distant, therefore, the new role of governments should be a “god father” to support education and training, public research and universities, and to encourage firms to invest in research. It will give indigenous innovation support, including finance to R&D activities and special funds support. These actions could stand for public knowledge base; enhance the attraction of the region and development of China’s pharmaceutical sector. Further, it will foster some pharmaceutical cluster, which embody MNCs in the innovation network, and make full use of the MNCs to pull the local innovation in the global economic.

REFERENCES

[1] Abernathy, W. J. and K. B. Clark. "Innovation: Mapping the winds of creative destruction." *Research Policy* 14(1), 1985, pp. 3-22.

[2] Arundel A., I. Kabla. What percentage of innovations are patented? Empirical estimates for European firms. *Re-*

search Policy 27(2), 1998, pp. 127-141.

[3] Archibugi, D. and M. Pianta. "Specialization and size of technological activities in industrial countries: The analysis of patent data." *Research Policy* 21(1), 1992, pp. 79-93.

[4] Bogner, W. C. and H. Thomas, 《Drug to Market: Creating Value and Advantage in the Pharmaceutical Industry》, *New York: Pergamon,* 1996.

[5] Breschi, S., F. Lissoni, *et al*. "Knowledge-relatedness in firm technological diversification." *Research Policy* 32(1), 2003, pp. 69-87.

[6] Castaldi, C., M. Milakovic, *et al*. Scale and technological factors in the diversification structure of business firms. *Economics Letters* 91(1), 2006, pp. 117-121.

[7] Cantwell, J. and Vertova. Historical evolution of technological diversification. *Research Policy* 33(3), 2004, pp. 511-529.

[8] Hara, T., 'Industrial Technological Trajectories and Corporate Technology Traditions: the Development of Antibacterial Drugs in Japan', *Annals of the School of Business Administration*, Kobe University, 41, 1997, pp. 1-18.

[9] McMillan, G. S., F. Narin, et al. An analysis of the critical role of public science in innovation: the case of biotechnology. *Research Policy* 29(1), 2000, pp. 1-8.

[10] Marx, L. and M. R. Smith. Introduction [A]. in M. R. Smith and L. Marx (eds) , *Does Technological Determinism* [C]. Cambridge,MA: M IT Press, 1994. pp. 9 - 15.

[11] Piscitello, L.. "Relatedness and coherence in technological and product diversification of the world's largest firms." S*tructural Change and Economic Dynamics* 11(3), 2000, pp. 295-315.

[12] Swann, Peter and Prevezer, Martha. A Comparison of the Dynamics of Industrial Clustering in Computing and Biotechnology. *Research Policy*, 25(7), 1996, pp.1139-1157.

[13] Tushman, M.L., Anderson, P.C., 'Technological discontinuities and organizational environments', *Administrative Science Quarterly*, 31, 1986, pp.439–465.

[14] Tushman, M., P. Anderson. Technological discontinuities and organizational environments. *Admin. Sci. Quart*.31, 1986, pp. 439-465.

[15] von Zedtwitz, M. and O. Gassmann, "Market versus technology drive in R&D internationalization: four different patterns of managing research and development, *Research Policy* 31(4), 2002, pp. 569-588.

[16] Zander, I. Technological diversification in the multinational corporation--historical evolution and future prospects. *Research Policy* 26(2), 1997, pp. 209-227.

[17] Taylor, C. T., Silberston, *Z.A. The Economic Impact of the Patent System: a Study of the British Experience* [M]. Cambridge University Press, Cambridge, 1973.

[18] Pavitt. Technology competencies in the world's largest firms: complex and path-dependent, but not much variety [J]. *Research Policy*, 26 (2), 1997, pp. 141-156.

[19] Chen xiangdong, Caolili, Comparative Study on indigenous Innovation resource in Pharmaceutical Fields: An Entropy Measure over Patent Data in China[C]， *international symposium on management of technology* 2007, pp.817-821.

J. Serv. Sci. & Management, **2008, 1: 165-171**
Published Online September 18, 2008 in SciRes (www.SRPublishing.org/journal/jssm)

The Relation Analysis of Equity Structure and Its Performance on Chinese Listed Bank

Chang-jun Zheng[1], Zhou Yan[2] & Li Wei[3]

School of Management of HUST, Wuhan, Hubei, P.R. China
Email: [1] zhchjun@hust.edu.cn, [2]zoeyet@163.com, [3]tianyoulw@yahoo.cn

ABSTRACT

With the development of reform of stock system among Chinese commercial banks, more and more listed banks are coming forth to the market, so it is important to analyze the relationship between the equity structure and its performance of Chinese listed bank. For representatives, this article selected 6 banks from the total of 14 listed banks, and then made the demonstration analysis about the relationship between the performance and equity structure, according to their different equity structures. The result shows the first major shareholder is significantly relative to banks' performance, by contract, there's no notable correlation between the big five /ten and bank's performance, but square sum of equity by the ten big is correlate to the performance.

Keywords: *equity structure, performance, listed banks*

1. Introduction

As putting up shares in domestic stock market in 1997, Shenzhen Development Bank became the first Chinese commercial bank raising money in public, and then followed by another 14 banks, such as SPD bank, Minsheng bank, Huaxia bank and so on. Since 2005, the Chinese commercial banks have sped up to issue, take three state-owned Commercial banks as the symbol, the China ICBC, the CCB and the Bank of China were all offered to the public through the issuance of shares, and this fact shows that the scale and quantity of issue has improved, and equity structure also has a big change.

Equity structure is the proportion which the shareholders have partial ownership of the corporation. Its has two meanings: the first is centralized degree which has many standard, including the sum and the square sum of equity proportion which the big five or ten shareholders has; the second is capital structure, which means the proportion of different kind of capital source, including state-owned shares, corporate shares and public shares.

As different in time and various in background of being listed bank, the data information about these 14 banks is numerous and complex. So we take 6 banks as representatives, including the Bank of China, the China ICBC, the China Merchants Bank, the Bank of Communication, the Shanghai Pudong Development Bank, and Minsheng Bank. We want to examine the results of relationship between the equity structure and performance by utilizing these six banks and offer a reference to other listed banks and listing banks.

2. Literature Review

Berle & Means firstly put forward the hypothesis in their book titled *Modern Corporation and Private Property*, which shows the more decentralize, the worse performance is, Jensen & Meckling considered that corporate value lie on the proportion of shares held by internal shareholder, the bigger proportion is, the higher company value is in theoretically.[1, 3]

In past 20 years, more and more scholars have studied the relationship between equity structure and performance, but they didn't come to the same conclusion. Some scholars said share centralized degree was related to performance. Grossman & Hart's model [9] indicates that the single shareholder would lack stimulant to monitor company's management, to participate in the company's operation, or to drive company's value to grow if equity structure is too scattered, because the cost in supervision is more than return. The model developed by Shleifer & Vishny shows the wealth brought by stock price rising causes benefit of holding shareholder and small shareholders being consistent, so some share concentration is necessary. [5]

Other scholars took the opposite view about the relationship between performance and share concentration. Pound [5] proposed conflict of interest hypothesis and strategic alignment hypothesis, which deduced the inverse correlation between share and performance. The study by Qian Sun and Wilson H.S.Tong [6] indicate state-owned shares and corporate shares make a negative effect on company's performance.

Also some scholars believed there has a curve relation between the equity structure and performance. Morck, Shleife & Vishny studied [71] large companies list in the "Fortune", and found that it was inverted U-shaped curve

relationship between the proportion of shareholding by company's manager and company's value. McConnell & Servaes analyzed 1173 companies as sample in 1976 and 1093 company as sample in 1986, they discovered that the Tobin's Q is relate to shares holding by interior shareholders in curvilinear.[4,5]

Other scholars thought the performance was dependent to shares concentration. Demsets & Lehn's research [7] found there was no significant correlation between shares concentration degrees and accountant margin. Holdmess & Sheehan [11] concluded that there is no difference between the performance in the year existing absolutely shareholders and the year they are not existing in the same company So they considered there was no correlation between equity structure and company's performance.

Xu Xiaonian [12] is the first domestic scholar who used the empirical method to analyze the relationship between the company's performance and the equity structure of Chinese listed companies, he found that performance was relate to share concentration degree; the proportion of corporate shares has a remarkable positive influence to the company's performance, while the proportion of state-owned shares has a negative influence to performance, the circulation stock proportion is not remarkable to the company performance affects, but proportion of circulation stock had no significant impact on performance.

Similarly, other scholars have made empirical analysis from the perspective of kind of share. Zhou Yean [14] selected 160 samples from 745 listed companies at random at the end of 1997, inspected the relation between equity structure and net assets, and drew the conclusion: the ratio of returns on net assets were significally related to proportion of A stock, the state-owned shares and the corporate shares, but B stock, H stock had the inverse correlation relations with the ROA ratio. By observing the relation between ROA and A stock, Yu Zhidong [15] concluded that the company performance was low ralative to state-owned shares and the corporate shares, by contract, the performance was inverse correlation to A stock, the staff share and circulation stock in foreign capital, and all kinds of shareholders didn't play the due role in the company development.

Some scholars did not agree with this point of view. The study by Xu Xiaonian, Wang Yan [16] showed that the company's profit margins declined with increasing of proportion of state-owned shares, but increased with increasing of the proportion of corporate shares, and had no impact on public shares. Chen Xiaoyue and Xu Xiaodong [13] took empirical study in relationship between equity structure and performance among the listed companies besides financial industries in Shenzhen Stock Exchange in1996-1999 years, they deduced that in non-protection industries (including10 industries except for petroleum chemical industry, energy and raw material industry), the performance was related to scale of share holding by first major shareholder, was inverse correlate to scale of circulation stock, and no significant relation to state-owned shares or corporate shares.

Wushu Kun's [18] Empirical analysis shows that the relationship was significantly inverted U-shaped curve between performance and shares concentration degrees or internal holdings (including the staffs holding and executives manager holding), while the relationship was significantly U-shaped curve between performance and scales of circulation stock, state-owned shares and domestic corporate shares. Yang Deyong, Cao Yongxia [23] studied five listed banks on Chinese market, and found bank's performance was inversing to the biggest shareholder, and was relate to holding scale of the top five or the top 10 shareholders, and also inversing to circulation stock scales.

However, not all research shows that the equity structure and the achievements are correlative. Zhang Zongyi, Song Zengji [17] pointed out that equity structure and scales of manager's holding did not affect on the company's performance.

Getting a review of studies in equity structure and company administration in recent years, you would find following characteristic: firstly, the theory almost focus on benefit balance in shareholders and manage operators. Secondly, some scholars neglected the fact that different way of administration would make different influence on performance; also they didn't make quantitative analysis to situation and characteristic of Chinese listed companies. Thirdly, more empirical analysis, but few canonical parse analysis. Lastly, some studies neglected Chinese situation, such as "forge", effectiveness of Chinese stock market, authenticity of the stock price, and so on, and precisely was this problem leading to conflict among different studies.

3. Methods and Tools

3.1. Samples

The samples was select from annual report of six banks in the five year when is from 2003 to 2007, including the Bank of China (BOC for short), the China ICBC (i.e. ICBC), the China Merchants Bank (i.e. CMB), the Bank of Communication (i.e. BC), the Shanghai Pudong Development Bank (i.e. SPD) and the Minsheng Bank (i.e. MB), the objects are these 6 banks, the number of samples is 30.

3.2. Selection and Difinition of Variables

- Be interpreted variable S, the composite index which measure the performance of listed banks.
- Interpreted variables

F1: Proportion of share holding by the biggest shareholders

F5: Sum of proportion of share holding by top five shareholders

F10: Sum of proportion of share holding by top ten shareholders

HERF10: square sum of proportion of share holding by top ten shareholders, measuring the degree of shares concentration.

CSP: Proportion of circulation stocks

LSP: Proportion of corporate shares

SOSP: Proportion of state-owned shares

In this article, variable refer to the dependent variable, the independent variable and the controlled variable. The dependent variable is displayed with the performance of listed banks; the independent variable is display with CSP, LSP, SOSP and indexes such as F1, F5, F10 and HERF10. In addition, we selected ratio of debt to assets and asset size, etc., as control variables in order to control other characteristics which would make influence to performance.

4. Qualitative Analysis and Assumptions

4.1. About Degree of Share Concentration

Generally, there are 3 methods to measure equity concentration .The first is CR index, which is the sum of proportion of share holding by top N shareholders, and we choose F1, F5, and F10 to represent it. The second is herfindahl index, which is the square sum of proportion of share holding by top N shareholders, and we choose ten to instead of N. The third is Z index, which is the ratio of proportion of the share holding by the biggest shareholder and the second one. The bigger the ratio, the wider the gap between the top two shareholders, which means the biggest one has remarkable predominance. In our study, we use first two indexes to measure equity concentration degree.

The larger F1, the stronger monitoring of the company, which would lead to continuing drive function and the corresponding restriction mechanism to the agent. Therefore, the bigger of first major stockholder's superiority is, the more helpful this will be in decision-making and coordination in bank it will. However, there are differences among the first major shareholders, such as their characteristics, ability of acting, which would then have a significant impact on performance of commercial banks. Only when holding shareholder is good at making correct decisions, bank performance would be improved; otherwise, it will reduce the performance. So we can make hypothesis as follows:

H1. Bank's performance is not related to proportion of share holding by first major shareholder

If shares are concentrated in some independent shareholders, not only in one shareholder, it will keep the benefits and rights in balance, which will improve on the shortcoming taken by the fact that there is only one control stockholder. However, when the proportion of these shareholders is not big enough, their own benefits may not be coherence to the banks. The bank performances should be U-shape curve relation to the equity concentration degree, and the equity concentration degree should have a best sector theoretically. So we can make hypothesis as follows:

H2. Performance of bank is related to the sum of proportion of the top five shareholders' holding, and so as to the top ten, but is U-shape curve relation to the square sum of proportion of top ten shareholders' holding.

4.2. About Property of Shares

In the perspective of property of shares, equity can be divided into state-owned shares, corporate shares and circulation shares. We select state-owned shares and circulation shares to measure property of shares in this article. As a major shareholder, the state is in a special status differ from other major shareholders. This is mainly because performance largely depends on the behaviors of governing by agents of the state. When proportion of the state-owned shares is low, there must be lack of effective monitoring to the bank as lacking impetus, however, if that proportion rose to a certain degree, which could make government's attention to improving in monitoring to banks, it will help to improve the performance of banks. Thus we make hypothesis as follows:

H3. The proportion of state-owned shares is related to performance in curve relation. The performance is inverse correlate to this proportion when the proportion is low, but performance is related to this proportion when the proportion is high.

Investors of circulation stocks and corporate stocks all concern about bank's performances, because they take maximum profit as the only goal. Compared to corporate stocks, investors of circulation stocks are in inferior position, but they can vote with "feet" if cannot with hands, that means they can influence the banks' performance by buying or selling stocks of banks in the stock market. So we make hypothesis as follows:

H4. Sum of proportion of the corporate shares is related to the banks' performance; and the sum of proportion of circulation stock is also related with the banks' performance.

5. Empirical Study on Relationship between Performance and Equity Sructure

5.1. The Data We Collect about the Chinese Listed Banks

We get the data from the Chinese statistical yearbooks and the Chinese finance yearbooks between 2003 and 2007, and some missing data from the listed banks' an-

nual reports .(see the table 1).

5.2. The Model of Equity Structure Attributing to Banks' Performance

Theoretical analysis shows that the equity structure will ultimately affect the bank's profitability and performance in the market, but the relationship between them also needs empirical analysis. In order to get the function relation between the equity attributes and the banks' per-

Table1.The banks' performance index and the variables. S—the composite index which measure the performance of listed banks, we get it by the weighted average between PE and the net assets returns ratio; F1—Proportion of share holding by the biggest shareholders; F5—Sum of proportion of share holding by top five shareholders; F10—Sum of proportion of share holding by top ten shareholders; HERF10—square sum of proportion of share holding by top ten shareholders, measuring the degree of shares concentration; CSP—Proportion of circulation stocks; LSP—Proportion of corporate shares; SOSP—Proportion of state-owned shares

S	F1	F5	F10	HERF10	CSP(%)	LSP(%)	SOSP(%)
19.54	35.30	92.80	95.66	2709.32	100.00	0.00	0.00
12.72	35.30	92.60	95.50	2704.13	100.00	0.00	0.00
14.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
63.21	17.88	43.23	51.89	547.46	100.00	0.00	0.00
33.94	17.79	43.09	51.75	544.01	100.00	0.00	0.00
26.97	17.78	37.61	47.06	454.86	27.00	63.90	3.70
30.53	17.95	37.98	47.14	462.22	26.30	69.90	3.80
25.61	17.95	37.98	47.03	458.68	26.30	72.00	1.80
28.99	22.02	69.11	72.68	1289.49	100.00	0.00	0.00
21.15	21.78	62.49	65.62	1067.13	50.40	27.90	21.80
16.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17.62	67.49	95.32	97.26	4800.41	100.00	0.00	0.00
13.90	67.49	95.09	97.07	4794.88	100.00	0.00	0.00
12.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00
72.72	23.57	41.01	47.40	652.20	100.00	0.00	0.00
45.29	23.57	41.37	47.11	653.43	100.00	0.00	0.00
39.76	7.01	26.08	37.32	170.28	27.60	63.30	9.10
31.80	7.01	26.08	37.06	168.50	27.70	62.50	9.80
26.52	6.44	24.80	33.27	140.74	27.70	65.20	7.10
28.19	5.90	25.57	44.35	202.45	100.00	0.00	0.00
25.43	5.98	25.97	42.00	190.26	100.00	0.00	0.00
27.24	5.99	25.98	42.06	190.59	100.00	0.00	0.00
27.40	6.98	29.46	48.40	251.61	30.10	69.90	0.00
26.21	7.40	32.61	53.27	303.54	26.00	74.00	0.00

Table 2. the regression analysis between the banks' performance and the equity attribution

	Coefficients	Std. Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	12.469	4.37929	2.8473	0.0085	3.467373	21.4709118
CSP(%)	0.1979	0.05791	3.4164	0.0021	0.0788103	0.31689295
LSP(%)	0.1635	0.09333	1.7514	0.0917	−0.028383	0.35528582
SOSP(%)	−0.011	0.58503	−0.02	0.9845	−1.214039	1.19104615

Table 3. The regression analysis between the banks' performance and the equity attribution. A Dependent Variable: S

	model	Unstandardized Coefficients		Standardized Coeffi-	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	12.469	4.379		2.847	0.008
	CSP	0.198	0.058	0.577	3.416	0.637
	LSP	0.163	0.093	0.327	1.751	0.984
	SOSP	−0.011	0.585	−0.004	−0.020	0.276

Table 4. The relation between the banks' performance and the proportion of CSP. Model Summary and Parameter Estimates. Dependent Variable: S; the induced variable is the banks' performance index S,the independent variable is the proportion of CSP. The independent variable is proportion of circulation stocks

Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig	Constant	b1
Linear	.229	8.335	1	28	.007	17.188	.164

Table 5. The relation between the banks' performance and the concentration ratio of t he equity. A Dependent Variable: S

	model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	12.772	3.299		3.871	0.001
	F1	2.020	0.457	2.376	4.421	0.000
	F5	−.182	0.435	−0.370	−0.418	0.680
	F10	0.294	0.341	0.621	0.862	0.397
	HERF10	−0.030	0.006	−2.563	−5.31	0.000

formance, we have done the regression analysis by using the tools Excel and SPSS based on the related data of the listed banks. The results which get from Excel are showed in table 2.

The results which get from SPSS are showed in Table 3.

From the analysis of P-value and the coefficient of significant level (sig), we get the conclusions: The P-value and the sig about CSP is 0.02 (smaller than 0.05), which means there is a linear relation between the proportion of circulation stocks and the banks' performance; The related coefficients about LSP and SOSP are bigger than 0.05, which means there is no linear relation between the proportion of corporate shares (the proportion of state-owned shares) and the banks' performance.

In the following we have done specific research about the linear relation between the proportion of circulation stocks and the banks' performance, assume $S= b_0 + b_1 * CSP+e$, we did the linear regression by SPSS, the result is showed in Table 4.

By the result we can get S=17.188+0.164CSP+e, that means there is a positive linear relation between the banks' performance and the proportion of circulation stocks.

5.3. The Model of the Equity Concentrative Degree's Effect to Banks' Performance

To test the relationship between the equity concentrative degree and banks' performance we take banks' performance as the result of the equity concentrative degree, financial leverage and the company's assets complicated action, as the corporation's data shows, the explanatory ability and the fitting degree which independent variable to dependent variable is the strongest, the relationship between them is shown below:

$$S = a_0 + a_1F_1 + a_2F_5 + a_3F_{10} + a_4HERF10 + e \quad (1)$$

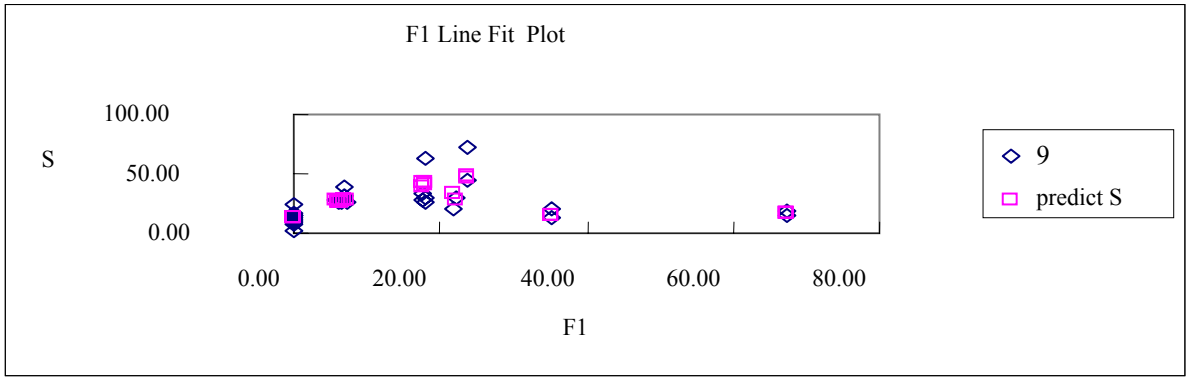


Figure 1. Fitting chart of real S and predict S used by F1

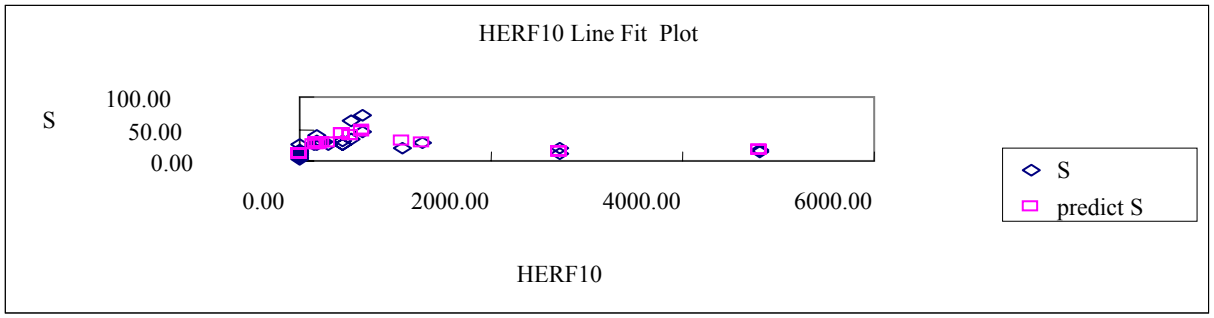


Figure 2. Fitting chart of real S and predict S used by HERF10

In which the variable S is the banks' performance, measured by earnings per share and net assets yield, a0 is a virtual variable delegates the intercept, F1 is the stock proportions that the first major shareholder holds, F5 is the stock proportions that the top five shareholders hold, F10 is the stock proportions that the top ten shareholders hold, HERF10 is the square sum of the top ten shareholders' proportions. Regression results are as follows:

The results: coefficient of significant level (sig) of proportion of shareholding by the biggest shareholders (F1) is 0.000, far less than 0.05, and the same to the square sum of proportion of share holding by top ten shareholders (HERF10) , which explains proportion of shareholding by the biggest shareholders (F1) and the square sum of proportion of share holding by top ten shareholders (HERF10) can interpret the variables S significantly. But coefficient of significant level (sig) of Sum of proportion of share holding by top five shareholders (F5) and proportion of share holding by top ten shareholders (F10) are far more than 0.05, which means F5 and F10 is not suitable for interpreting the variables S.

Actually, from the picture drew by SPSS, we can find the performance (S) predicted by F1 and HERF10 is highest fitting to real performance, which means a significant interpreting to the variables S. pictures drew by SPSS are as follows:

6. Conclusions and Suggestions

In summary, through the qualitative analysis and the empirical research of listed banks in China, we obtain some conclusions below:

First, choose the net asset yield and per share yield, we have constructed an Index S to reflect the particularity of banks' performance. However, because it is difficult to acquire relevant data, the construction of the index is not perfect; it is just an idea of reference, to make the indicators used to measure the banks' performance more reasonable.

Second, between the proportions of the first major shareholders in commercial bank and the performance of banks there is a significant correlation. As we have described in the qualitative analysis, the largest shareholder of governance will have a direct effect on performance of the bank, and to strengthen the shareholding ratio of the top five and the top 10 shareholders, forming a certain degree of control over checks and balances will help improve the listed Banks operating performance.

Third, the "multiple checks and balances" is a better equity structure. Decentralization and lower the largest shareholder's equity proportions can avoid the major shareholder controlling the bank effectively, and improving the structure of the bank's corporate management.

Fourth, changing the constitute of the circulation share that the shareholders hold, introducing a treatment capacity of institutional investors, one hand it is helpful of introducing and making use of foreign capital, the other hand it is helpful of introducing the advanced experience of foreign country. It is better for our country to optimize the combination of capital ,and to drive the fund flow to

the shareholders more reasonable, to improve the unsuitable structure of equity , further to improve corporate governance structure of China's listed companies and also it is helpful to enhance China's listed banks operating performance.

In conclusion because the banks in China is special, and in some state-owned commercial banks, the special performance is even more obvious, the interest rate is still controlled by the government, information disclosure is not entirely, but because the stock market is a bull market in 2007, Listed bank's annual report data generally higher, so the data's reliability remains to be improved. It can be said that the relationship between shareholding structure and listed banks' performance cannot be fully reflected from data of the listed banks' e annual report, but we can get some e enlightenments from this research, as a reference for the just listed and the forthcoming listed commercial banks in China.

REFERENCES

[1] Berle & Means, Modern, "Corporation and Private Property", *The Macmillan Company*, 1944, pp.69-70.

[2] Fama, E. &Jensen, M., "Seperation of ownership and control", *Journal of Law and Economy*, Vo1.26, 1983 pp.301-325.

[3] Jensen, M.C.and Meckling, W.H."Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, vol.3, 1976, pp. 305-360.

[4] McConnell and Servaes, "Additional Evidence on Equity Ownership and Corporate Value", *Journal of Financial Economics*, Vol.27, 1990, pp.596-612.

[5] John Pound; Richard J Zeckhauser 1988: "The market for corporate control :the economics of corporate takeovers and the New Zealand takeover code : an analysis and proposals for reform " [Auckland, N.Z.] : *New Zealand Centre for Independent Studies.*

[6] Qian Sun and Wilson H. S. Tong,"China share issue privatization: the extent of its success" , *Journal of Financial Economics*, Volume 70, Issue 2, Nov., 2003, pp.183-222.

[7] Demsetz and Lehn, "The Structure of Ownership : Causes and Consequences", *Journal of Political Economy* , Vol.93, 1985, pp.11155-11771.

[8] Morck, Shleifer and Vishny ,"Management Ownership and Market Valuation: An Empirical Analysis", *Journal of Financial Economics*,Vol.20, 1988, pp.293-315.

[9] Sanford J Grossman; Oliver D Hart; Hayne E Leland 1980: "Corporate financial structure and managerial incentives, Cambridge, Mass. : *National Bureau of Economic Research*,".

[10] Thomsen, Steen, "Ownership structure and economic performance in the largest European companies" *Strategic Management Journal*, vol.21, no.6, 2000, pp.689-705.

[11] Holdemess, C. and Sheehan, D., "The Role of Majority Shareholders in Publicly Held Corporations", *Journal of Financial Economics*, 1988, pp.20.

[12] Xu Xiaonian 1997: "Building a company governance mechanism and capital market on the base of corporation structure", *reform* ,vol 5, 1997.

[13] Chen xiaoyue, Xu xiaodong, "equity structure, corporate performance and protection of investors' interests" *Economic Research*, no.11, 2001.

[14] Zhou Ye an 1999: "The empirical research on influence of financial repression to enterprise financing ability " *Economic Research Journal*, issue 2, 1999.

[15] Yu Zhidong 2002: "*The corporate governance of listed company in a changing economic environment*" publishing house of Renmin University of China, 2002.

[16] Xu Xiaonian, Wang Yan, 2000: "*The Structure of Ownership and corporate governance on Chinese listed company*", publishing house of Renmin University of China.

[17] Zhang zongyi, Song zengji, "Empirical Study in equity structure and corporate performance of listed companies", *Quantity economical technical economy research*, no.1, 2003.

[18] Wu Shu-kun 2002: "Ownership structure and firm performance: an empirical research on Chinese public companies", *China Industrial Economy*, issue 1, 2002.

[19] Xu xiaodong, Chen xiaoyue, "analysis of the first major shareholders' impact on corporate governance and performance", *Economic Research*, no.2, 2003.

[20] Bai chongen, etc. *Empirical Study in governance structure of Chinese listed companies Economic Research*, no.2, 2005.

[21] Li hanjun, Zhang junxi, 2006, "The endogenetic degree of Listed corporate governance and performance." , *Management World*, no.5.

[22] Xu liping, etc. "Equity Concentration & Multi-shares Balance and impact in corporate performance", *Economic Research,* no.1, 2006.

[23] Yang deyong, Cao yongxia, "Empirical Study of equity structure and performance about Chinese listed banks", *Financial Research*, no5, 2007.

[24] Li hanjun, Zhang junxi, "The endogenetic degree of Listed corporate governance and performance." , *Management World*, no5, 2006.

J. Serv. Sci. & Management, **2008,1: 172-192**
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

The Theory of the Revenue Maximizing Firm

Beniamino Moro

Department of Economics, University of Cagliari, Viale Sant'Ignazio 17 - 09123 Cagliari (Italy)
Email: moro@unica.it

ABSTRACT

An endogenous growth model of the revenue maximizing firm is here presented. It is demonstrated that, in a static analysis, a revenue maximizing firm in equilibrium equates the average product of labor to the wage rate. In a dynamic analysis, the maximization rule becomes the balance between the rate of marginal substitution - between labor and capital - and the ratio of the wage rate over the discount rate. When the firm satisfies this rule, it grows endogenously at the rate of return on capital. The firm may also have multiple stationary equilibria, which are very similar to the static equilibrium.
JEL classification: D21, O41.

Keywords: *firm, theory of the firm, revenue maximization, endogenous growth*

1. Revenue Maximization Versus Profit Maximization and the Theory of the Firm

The original idea of a firm that maximizes revenue instead of profit was put forward by Baumol [2, 3], and further investigated during the sixties by Cyert-March [12], Galbraith [19], Winter [39] and Williamson [36]. Autonomously, a similar idea was also investigated by Rothbard [31], a precursor of the Austrian theory of the firm.[1] Nonetheless, the main stream economic thought, as Cyert-Hedrick [11] pointed out in their review article, remains characterized by an ideal market with firms for which profit maximization is the single determinant of behavior. [2] Indeed, the relevance of pure profit-maximization is not so obvious for modern corporations when ownership and control of the firm are separated and there are no dominant owners that merely maximize their profits [27].

More recently, the revenue-maximization dominance hypothesis has been re-proposed by Uekusa-Caves [34], Komiya [25, 26], Blinder [4, 5] and Tabeta-Wang [32, 33]. In all these papers it is argued that the separation of ownership and control in public companies causes a deviation of management from the pure profit maximization principle and provides a considerable degree of decision-making autonomy for managers. In fact, in an oligopolistic market, each firm may set up its own goal, and the choice to maximize revenue or profit depends on the real interests of the managers, and is also influenced by the corporate culture and institutional arrangements of the country where the firm operates. According to Kagono *et al*. [23], the principal objectives of Japanese firms are growth and market-share gaining, which imply that they are revenue maximizers, while US corporations emphasize more on short run investment returns and capital gains, which means that they are profit maximizers.[3] In

[1] See Anderson [1], and for the Austrian school see Foss [15, 18], and Witt [41].

[2] In fact, during the '30s and '40s, a great dispute was due to the "Old Marginalist Debate" which questioned the relevance of the profit-maximization assumption in neoclassical theory of the firm. In the '70s, the marginalist debate changed tone with the emergence of the theories of agency costs, property rights, and transactions costs theories of the firm. These gave rise to the "New Institutionalist" field of research, where the object of study changed from how to reconcile firm behavior with marginalist principles to how to reconcile firm structure with marginalist principles. Following the seminal work by Coase [8], papers belonging to the Institutionalist debate can be divided in the transactions costs economics [24, 37, 38], and the contractual field of research [6, 13, 21, 22]. The old marginalist debate re-emerged in the '80s with the evolutionary theory of the firm by Nelson-Winter [30], Winter [40], and Foss [14]. Finally and more recently, we also have a "Knowledge-based" theory of the firm [16, 17, 20], and a "Resource-based" theory [9, 10].

[3] Blinder [5] builds up a model to demonstrate that revenue-maximizers like Japanese firms have an advantage when competing with profit-maximizers. Particularly, he points out that the revenue-maximizer is likely to drive its profit-maximizers rivals out of business if either average costs are declining or learning is a positive function of cumulative output. Tabeta-Wang [33] find the following four reasons to explain why Japanese firms are in general able to act like revenue-maximizers. First of all, in Japan, expansion in firm-size is a necessary condition to maintain the life-time employment system and internal promotion. Secondly, a faster growth of the firm helps hiring new young employees, and keeping low the average age of the work force helps to maintain low labor costs. Thirdly, Japanese firms pursue the growth-oriented strategy also because there is little external pressure for short-term earnings, and tax rates on reinvested earnings are lower than tax rates on dividends.

fact, as Anderson [1] points out, the profit maximizing versus the revenue maximizing strategy of the firm still stays as an open question, the answer to which only time will tell.

A parallel problem to this dispute is how to formalize the firm behavior in the two cases. At this regard, the mainstream microeconomic analysis has been mainly oriented to the profit maximization strategy, while very little attention has been devoted to the revenue maximizing case. Apart from the static analysis during the sixties, the latter field of research is very poor. After the seminal work by Leland [28], Van Hilton-Kort-Van Loon [35] and Chiang [7] put the problem in the contest of the optimal control theory and demonstrated that a revenue maximizing firm subject to a minimum profitability constraint is in equilibrium at a smaller capital-labor ratio than a profit maximizing one. This result is also obtained here. Anyway, Leland's model suffers of some limitations - e.g. he considers constant the share of profits used to self-financing the accumulation of capital - which preclude him to develop a complete dynamic model which fully describes the dynamics of a revenue maximizing firm. The aim of this paper is to fill the gap in this field of research, presenting a complete endogenous growth model of a revenue maximizing firm.

The paper is organized as follows. In section 2, the problem of a revenue maximizing firm versus the classical problem of profit maximization is analyzed from a static point of view. First of all, the analysis is made without taking into account a minimum acceptable return on capital constraint (section 2.1) and then with such a constraint (section 2.2). In section 2.3, the analysis is generalized into a rate of profit maximization problem and into a revenue per unit of capital maximization problem, respectively. We obtain the fundamental rule followed by a static revenue maximizing firm, according to which the firm equates the average product of labor to the wage rate. The same rule also applies in a dynamic context.

In section 3, we use the optimal control theory to describe the dynamics of a revenue maximizing firm. With respect to Leland's model, this paper differs on the following two assumptions: a) we suppose that the firm's accumulation of capital is limited to the non distributed profits and b) we also suppose that the share of the reinvested profits is endogenously determined by the firm, while in Leland's model this is a constant. We also demonstrate that only some of the possible dynamic equilibria (stationary equilibria) correspond to those discussed in the static analysis. Anyway, it is also demonstrated that an endogenous growth equilibrium of the firm does exist, where the rate of growth is obtained from the solution of a system of differential equations which fully describes the dynamics of the model. Also in this section, the problem is first analyzed without taking into account any minimum acceptable return on capital constraint (sections 3.1-3.4) and then with such a constraint (section 3.5). Our main conclusion is that, in a dynamic context, the equilibrium of a revenue maximizing firm requires not only that the marginal rate of substitution between labor and capital to be equal to the shadow value of the capital-labor ratio, but that this value also balances the ratio of the wage rate over the discount rate.

Further, if we introduce a minimum acceptable return on capital constraint, this must be added to the discount rate when determining the equilibrium equality with the rate of marginal substitution between labor and capital. As a consequence, a change of the minimum acceptable return on capital rate has the same effect as a variation of the discount rate. Finally, section 4 is devoted to the concluding remarks.

2. The Static Analysis of the Firm Behavior

2.1. The Equilibrium Conditions Without a Minimum Acceptable Return on Capital Constraint

We make the following neoclassical assumptions on the firm production function $Q=Q(K, L)$, where K is capital and L is labor:

a) Q is linear homogenous and strictly quasi-concave, which implies that $Q=Q(K, L)=Lf(k)$, where $f(k)=Q(K/L, 1)$ and $k=K/L$; $f(0)=0$ and $\lim_{k\to\infty} f(k)=\infty$;

b) the marginal productivity of capital $Q_K = f'(k)$ has $\lim_{k\to 0} f'(k)=\infty$ and $\lim_{k\to\infty} f'(k)=0$;

c) the marginal productivity of labor $Q_L = f(k)-kf'(k)$ has $\lim_{k\to 0}=0$ and $\lim_{k\to\infty}=\infty$.

We also assume that the price of the firm's output is normalized to one, so that both the nominal and the real wage rate can be indicated by w. First of all, we demonstrate that if the firm program is:

$$\text{Maximize } [Q(K,L)-wL] \tag{1}$$

subject to $Q(K, L)\geq 0$

then there are no limits to the expansion of capital, which means that no finite capital-labor ratio exists in equilibrium. To see this, let us form the Lagrangian function:

$$\Im = Q(K,L)-wL+\lambda Q \tag{2}$$

where λ is a Lagrangian multiplier. The Kuhn-Tucker conditions state that in equilibrium we have:

$$\frac{\partial \Im}{\partial K} = Q_K + \lambda Q_K = 0 \rightarrow (1+\lambda)Q_K = 0 \tag{3}$$

Lastly, there is some possibility that administrative guidance and controls lead Japanese firms to act like revenue-maximizers. At this regard, Nakamura [29] clamed that administrative guidance and controls play a role as a "shelter from the storm" once the firm grows beyond the limits of a market accepted profitability.

$$\frac{\partial \Im}{\partial L} = Q_L - w + \lambda Q_L = 0 \rightarrow (1+\lambda)Q_L = w \quad (4)$$

$$\frac{\partial \Im}{\partial \lambda} = Q \geq 0, \quad \lambda \geq 0, \quad \lambda Q = 0 \quad (5)$$

Clearly we see that if $Q>0$, which means that the firm produces something, from (5) we have $\lambda=0$, so as equation (3) reduces to $Q_K = 0$ and equation (4) to $Q_L = w$. Thus, while equation (4) states a limit to the decreasing of the marginal productivity of labor, which cannot fall under the level of the real wage rate, from equation (3) it follows that no limits to capital accumulation exist in this problem. Given that $Q_K = f'(k) \rightarrow 0$ for $k \rightarrow \infty$ and $Q_L = f(k) - kf'(k) \rightarrow \infty$ for $k \rightarrow \infty$, it follows that no finite k exists which maximizes the profit of the firm.

Under the same conditions, no equilibrium exists for the revenue maximizing firm too. To see this, let the firm maximization program be:

$$\text{Maximize } Q(K, L) \quad (6)$$

subject to $Q(K,L) - wL \geq 0$

where $Q - wL \geq 0$ can be interpreted as the non bankruptcy constraint. The Lagrangian of this problem takes the form:

$$\Im = Q(K,L) + \lambda[Q(K,L) - wL] \quad (7)$$

from which we derive the following Kuhn-Tucker conditions:

$$\frac{\partial \Im}{\partial K} = Q_K + \lambda Q_K = 0 \rightarrow (1+\lambda)Q_K = 0 \quad (8)$$

$$\frac{\partial \Im}{\partial L} = Q_L + \lambda Q_L - \lambda w = 0 \rightarrow Q_L + \lambda(Q_L - w) = 0 \quad (9)$$

$$\frac{\partial \Im}{\partial \lambda} = Q - wL \geq 0, \ \lambda \geq 0, \ \lambda(Q - wL) = 0 \quad (10)$$

If the non bankruptcy constraint is not binding, that is if $Q - wL > 0$, which implies $f(k)>w$, then from (10) we deduce that $\lambda=0$. In this case, from equation (8) we have $Q_K = 0$ and from equation (9) we have $Q_L = 0$. These conditions are not consistent, because the former is satisfied for $k \rightarrow \infty$, while the latter for $k \rightarrow 0$. On the contrary, if the non bankruptcy constraint is binding, that is if $Q=wL$, which implies $f(k)=w$, then from (10) we deduce that $\lambda>0$. In this case, from equation (8) we again have $Q_K = 0$, while from equation (9) we obtain $\lambda = Q_L/(w - Q_L)$. But, once again, the capital-labor ratio $k = k_w$ for which the condition $f(k_w) = w$ is satisfied is not an equilibrium ratio, because we have $Q_K = f(k_w) > 0$, while equation (8) requires that $Q_K \rightarrow 0$.

This inconsistency depends on the fact that, if no rental market price exists for capital, the firm takes advantage of accumulating capital without limits. So, the only way to avoid that is to fix the level of capital K_0. If we do that, the problem becomes definite, both for the profit maximizing firm and for the revenue maximizing one.

For the profit maximizing firm, the Lagrangian function is maximized only with respect to the labor factor, while capital stays constant. In this case, from equation (4), when $Q>0$ and $\lambda=0$, we have:

$$Q_L = f(k) - kf'(k) = w \quad (11)$$

which is the well known rule of the profit maximizing firm that equates the marginal productivity of labor to the real wage rate. In the same way, for the revenue maximizing firm, if $f(k)>w$, so as from (10) we have $\lambda=0$, then from (9) we obtain $Q_L = 0$, which implies, according to assumption c), that $k = 0$. In this case, the firm does not produce anything. On the contrary, if the firm does produce something, it must be:

$$\frac{Q}{L} = f(k) = w \quad (12)$$

Therefore, we can conclude that the rule for a profit maximizing firm is given by equation (11), and the capital-labor ratio that satisfies it can be indicated with $\hat{k}$, so that we can write:

$$f(\hat{k}) - \hat{k} f'(\hat{k}) = w \quad (13)$$

whereas the rule for a revenue maximizing firm is given by equation (12) and the capital-labor ratio that satisfies

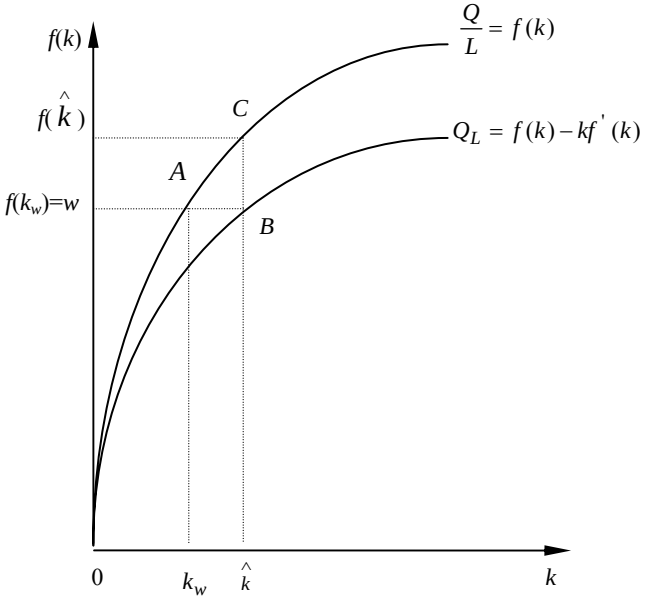


Figure 1. The average productivity of labor function *f(k)* and the marginal productivity of labor function *f(k)-kf'(k)* with respect to the capital-labor ratio *k*

it can be indicated with k_w, so that we have:

$$f(k_w) = w \tag{14}$$

Both these conditions are shown in figure 1, where the functions $f(k)$ and $Q_L = f(k) - kf'(k)$ are depicted. Since $f'(k)>0$, we deduce that Q_L always stays below $f(k)$. Once the real wage rate w and the level of capital K_0 are given, the profit maximizing firm is in equilibrium at the point B, where $f(\hat{k}) - \hat{k} f'(\hat{k}) = w$. In this point, the output per worker is $f(\hat{k})$, so the profit per worker is given by the difference:

$$f(\hat{k}) - w = \hat{k} f'(\hat{k}) \tag{15}$$

If we indicate with r the rate of profit or the rate of net return on capital, so that:

$$r = \frac{Q - wL}{K} = \frac{f(k) - w}{k} \tag{16}$$

then, from equation (15) we deduce that the maximum rate of profit $\hat{r}$ is given by:

$$\hat{r} = \frac{f(\hat{k}) - w}{\hat{k}} = f'(\hat{k}) \tag{17}$$

which says that $\hat{r}$ is equal to the marginal productivity of capital corresponding to the optimal capital-labor ratio $\hat{k}$. A revenue maximizing firm is in equilibrium at point A, where the average productivity of labor equals the real wage rate, that is $f(k_w) = w$. At this point the profit rate is zero, as we have:

$$r_w = \frac{f(k_w) - w}{k_w} = 0 \tag{18}$$

Each level of w corresponds to a minimum capital-labor ratio k_w for which we have a null rate of profit. Given the amount of capital, the firm employs more labor if it maximizes revenue than if it maximizes profits; and this explains why the equilibrium ratio k_w is smaller than the profit maximizing one $\hat{k}$. If w increases, both ratios k_w and $\hat{k}$ increase, and their difference increases too.

2.2. The Equilibrium Conditions with a Minimum Acceptable Return on Capital Constraint

If we introduce a minimum acceptable return on capital constraint of the form:

$$\frac{Q - wL}{K_0} = \frac{f(k) - w}{k} \geq r_0 \tag{19}$$

then a revenue maximizing firm solves the following maximization program:

$$\text{Maximize } Q(K_0, L) \tag{20}$$

subject to $Q(K_0,L) - wL \geq r_0 K_0$.

The Lagrangian of this problem is:

$$\Im = Q(K_0, L) + \lambda\left[Q(K_0, L) - wL - r_0 K_0\right] \tag{21}$$

from which we derive the following Kuhn-Tucker conditions:

$$\frac{\partial \Im}{\partial L} = Q_L + \lambda(Q_L - w) = 0 \tag{22}$$

$$\frac{\partial \Im}{\partial \lambda} = (Q - wL - r_0 K_0) \geq 0,\ \lambda \geq 0,\ \lambda(Q - wL - r_0 K_0) = 0 \tag{23}$$

If the constraint is not binding, that is if $Q - wL - r_0 K_0 > 0$, from (23) we have $\lambda = 0$ and from equation (22) we deduce that $Q_L = 0$, hence the firm does not produce anything. If on the contrary the firm does produce something, the constraint is binding, so that $Q - wL - r_0 K_0 = 0$ and $\lambda > 0$. In this case, from equation (22) we have:

$$\lambda = \frac{Q_L}{w - Q_L} \tag{24}$$

while, from the condition $Q - wL - r_0 K_0 = 0$, we have:

$$\frac{Q - wL}{K_0} = r_0 \tag{25}$$

which says that the rate of return on capital must be equal to the minimum acceptable rate. This condition can be put in the form:

$$f(k) - kr_0 = w \tag{26}$$

If we consider that the range of r_0 is:

$$0 \leq r_0 \leq f'(k) \tag{27}$$

it follows that the curve $f(k) - kr_0$ has an intermediate mapping between $f(k)$ and $f(k) - kf'(k)$, as it is depicted in figure 2(a). Given the level of capital K_0, the real wage rate w and the minimum acceptable return on capital r_0, the equilibrium point of a revenue maximizing firm is no more A, but A', where equation (26) is satisfied. Let $k_0 = K_0/L_0$ be the capital-labor ratio which satisfies equation (26), in point A' the output per worker of the firm is $f(k_0)$. Clearly, at this point the rate of return on capital is:

$$r_0 = \frac{f(k_0) - w}{k_0} \tag{28}$$

where $k_w \leq k_0 \leq \hat{k}$. When r_0 varies between zero and

$f'(k)$, the equilibrium point *A*' ranges between *A* and *B*, while the equilibrium capital-labor ratio ranges between k_w and $\hat{k}$. In figure 2(*b*), both the revenue-capital ratio Q/K_0 (which is equal to the output-capital ratio, because *P*=1) and the rate of profit *r* are depicted. The rate of profit varies with respect to *k* according to the rule:

$$r(k) = \frac{f(k) - w}{k} \tag{29}$$

For each level of *w* and K_0, the rate of profit varies as follows. In the interval $0 \leq k < k_w$, it is negative and increases with respect to *k*, ranging from $-\infty$ to zero. In the interval $k_w \leq k \leq \hat{k}$, it is positive and increasing, and

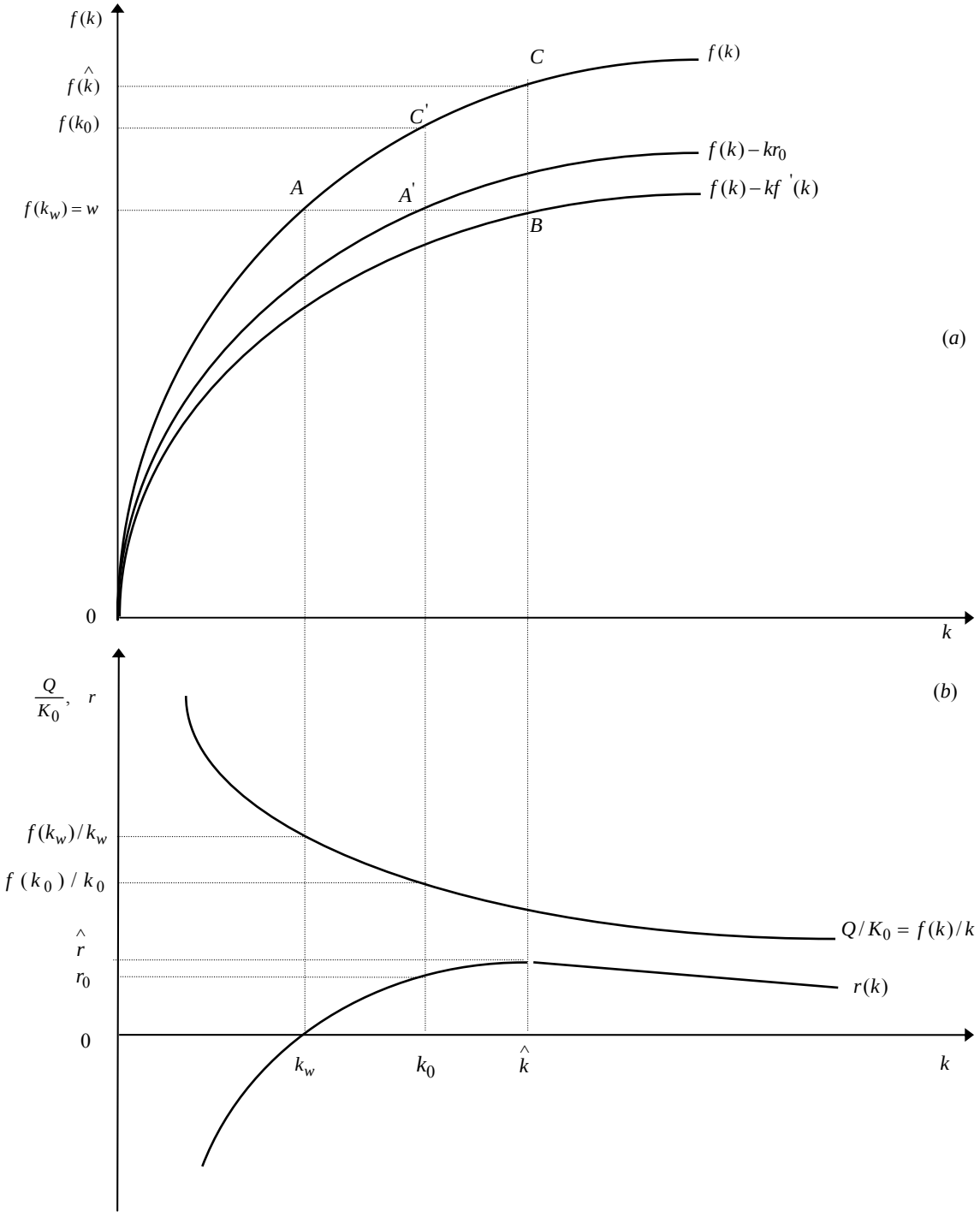


Figure 2. The mapping of the revenue-capital ratio (output-capital ratio) Q/K_0 and the rate of profit $r(k)$ with respect to k

varies between zero and its maximum value $\hat{r}$ given by:

$$\hat{r} = \frac{f(\hat{k}) - w}{\hat{k}} \tag{30}$$

Finally, in the interval $\hat{k} < k < \infty$, it is positive and decreasing, and tends asymptotically to zero for $k \to \infty$.

From figure 2(*b*), we see that for $k = k_w$ the rate of profit *r* is zero, for $k = \hat{k}$ it takes the maximum value $\hat{r}$, while for $k = k_0$ it equals r_0. In figure 2(*b*), the revenue-capital ratio or output-capital ratio $Q/K_0 = f(k)/k$ is also depicted. Given the level of capital K_0, this increases with respect to *L*, so it decreases with respect to the capital-labor ratio *k*. Using de L'Hôpital's rule we have:

$$\lim_{k \to 0} \frac{f(k)}{k} = \lim_{k \to 0} \frac{\frac{d}{dk} f(k)}{\frac{d}{dk} k} = \lim_{k \to 0} f'(k) = \infty \tag{31}$$

$$\lim_{k \to \infty} \frac{f(k)}{k} = \lim_{k \to \infty} f'(k) = 0 \tag{32}$$

so the ratio $f(k)/k$ ranges between infinity and zero for $0 < k < \infty$. Furthermore, we can check if the sign of its derivative is negative:

$$\frac{d}{dk} \frac{f(k)}{k} = -[f(k) - kf'(k)] / k^2 < 0 \tag{33}$$

But for $k < k_w$ the ratio $Q/K_0 = f(k)/k$ does not make sense, because it does not respect the non bankruptcy constraint. On the contrary, for $k \geq k_w$, this ratio is economically meaningful and decreases with *k*, tending asymptotically to zero as $k \to \infty$. Hence, its maximal economically meaningful value $Q_w / K_0 = f(k_w)/k_w$ corresponds to a capital-labor ratio equal to k_w.

Therefore, once the level of capital K_0 is given, if the objective function of the firm is to maximize profits, this corresponds to maximize the rate of profit. In this case, the optimal quantity of labor to be employed is $\hat{L}$ for which $\hat{k} = K_0 / \hat{L}$ and the equilibrium point is *B* in figure 2(*a*). On the contrary, if the objective function of the firm is to maximize revenue, this corresponds to maximize the revenue or output per unit of capital. In this case, the optimal quantity of labor to be employed is $L_w > \hat{L}$ for which $k_w = K_0 / L_w$ and $k_w < \hat{k}$. Then, the equilibrium point is *A* in figure 2(*a*), which corresponds to the maximum level of employment compatible with the respect of the non bankruptcy constraint.

If the revenue maximizing firm must respect a minimum acceptable return on capital constraint of the form $r \geq r_0$, then the equilibrium point is *A*' in figure 2(*a*), where $f(k) - kr_0 = w$. Given the level of capital K_0, the employment in point *A*' is equal to L_0 for which $\hat{L} \leq L_0 \leq L_w$ and $k_0 = K_0/L_0$. Therefore, in this point equation (28) is verified.

Thus, if r_0 increases from zero to $\hat{r}$, the equilibrium point *A*' moves along the segment *AB*, going away from point *A* to point *B*. Point *A'* is as far from point *A* as higher the minimum acceptable return on capital r_0 is. An optimal level of the capital-labor ratio k_0 for which $k_w \leq k_0 \leq \hat{k}$ corresponds to each predetermined level of r_0. Given the level of capital K_0, this also corresponds to an employment level L_0 for which $\hat{L} \leq L_0 \leq L_w$, where L_w is the maximum level of employment compatible with the respect to the non bankruptcy constraint and $\hat{L}$ is the level of employment that maximizes the rate of profit. Introducing a minimum acceptable rate of return on capital amounts to introduce a limit to the expansion of the production and the employment of the firm.

2.3. A Generalization of the Static Theory of a Revenue Maximizing Firm

Once we have proven that, given the absolute value of capital K_0, the profit maximization of a firm amounts to the maximization of the rate of profit (or the rate of return on capital), while the maximization of the revenue amounts to the maximization of the revenue (or output) per unit of capital, these results can be generalized for each given level of capital K_0. So, the equilibrium conditions, both for a profit or for a revenue maximizing firm, become independent from the absolute value of capital and labor employed in the production process. They only depend on the capital-labor ratio. Therefore, if we define:

$$r = \frac{Q - wL}{K} = \frac{f(k) - w}{k} \tag{34}$$

as the rate of profit (or the rate of *net* return on capital), then the program of a profit maximizing firm becomes:

$$\text{Maximize} \frac{Q - wL}{K} \tag{35}$$

subject to $\frac{Q}{K} \geq 0$

The Lagrangian of this problem takes the form:

$$\Im = \frac{Q - wL}{K} + \lambda \frac{Q}{K} \tag{36}$$

from which the following Khun-Tucker conditions can be derived:

$$\frac{\partial \Im}{\partial K} = \frac{KQ_K - (Q - wL)}{K^2} + \frac{\lambda KQ_K - \lambda Q}{K^2} = 0 \quad (37)$$

$$\frac{\partial \Im}{\partial L} = \frac{Q_L - w}{K} + \frac{\lambda Q_L}{K} = 0 \quad (38)$$

$$\frac{\partial \Im}{\partial \lambda} = \frac{Q}{K} \geq 0, \ \lambda \geq 0, \ \lambda \frac{Q}{K} = 0 \quad (39)$$

Equations (37) and (38) can also be stated in the form:

$$(1+\lambda)\left[Q - KQ_K\right] = wL \quad (40)$$

$$(1+\lambda)\, Q_L = w \quad (41)$$

If the firm produces, then $Q/K>0$ and from equation (39) we deduce that $\lambda=0$, thus equations (40) and (41) take the form:

$$Q - KQ_K = wL \rightarrow f(k) - kf'(k) = w \quad (42)$$

$$Q_L = f(k) - kf'(k) = w \quad (43)$$

These two equations state the same maximization rule, corresponding to the equality of the marginal productivity of labor to the real wage rate. This again re-asserts that the equilibrium capital-labor ratio for a profit maximizing firm is $\hat{k}$.

Likewise, the program of a revenue maximizing firm becomes:

$$\text{Maximize } \frac{Q}{K} \quad (44)$$

subject to $\frac{Q - wL}{K} \geq 0$

where the firm respects a non bankruptcy constraint. To solve this problem, let us form the Lagrangian function:

$$\Im = \frac{Q}{K} + \frac{\lambda(Q - wL)}{K} \quad (45)$$

from which the following Kuhn-Tucker conditions can be derived:

$$\frac{\partial \Im}{\partial K} = \frac{KQ_K - Q}{K^2} + \frac{\lambda\left[KQ_K - (Q - wL)\right]}{K^2} = 0 \quad (46)$$

$$\frac{\partial \Im}{\partial L} = \frac{Q_L}{K} + \frac{\lambda(Q_L - w)}{K} = 0 \quad (47)$$

$$\frac{\partial \Im}{\partial \lambda} = \frac{Q - wL}{K} \geq 0, \ \lambda \geq 0, \ \frac{\lambda(Q - wL)}{K} = 0 \quad (48)$$

Equations (46) and (47) can be put in the form:

$$(1+\lambda)\,(Q - KQ_K) = \lambda wL \quad (49)$$

$$(1+\lambda)\, Q_L = \lambda w \quad (50)$$

respectively. If the non bankruptcy constraint is not binding, that is if $Q-wL>0$, then from (48) we deduce $\lambda=0$. In this case, from equation (50) we have $Q_L=0$, which implies:

$$Q_L = f(k) - kf'(k) = 0 \quad (51)$$

or:

$$f(k) = kf'(k) \quad (52)$$

This last equation is satisfied only for $k=0$, that is, when the firm does not produce anything. Analogously, for $\lambda=0$, also equation (49) leads to the same conclusion. In fact, if $\lambda=0$, then $Q=KQ_K$, and dividing by L we obtain (52). It follows that, if the firm produces, the non bankruptcy constraint must be binding, which implies:

$$Q - wL = 0 \rightarrow f(k) = w \quad (53)$$

In this case, from (48) we deduce that $\lambda>0$ and the value of λ can be determined either from (49) or (50). Using the latter, we have:

$$(1+\lambda)[\, f(k) - kf'(k)\,] = \lambda f(k) \quad (54)$$

The same result is obtained from (49) dividing by L. From (54) we have:

$$\lambda = \frac{f(k) - kf'(k)}{kf'(k)} = \frac{Q_L}{kQ_K} \quad (55)$$

which defines the equilibrium value of λ. Equation (55) can also be put in the form:

$$MRS_{L/K} = \frac{Q_L}{Q_K} = \lambda k \quad (56)$$

where $MRS_{L/K}$ stays for the marginal rate of substitution between labor and capital. If we interpret λ as the shadow price of capital, then (56) means that in equilibrium the marginal rate of substitution must equal the shadow value of per capita capital. As we shall see later, this rule also applies in a dynamic analysis. Because equations (53) and (56) are satisfied for $k=k_w$, it is confirmed that the equilibrium capital-labor ratio for a revenue maximizing firm is k_w, that is the value of k which guaranties the equality between the average productivity of labor and the real wage rate.

Finally, if the firm must satisfy a minimum return on capital constraint, the maximization program is as follows:

$$\text{Maximize } \frac{Q}{K} \quad (57)$$

subject to $\frac{Q - wL}{K} \geq r_0$

where r_0 is the minimum acceptable rate of return on capital. The Lagrangian of this problem is:

$$\Im = \frac{Q}{K} + \lambda\left[\frac{Q - wL}{K} - r_0\right] \quad (58)$$

from which we derive the following Kuhn-Tucker conditions:

$$\frac{\partial \Im}{\partial K} = \frac{KQ_K - Q}{K^2} + \lambda\left[\frac{KQ_K - (Q - wL)}{K^2}\right] = 0 \quad (59)$$

$$\frac{\partial \Im}{\partial L} = \frac{Q_L}{K} + \frac{\lambda(Q_L - w)}{K} = 0 \quad (60)$$

$$\frac{\partial \Im}{\partial \lambda} = \frac{Q - wL}{K} - r_0 \geq 0,\ \lambda \geq 0,\ \lambda\left[\frac{Q - wL}{K} - r_0\right] = 0 \quad (61)$$

Given that equations (59) and (60) are similar, respectively, to equations (46) and (47), they can again be put in the form of equations (49) and (50). From (61) we can conclude that, if the minimum rate of return on capital constraint is not binding, then λ=0. In this case, equations (51) and (52) follow, and k=0. On the contrary, if the firm does produce, the minimum rate of return on capital constraint must be binding. Therefore, it must be:

$$r_0 = \frac{Q - wL}{K} = \frac{f(k) - w}{k} \quad (62)$$

or

$$f(k) - kr_0 = w \quad (63)$$

and λ is positive. In this case, either from (49) or (50), we conclude that

$$(1+\lambda)[\ f(k) - kf'(k)\ ] = \lambda\ [\ f(k) - kr_0] \quad (64)$$

and

$$\lambda = \frac{f(k) - kf'(k)}{k[\ f'(k) - r_0\ ]} = \frac{Q_L}{k(Q_K - r_0)} \quad (65)$$

which is positive for $k < \hat{k}$, as the difference $f'(k) - r_0 = Q_K - r_0$ is also positive.

Equation (65) can be put in the form:

$$NMRS_{L/K}(r_0) = \frac{Q_L}{Q_K - r_0} = \lambda k \quad (66)$$

which states the equilibrium condition for a revenue maximizing firm in the case that a minimum rate of return on capital constraint must be satisfied. In this case, the equilibrium rule becomes the equality of the net marginal rate of substitution between labor and capital ($NMRS_{L/K}$), which is a function of r_0, to the shadow value of per capita capital λk. Now, the $NMRS_{L/K}$ is defined as the ratio of the marginal productivity of labor over the marginal productivity of capital net of the minimum acceptable rate of return r_0.

Since equations (63) and (66), as is shown in figure 2, are satisfied for $k = k_0$, it is confirmed that the equilibrium capital-labor ratio for a revenue maximizing firm, when a minimum rate of return on capital constraint must be satisfied, is $k = k_0$. To this value of k we have the balance between the average productivity of labor, net of the minimum acceptable rate of return on capital, and the real wage rate. This condition is also obtained in the following dynamic analysis.

3. The Dynamic Analysis of the Revenue Maximizing Firm

3.1. The Equilibrium Conditions without a Minimum Acceptable Return on Capital Constraint

In the dynamic analysis, K, L and Q are all functions of time t. Therefore, the instant dynamic production function is

$$Q = Q\left[K(t),\ L(t)\right] \quad (67)$$

where Q has the same properties as in *a*)-*c*) of section 2.1. Let P be the price of the output, so the instant revenue is PQ. The objective function of a revenue maximizing firm then is the *present value* (PV) of all future revenues, that is

$$PV = \int_0^\infty PQ\left[K(t),\ L(t)\right] e^{-\rho t}\, dt \quad (68)$$

where ρ is the instantaneous discount rate.

The dynamics of capital accumulation can be defined as follows:

$$\dot{K} = \alpha\{PQ[K(t),\ L(t)] - WL\} \quad (69)$$

where $\dot{K} = dK/dt$ is the net instantaneous investment (there is no depreciation of capital), while W is the nominal wage rate. Finally, $0 \leq \alpha \leq 1$ is the share of instantaneous profits the firm decides to accumulate.

Normalizing P=1 so as $w = W/P = W$ indicates both the real and the nominal wage rate, equation (69) must satisfy the condition:

$$\dot{K} = \alpha\ (Q - wL) \geq 0 \quad (70)$$

which is both the law of motion of capital and the instantaneous non bankruptcy constraint.

Starting from an initial level of capital K_0, the program of the revenue maximizing firm is

$$\text{Maximize} \int_0^\infty Q(K,L) e^{-\rho t}\, dt \quad (71)$$

subject to $\dot{K} = \alpha\ (Q\text{-}wL) \geq 0$, $0 \leq \alpha \leq 1$, and $K(0) = K_0$, K_∞ free, K_0, w given.

In this problem, K is the state variable while L is a control variable. What about α, which is the share of instantaneous profits the firm decides to accumulate?[4] In fact, we can suppose that the firm can decide in each instant of time to accelerate the speed of capital accumulation by increasing the value of α. We have the maximum speed if α=1, while there is no capital accumulation if α=0. In the latter case, all profits are distributed to the stockholders. Hence, it is obvious to consider α as a second control variable of the problem.

To solve the program (71), let us indicate with H_C the current value Hamiltonian:

$$H_C = Q(K, L) + \lambda\alpha(Q - wL) \tag{72}$$

where λ is a dynamic Lagrangian multiplier. Furthermore, as the costate variable is strictly adherent to the state variable K, λ can also be interpreted as the shadow price of capital.

Because the Hamiltonian is linear with respect to α, the maximization rule $\partial H_C / \partial \alpha = 0$ does not apply in this case. Instead, we have a maximum in one of the extreme values of the α interval [0, 1]. Moreover, since λ is positive, as we will see later, H_C is maximized for α=1 if Q–wL>0, while, for Q–wL=0, α is indeterminate. Furthermore, since Q is a strictly quasi-concave function, the Hamiltonian H_C is a concave function of K and L. So, the following conditions of the maximum principle are necessary and sufficient for the solution of the program (71):

$$\frac{\partial H_C}{\partial L} = Q_L + \lambda\alpha(Q_L - w) = 0 \tag{73}$$

$$\dot{K} = \frac{\partial H_C}{\partial \lambda} = \alpha(Q - wL) \tag{74}$$

$$\dot{\lambda} = -\frac{\partial H_C}{\partial K} + \rho\lambda = -Q_K - \lambda\alpha Q_K + \rho\lambda \tag{75}$$

$$\lim_{t\to\infty} H_C e^{-\rho t} = 0, \lim_{t\to\infty} \lambda e^{-\rho t} = 0 \tag{76}$$

Condition (73) guarantees the employment of the labor factor is optimized along the time path $L(t)$. If the firm does make profits, Q–wL>0 and α=1, and from condition (73) we obtain:

$$\lambda = \frac{Q_L}{w - Q_L} = \frac{f(k) - kf'(k)}{w - [f(k) - kf'(k)]} \tag{77}$$

This is the dynamic optimization rule the firm follows in the employment of the labor factor, which corresponds to condition (9) and conditions (22) or (24) from a static point of view. This rule requires that the firm must balance the ratio of the marginal productivity of labor over the wage rate, net of the same marginal productivity of labor, to the shadow price of capital along the entire time path of the costate variable λ.

Condition (74) again represents the dynamics of capital, while condition (75) is the equation of motion of the costate variable λ. Equations (74) and (75) together form the *Hamiltonian* or *canonical system*.

Finally, equations (76) are the transversality conditions. The first one requires the limit of the present-value Hamiltonian vanishes as $t\to\infty$. The second one requires the shadow price of capital in discounted value vanishes as $t\to\infty$. Since the term $e^{-\rho t}\to 0$ as $t\to\infty$, both these transversality conditions are satisfied for finite values of $H_C \neq 0$, and $\lambda \neq 0$. This will be demonstrated in section 3.4.

Condition (74), dividing by L and remembering that α=1, can be put in the form:

$$\frac{\dot{K}}{L} = f(k) - w \tag{78}$$

Since from the definition of k=K/L we have K=kL, and $\dot{K} = \dot{k}L + k\dot{L}$, dividing by L and substituting we obtain:

$$\frac{\dot{K}}{L} = \dot{k} + k\frac{\dot{L}}{L} = f(k) - w \tag{79}$$

and finally:

$$\dot{k} = f(k) - w - k\frac{\dot{L}}{L} \tag{80}$$

Given the real wage rate w, equation (80) describes the dynamics of the capital-labor ratio as a function of the rate of growth of employment $\dot{L}/L$. Analogously, from condition (75), setting α=1, we obtain:

$$\dot{\lambda} = \lambda[\rho - f'(k)] - f'(k) \tag{81}$$

which describes the dynamics of λ, given ρ, as a function of the capital-labor ratio.

Conditions (77), (80) and (81), taken together, form a system of three equations, two of which are differential equations, in three unknowns, given by k, λ and $\dot{L}/L$. These equations fully describe the dynamics of a revenue maximizing firm. Given the real wage rate w and the discount rate ρ, we can find a *steady state* equilibrium for $\dot{k} = \dot{\lambda} = 0$. This equilibrium is characterized by a time constant value of the triple (k, λ, $\dot{L}/L$). Furthermore, our system can be decomposed into two sub-systems: the first one made up by equations (77) and (81), where the variable $\dot{L}/L$ does not appear, and the second one corre-

[4] In Leland's model [28], α is an exogenous parameter on which the firm exerts no influence. This assumption is not realistic and, at the same time, it contributes to limit Leland's analysis.

sponding to equation (80) alone, which is the only equation where the employment rate of growth $\dot{L}/L$ is present. So, the first subsystem made up by equations (77) and (81) can be solved autonomously, and its solution gives, for $\dot{\lambda} = 0$, the steady state values of k and λ. Substituting the equilibrium value of k in equation (80) and setting $\dot{k} = 0$, the second subsystem yields the equilibrium growth rate of employment $\dot{L}/L$.

Now, let us concentrate, first of all, on the subsystem made up by equations (77) and (81), which can be depicted in a phase diagram on the plane k, λ. To construct this, we must pay attention to equation (77). Like in the static analysis, let us indicate with $\hat{k}$ the capital-labor ratio which satisfies the equality between the marginal productivity of labor and the real wage rate, that is $Q_L = f(\hat{k}) - \hat{k} f'(\hat{k}) = w$; and let us indicate with k_w the capital-labor ratio which satisfies the equality between the average productivity of labor and the real wage rate, that is $f(k_w) = w$. From equation (77), since $Q_L \to 0$ as $k \to 0$, whereas the ratio $Q_L/(w - Q_L) \to \infty$ as $(w - Q_L) \to 0$ and $k \to \hat{k}$, we deduce that λ is an increasing function of k, which goes from zero to infinity as k goes from zero to $\hat{k}$. Furthermore, we can find that the sign of the derivative of λ with respect to k is positive. To find this, let us write equation (77) in the implicit form:

$$G_w(k,\lambda) = f(k) - kf'(k) + \lambda[f(k) - kf'(k) - w] = 0 \qquad (82)$$

where G_w is the implicit function existing between k and λ. The sub-index w in G means that this function is defined for any given w. Differentiating (82), we have:

$$\frac{d\lambda}{dk} = -\frac{\partial G_w/\partial k}{\partial G_w/\partial \lambda} = \frac{(1+\lambda)kf''(k)}{f(k) - kf'(k) - w} > 0 \qquad (83)$$

which is positive because for $k < \hat{k}$ we have $w > f(k) - kf'(k)$ and $f''(k) < 0$.

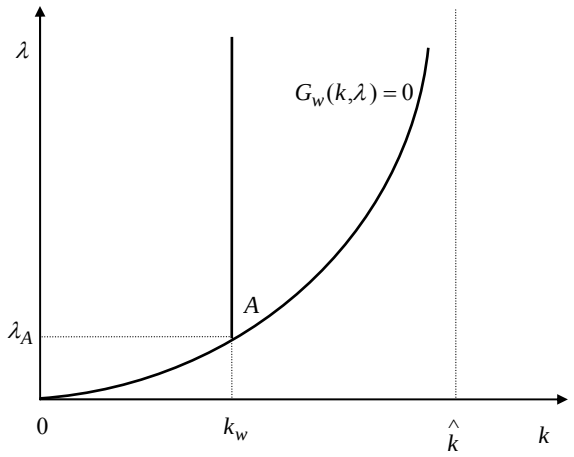


Figure 3. The mapping of the function $G_w(k, \lambda)$

However, equation (77) is not economically meaningful for $0 \leq k < k_w$, because in this range the non bankruptcy constraint is not satisfied. Anyway, for $k = k_w$, the firm does not make profits, so we have $Q - wL = 0$ and α is indeterminate in the open range $\alpha \in [0,1)$. In this case, from (73) we obtain:

$$\lambda = \frac{Q_L}{\alpha[w - Q_L]} = \frac{f(k_w) - k_w f'(k_w)}{\alpha[w - f(k_w) + k_w f'(k_w)]} \qquad (84)$$

where λ is defined with respect to α for $k = k_w$. The mapping of the function $G_w(k,\lambda) = 0$ is depicted in figure 3. The part of this function that has an economic meaning belongs to the range $[k_w, \hat{k})$; for $k = k_w$, the function $G_w(k,\lambda) = 0$ becomes a truncated vertical line, where the truncation point A is given by:

$$\lambda_A = \lim_{k \to k_w} \frac{f(k) - kf'(k)}{w - [f(k) - kf'(k)]} \qquad (85)$$

where the limit is calculated as $k \to k_w$ from the right. For $k_w < k < \hat{k}$, instead, the curve $G_w(k,\lambda) = 0$ is defined by the expression (77).

Since the equilibrium condition the firm must respect in employing labor is satisfied only along the curve $G_w(k,\lambda) = 0$, it follows that the dynamics of the firm can be fully described only by a point of this curve belonging to the range $[k_w, \hat{k})$. To find this point, we need to take into account condition (81), which describes the dynamics of the shadow price of capital λ.

The static analysis has suggested that a profit maximizing firm is in equilibrium in $\hat{k}$, while a revenue maximizing firm is in equilibrium in k_w. Furthermore, if a revenue maximizing firm must also respect a minimum ac

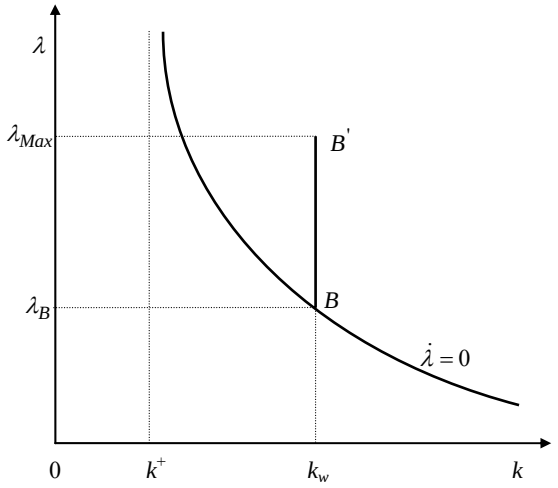


Figure 4. The mapping of the $\dot{\lambda} = 0$ curve

ceptable return on capital constraint, its equilibrium point is intermediate between k_w and $\hat{k}$. We are led to an analogous conclusion in a dynamic context too, but with a fundamental specification. As long as the firm grows, the dynamic equilibrium point is intermediate between k_w and $\hat{k}$. Once the firm stops growing, the stationary equilibrium of a revenue maximizing firm is the same as the static one and corresponds to a capital-labor ratio equal to k_w.

When a revenue maximizing firm must also respect a minimum acceptable return on capital constraint in a dynamic context, we have two cases. If the constraint is not binding, in the equilibrium point the firm grows at a rate which is smaller than the rate of profit. Otherwise, if the constraint is binding, the firm does not grow any more. In the latter case, all profits that the firm realizes are distributed to the stockholders to satisfy the minimum acceptable return on capital constraint, and nothing remains to the firm for self-financing and capital accumulation. The equilibrium point k_0, being $k_w \le k_0 \le \hat{k}$, is the capital-labor ratio to which it corresponds a rate of profit exactly equal to the minimum return on capital rate r_0. Obviously, also in this case the firm rate of growth is zero. We demonstrate all this in the following sections.

3.2. The Equilibrium Capital-labor Ratio

As previously seen, the dynamics of the costate variable λ is given by equation (81). First we need to map in a phase diagram the stationary points of λ, given by:

$$\dot{\lambda} = \lambda [\rho - \alpha f'(k)] - f'(k) = 0 \tag{86}$$

from which we obtain:

$$\lambda(\dot{\lambda} = 0) = \frac{f'(k)}{\rho - \alpha f'(k)} \tag{87}$$

If the firm makes profits, $Q - wL > 0$ and $\alpha = 1$. In this case, equation (87) reduces to:

$$\lambda(\dot{\lambda} = 0) = \frac{f'(k)}{\rho - f'(k)} \tag{88}$$

which describes the stationary points of the shadow price of capital λ with respect to the capital-labor ratio k. The map of this curve, from now on defined as the $\dot{\lambda} = 0$ curve, which is depicted in figure 4, depends on the value of ρ. For $\rho < f'(k)$, the curve $\dot{\lambda} = 0$ gives negative values of λ, which are not economically meaningful. Therefore it must be $\rho \ge f'(k)$ in order to have significant economic values of λ.

Let k^+ be the level of k which satisfies the equation $\rho - f'(k^+) = 0$, so if $k \to k^+$ from the right, then we have $\lambda \to \infty$. Since $f(k)$ decreases with k, for $k > k^+$ we have $\rho - f'(k) > 0$, and $\lambda > 0$, so we have positive values of the shadow price of capital. Furthermore, always for values of $k > k^+$, differentiating equation (86), we have:

$$\frac{d\lambda}{dk} = -\frac{\partial \dot{\lambda}/\partial k}{\partial \dot{\lambda}/\partial \lambda} = \frac{(1+\lambda) f''(k)}{\rho - f'(k)} < 0 \tag{89}$$

The sign of this expression is negative because the denominator is positive, while $f''(k) < 0$. As $k \to k^+$ from the right, $[\rho - f'(k)] \to 0$, so $d\lambda/dk \to -\infty$. This means that the curve $\dot{\lambda} = 0$ is asymptotic as $k \to k^+$ from the right, decreases as k increases and becomes zero as $k \to \infty$, being $\lim_{k\to\infty} f'(k) = 0$.

By definition, the value of k^+ depends on the discount rate ρ. For sufficiently high values of ρ, we find that k^+ has an intermediate value between zero and k_w. In this case, the curve $\dot{\lambda} = 0$ is economically meaningful only for $k \ge k_w$, that is to say only for those values of k for which the non bankruptcy constraint is satisfied. For $k > k_w$, the curve $\dot{\lambda} = 0$ is defined by equation (88), while for $k = k_w$ the curve is defined by equation (87) with a value of α which varies in the open interval $0 \le \alpha < 1$. In the latter case, the stationary values of λ are given by:

$$\lambda(\dot{\lambda} = 0) = \frac{f'(k_w)}{\rho - \alpha f'(k_w)} \tag{90}$$

As α varies, equation (90) describes the segment *BB*', whose extreme value at *B*' (obtained for $\alpha = 0$) is

$$\lambda_{Max} = \frac{f'(k_w)}{\rho} \tag{91}$$

while in point *B* the stationary value of λ is given by:

$$\lambda_B = \lim_{k \to k_w} \frac{f'(k)}{\rho - f'(k)} = \frac{f'(k_w)}{\rho - f'(k_w)} \tag{92}$$

where the limit is calculated as $k \to k_w$ from the right. Overlapping the two curves $G_w(k,\lambda) = 0$ and $\dot{\lambda} = 0$ on the same graph, we can depict the phase diagram of our model. We can have many cases, depending on the relative location of the points *A*, *B* e *B*'. For $\lambda_B > \lambda_A$, where λ_A is defined by equation (85), we have the situation depicted in figure 5. In this case, the dynamic equilibrium of a revenue maximizing firm may occur at two points of the capital-labor ratio, corresponding respectively to $\bar{k}$ and k_w. However, the only dynamic equilibrium is point *E*, to which an endogenous growth of the firm corresponds, which is defined by the pair of values $(\bar{k}, \bar{\lambda})$.

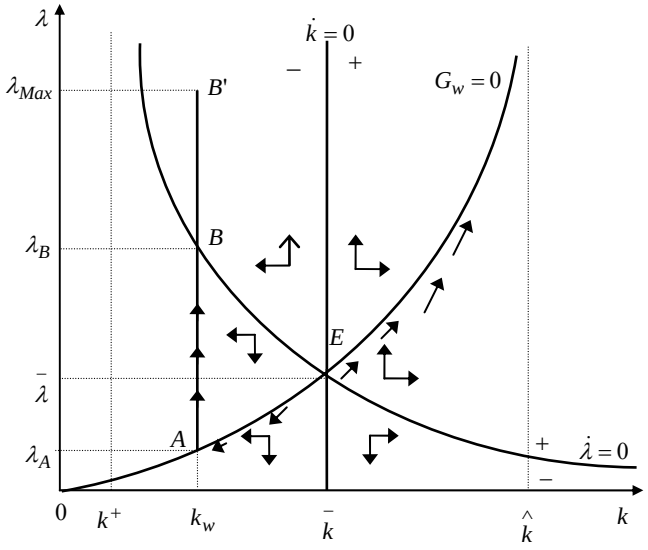


Figure 5. The phase diagram for $0 < k^+ < k_w$

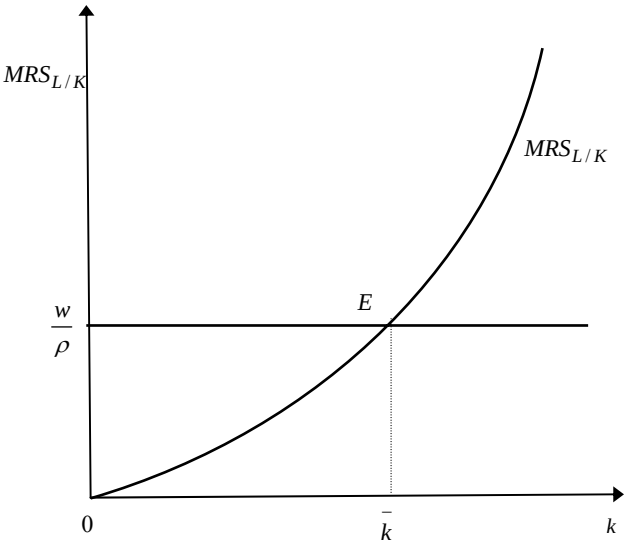


Figure 6. The mapping of the dynamic equilibrium condition $MRS_{L/K} = w/\rho$

At this point, the dynamic equilibrium of a revenue maximizing firm is characterized by the stationarity of both the capital-labor ratio at the level $\bar{k}$ and the shadow price of capital at the level $\bar{\lambda}$. The equilibrium value of the capital-labor ratio $\bar{k}$ can be found taking the equations (77) and (88) as a system of simultaneous equations, whose solution gives:

$$\frac{f(k)-kf'(k)}{w-f(k)+kf'(k)}=\frac{f'(k)}{\rho-f'(k)} \tag{93}$$

while the equilibrium value of the capital shadow price $\bar{\lambda}$ is given by:

$$\bar{\lambda}=\frac{f'(\bar{k})}{\rho-f'(\bar{k})}=\frac{f(\bar{k})-\bar{k}f'(\bar{k})}{w-[f(\bar{k})-\bar{k}f'(\bar{k})]} \tag{94}$$

From equation (93), rearranging we obtain:

$$\frac{f(k)-kf'(k)}{f'(k)}=\frac{w}{\rho} \tag{95}$$

which can be expressed in the form:

$$MRS_{L/K}=\frac{Q_L}{Q_K}=\frac{w}{\rho} \tag{96}$$

where, like in the static analysis, $MRS_{L/K}$ is the rate of marginal substitution between labor and capital. Equation (96) represents the optimizing rule which must be followed by a revenue maximizing firm in a dynamic context. This rule requires that the rate of marginal substitution between labor and capital be equal to the ratio between the real wage rate and the discount rate. Since the marginal productivity of labor increases with respect to the capital-labor ratio, whereas the marginal productivity of capital decreases, the $MRS_{L/K}$ is an increasing function of k. Furthermore, because as $k\to 0$ we have $Q_L \to 0$ and $Q_K \to \infty$, whereas as $k\to\infty$ we have $Q_L \to \infty$ and $Q_K \to 0$, it follows that $MRS_{L/K} \to 0$ as $k\to 0$ and $MRS_{L/K} \to \infty$ as $k\to\infty$. This is shown in figure 6. It is clear from this figure that an increase of the real wage rate w or a decrease of the discount rate ρ imply an increase of the equilibrium capital-labor ratio $\bar{k}$.

In a neighborhood of point E in figure 5, the value of λ tends to depart from its equilibrium value $\bar{\lambda}$. This can be demonstrated by differentiating (81) with respect to k:

$$\frac{\partial\dot{\lambda}}{\partial k}=-\lambda f''(k)-f''(k)>0 \tag{97}$$

which is positive, since $f''(k)<0$. According to (97), as k increases (going rightwards), $\dot{\lambda}$ should follow the (−, 0, +) sign sequence. So, the λ-arrowheads must point downward to the left of $\dot{\lambda}=0$ curve, and upward to the right of it. This means that $\dot{\lambda}<0$, implying λ decreases with respect to time, on the left of the $\dot{\lambda}=0$ curve, while $\dot{\lambda}>0$, implying λ increases with respect to time, on the right of the $\dot{\lambda}=0$ curve.

Since the value of the capital-labor ratio at the point E is constant at the level $\bar{k}$, from (80), which is the second subsystem that describes the dynamics of the revenue maximizing firm, we have:

$$\dot{k}=f(\bar{k})-w-\bar{k}\frac{\dot{L}}{L}=0 \tag{98}$$

from which we deduce:

$$\frac{\dot{L}}{L}=\frac{f(\bar{k})-w}{\bar{k}} \tag{99}$$

Therefore, in the dynamic equilibrium of point E, the

growth rate of employment is given by the return on capital rate corresponding to the equilibrium capital-labor ratio $\bar{k}$. As $\bar{k}$ is stationary, this implies that K grows at the same rate as L. Furthermore, the linear homogeneity of the production function implies that Q increases at the same rate, too. Thus, we have:

$$\frac{\dot{L}}{L}=\frac{\dot{K}}{K}=\frac{\dot{Q}}{Q}=\frac{f(\bar{k})-w}{\bar{k}}=\bar{r} \tag{100}$$

where $\bar{r}$ is the common growth rate of labor, capital and output.

Therefore, in point E we have a balanced growth of the firm, where the rate of growth $\bar{r}$ is endogenously determined. Therefore, this is a neoclassical endogenous model of the revenue maximizing firm, which goes further beyond Leland's analysis.

There is no transitional dynamics in this model. In fact, if the firm begins its activity with the capital $K(0)=K_0$, it must choose from the beginning the level L_0 of the labor factor for which we have $K_0/L_0=\bar{k}$ and it must keep permanently fixed this capital-labor ratio. So, the equilibrium time path of k is constant.

Analogously, the equilibrium time path of the shadow price of capital is also constant at the level indicated by (94), whereas the equilibrium time paths of labor, capital and output depend on the rate of endogenous growth $\bar{r}$ as defined by (100). They are defined respectively by the following exponential functions:

$$L=L_0e^{\bar{r}t} \tag{101}$$

$$K=K_0e^{\bar{r}t} \tag{102}$$

$$Q=Q_0e^{\bar{r}t} \tag{103}$$

where, given K_0, the value of L_0 is defined by $L_0 = K_0 / \bar{k}$, whereas that of Q_0 is given by $Q_0 = Q(K_0, L_0)$.

The $\dot{k}=0$ curve given by (98) must plot as a vertical straight line in figure 5, with horizontal intercept $\bar{k}$. In a neighborhood of point E along the $G_w(k,\lambda)=0$ curve, the value of k, and not only that of λ, tends to depart from its equilibrium value. This can be demonstrated differentiating (80) with respect to k:

$$\frac{\partial\dot{k}}{\partial k}=f'(k)-\frac{\dot{L}}{L} \tag{104}$$

Substituting to $\dot{L}/L$ its value given by (99), in a neighborhood of the point E we have:

$$\frac{\partial\dot{k}}{\partial k}=\frac{w-[f(\bar{k})-\bar{k}f'(\bar{k})]}{\bar{k}}>0 \tag{105}$$

which assumes a positive value because for $\bar{k}<\hat{k}$ the real wage rate is greater than the marginal productivity of labor. This suggests that $\dot{k}$ should follow the (−, 0, +) sign sequence as k increases. Hence the k-arrowheads should point leftwards to the left of the $\dot{k}=0$ curve, and rightwards to the right of it. This means that, for $k<\bar{k}$, we have $\dot{k}<0$, so k decreases with respect to time, while for $k>\bar{k}$ we have $\dot{k}>0$, so that k increases with respect to time.

Therefore, along the $G_w(k,\lambda)=0$ curve, the arrowheads point rightwards and upward to the right of the point E and they point leftwards and downward to the left of that point. So, both the capital-labor ratio k and the shadow price of capital λ, when moving along the $G_w(k,\lambda)=0$ curve to optimize the use of labor, tend to increase to the right of point E and to decrease to the left. It follows that point E is dynamically unstable, and the firm can remain in equilibrium at that point only if it has chosen to stay there from the beginning.

If, on the contrary, the firm begins its activity with a capital-labor ratio greater than $\bar{k}$, the dynamics of the model suggests that the firm must choose a point on the right of point E along the $G_w(k,\lambda)=0$ curve; so, it is induced to over-accumulate capital, approaching asymptotically the capital-labor ratio $\hat{k}$ which maximizes the rate of profit. But, in this case, a dynamically efficient equilibrium point for a revenue maximizing firm does not exist, because when k tends to $\hat{k}$ the shadow price of capital tends to infinity, so it has not a stationary value.

On the contrary, if the firm begins its activity with a capital-labor ratio smaller than $\bar{k}$, the dynamics of the model suggests that the firm must choose a point on the left of the point E along the $G_w(k,\lambda)=0$ curve. So, it is induced, following the arrowheads, to under-accumulate capital moving along this curve towards the point A. In the point A, the value of α becomes indeterminate in the open interval $0\leq\alpha<1$. For $\lambda_B\leq\lambda\leq\lambda_{Max}$, both equations $G_w(k,\lambda)=0$ and $\dot{\lambda}=0$ are satisfied, so the equilibrium capital-labor ratio becomes k_w. In this case, we have multiple equilibria along the segment BB', all characterized by the stationarity of the capital-labor ratio at the level k_w and by zero profits, so the endogenous growth rate of the firm is null.

3.3. Market Conditions and the Role of the Discount Rate

The firm can also be pushed to a stationary point by the market conditions given by the ratio w/ρ. Given the real wage rate w, the ratio k^+ depends, as seen, on the value of the discount rate ρ. If ρ increases, the ratio k^+ in figure

5 decreases, so the $\dot{\lambda}=0$ curve shifts down and left. This implies the point *E* also shifts down and left, along the $G_w(k,\lambda)=0$ curve, as shown by the arrowheads, until it overlaps to point *A*. At the same time, the movement to the left of the $\dot{\lambda}=0$ curve makes the point *B* also to coincide with the point *A*.

When these three points coincide, a revenue maximizing firm does not have an endogenous growth equilibrium any more; it only has multiple stationary equilibria along the segment *BB*', all characterized by the stationarity of the capital-labor ratio at the level k_w. On the contrary, the equilibrium values of the shadow price of capital range from λ_A, which now coincides with λ_B, and λ_{Max}. If we indicate with ρ_w the discount rate to which it corresponds a $\dot{\lambda}=0$ curve whose point *B* coincides with point *A*, so as $\lambda_B=\lambda_A$, the common value of the shadow price of capital is then given by:

$$\lambda_A=\frac{f(k_w)-k_w f'(k_w)}{w-f(k_w)+k_w f'(k_w)}=\frac{f'(k_w)}{\rho_w-f'(k_w)}=\lambda_B \qquad (106)$$

while the value of λ_{Max}, remembering equation (91), is given by:

$$\lambda_{Max}=\frac{f'(k_w)}{\rho_w} \qquad (107)$$

Because in this case $w = f(k_w)$, and taking account of (95), from equation (106) we get:

$$\lambda_A=\lambda_B=\frac{f(k_w)-k_w f'(k_w)}{k_w f'(k_w)}=\frac{w}{k_w\rho_w} \qquad (108)$$

which can also be put in the form:

$$MRS_{L/K}=\frac{Q_L}{Q_K}=\lambda_A k_w=\frac{w}{\rho_w} \qquad (109)$$

If we exclude the last equality, this condition corresponds to condition (56) of the static theory. It says that in a stationary equilibrium the marginal rate of substitution between labor and capital must equal the shadow value of the capital per worker. Furthermore, in a dynamic analysis, it also must equal the ratio between the real wage rate and the discount rate. So, in this case, the stationary values of λ belong to the range:

$$\lambda_A=\lambda_B=\frac{w}{k_w\rho_w}\le\bar{\lambda}\le\frac{f'(k_w)}{\rho_w}=\lambda_{Max} \qquad (110)$$

If ρ increases beyond ρ_w, the $\dot{\lambda}=0$ curve again shifts down and left, so point *B* shifts under point *A*, while point *B*' approaches point *A*. When point *B*' coincides with point *A*, this remains the only stationary equilibrium point. If we indicate with ρ_{Max} the corresponding discount rate, in this case we have:

$$\lambda_A=\lambda_{Max}=\frac{w}{k_w\rho_{Max}}=\frac{f'(k_w)}{\rho_{Max}} \qquad (111)$$

For $\rho>\rho_{Max}$, point *B*' is under point *A*, which implies that $\lambda_{Max}<\lambda_A$, so no equilibrium exists for a revenue maximizing firm.

On the contrary, when ρ decreases, always keeping a fixed real wage rate, the ratio k^+ increases. When k^+ equals k_w, the dynamic equilibrium of the firm becomes a unique equilibrium of endogenous growth, corresponding to point *E* in figure 5. Let us indicate with $\rho^+<\rho_w$ the level of the discount rate for which we have $k^+=k_w$. Then, for $\hat{r}<\rho\le\rho^+$, the revenue maximizing firm has a unique dynamic equilibrium of endogenous growth in point *E*. This equilibrium stays unique until the discount rate takes another critique value given by $\hat{r}$, which corresponds to the equality $k^+=\hat{k}$. In fact, remembering that, according to the (17), we have $\hat{r}=f'(\hat{k})$, from the definition of k^+ we deduce that, when $\rho=\hat{r}$, the equality $k^+=\hat{k}$ is satisfied.

For $\rho\le\hat{r}$, we have $k^+\ge\hat{k}$, so again no equilibrium of a revenue maximizing firm exists. In this case, as we shall see later, the objective integral is no longer convergent. As $\rho\to 0$, also $f'(k^+)\to 0$, so $k^+\to\infty$. This is why, to make the objective integral converging, the condition $\rho>\hat{r}$ must be satisfied, which implies that $\bar{r}<\hat{r}$. This also implies that the case $\bar{k}=\hat{k}$, when the equilibrium points of a revenue maximizing firm and of a profit maximizing firm should coincide, does not exist. In fact, we always have $\bar{k}<\hat{k}$.

Referring once again to figure 5, we can summarize the following cases:

- for $\rho\le\hat{r}$, no equilibrium of the revenue maximizing firm exists;
- for $\hat{r}<\rho\le\rho^+$, we have a unique equilibrium of endogenous growth corresponding to the stationary point $E=(\bar{k},\bar{\lambda})$ in figure 5;
- for $\rho^+<\rho<\rho_w$, we have an endogenous growth equilibrium in point *E* of figure 5 and multiple stationary equilibria without growth corresponding to the interval *BB*' in the same figure;
- for $\rho_w\le\rho<\rho_{Max}$, we only have multiple stationary equilibria without growth in the interval *AB*', being *A*, in this case, over *B*;
- for $\rho=\rho_{Max}$, we have a unique stationary equilibrium without growth corresponding to point *A*=*B*';

- for $\rho > \rho_{Max}$, again no equilibrium of the revenue maximizing firm exists.

3.4. The Present Value, the Net Present Value per Unit of Capital and the Transversality Conditions

In general, for $P=1$ and $\hat{r} < \rho \leq \rho_{Max}$, the present value ($PV$) of future revenues of the firm is:

$$PV = \int_0^{\infty} Lf(k)e^{-\rho t}dt \tag{112}$$

where $L=L_0e^{rt}$ and $L_0=K_0/k$, while r is any endogenous growth rate which must be smaller than $\hat{r}$. Substituting we have:

$$PV = \int_0^{\infty} L_0 f(k)e^{-(\rho-r)t}dt \tag{113}$$

It is clear that PV converges only if $\rho > r$. As $\rho > \hat{r} > r$, this condition is satisfied. Dividing (113) by K_0, we obtain the present value of future revenues per unit of initial capital:

$$\frac{PV}{K_0} = \int_0^{\infty} \frac{f(k)}{k} e^{-(\rho-r)t}dt \tag{114}$$

Given K_0, when L_0 varies between infinity and zero, the capital-labor ratio k varies between zero and infinity, while the instantaneous output per unit of capital $f(k)/k$, remembering (33), varies between infinity and zero, decreasing with respect to k. But, for $k< k_w$, the ratio $f(k)/k$ is not economically meaningful, because it does not respect the non bankruptcy constraint. Whereas for $k \geq k_w$, as in the static analysis, the ratio $f(k)/k$ is economically meaningful. It decreases with k and tends asymptotically to zero as $k \to \infty$. So, its maximum value is given for $k= k_w$, as we can see in figure 7(c).

In the same way, we can define the net present value NPV as follows:

$$NPV = \int_0^{\infty} \left[Lf(k) - wL\right]e^{-\rho t}dt \tag{115}$$

where $L=L_0e^{rt}$ and $L_0= K_0/k$. This equation can be put in the form:

$$NPV = \int_0^{\infty} L_0\left[f(k) - w\right]e^{-(\rho-r)t}dt \tag{116}$$

which is convergent as $\rho > \hat{r} > r$. Dividing (116) by K_0, we obtain the net present value per unit of capital:

$$\frac{NPV}{K_0} = \int_0^{\infty} \frac{f(k)-w}{k}\, e^{-(\rho-r)t}dt = \int_0^{\infty} re^{-(\rho-r)t}dt \tag{117}$$

As said, given K_0, when L_0 varies between infinity and zero, the capital-labor ratio k varies between zero and infinity, while the instantaneous rate of profit varies with respect to k from $-\infty$ to zero in the interval $0< k \leq k_w$, it becomes positive and increasing in the range $k_w < k < \hat{k}$, further it assumes its maximum value $\hat{r} = \left[f(\hat{k}) - w\right] \Big/ \hat{k}$ for $k= \hat{k}$, and then it decreases towards zero as $k \to \infty$. As a consequence, the ratio NPV/K_0 also varies from $-\infty$ to zero in the interval $0< k \leq k_w$, it becomes positive and increasing in the range $k_w < k< \hat{k}$, further it assumes its maximum value for $k = \hat{k}$, and then it decreases towards zero as $k \to \infty$. The mapping of r with respect to k is also depicted in figure 7(c). This corresponds to the mapping of the rate of profit in the static analysis.

The condition $\rho > \hat{r}$, which is necessary for the convergence of the objective integral, is also necessary and sufficient to satisfy the transversality conditions of our optimal control problem. These conditions are:

$$\lim_{t\to\infty} H_c e^{-\rho t} = 0 \tag{118}$$

$$\lim_{t\to\infty} \lambda e^{-\rho t} = 0 \tag{119}$$

which must be satisfied along the optimal paths of the involved variables.

It is immediate to verify the (119), as in equilibrium we have $\lambda = \bar{\lambda}$, where $\bar{\lambda}$ is constant if the firm has an endogenous growth equilibrium in point E of figure 5, whereas it takes a value which ranges between λ_B and λ_{Max} if the firm is in a stationary equilibrium corresponding to the capital-labor ratio k_w. In both cases, as the term $e^{-\rho t} \to 0$ for $t \to \infty$, condition (119) is satisfied.

To verify condition (118), we need to write it in the extended form:

$$\lim_{t \to \infty} \left[Q + \bar{\lambda}\alpha (Q - wL)\right] e^{-\rho t} = 0 \tag{120}$$

If we have a stationary equilibrium, Q and L are constant, so equation (120) is satisfied. If, otherwise, we have an endogenous growth equilibrium, then $Q = Q_0e^{\bar{r}t}$ and $L = L_0e^{\bar{r}t}$, where $\bar{r} = [f(\bar{k}) - w]/\bar{k}$ is the endogenous growth rate. Substituting these values, the condition (120) becomes:

$$\lim_{t \to \infty} \left[Q_0 + \bar{\lambda}\alpha(Q_0 - wL_0)\right]e^{-(\rho-\bar{r})t} = 0 \tag{121}$$

which is satisfied because $\rho > \hat{r} > \bar{r}$.

3.5. The Equilibrium Conditions with a Minimum Acceptable Return on Capital Constraint

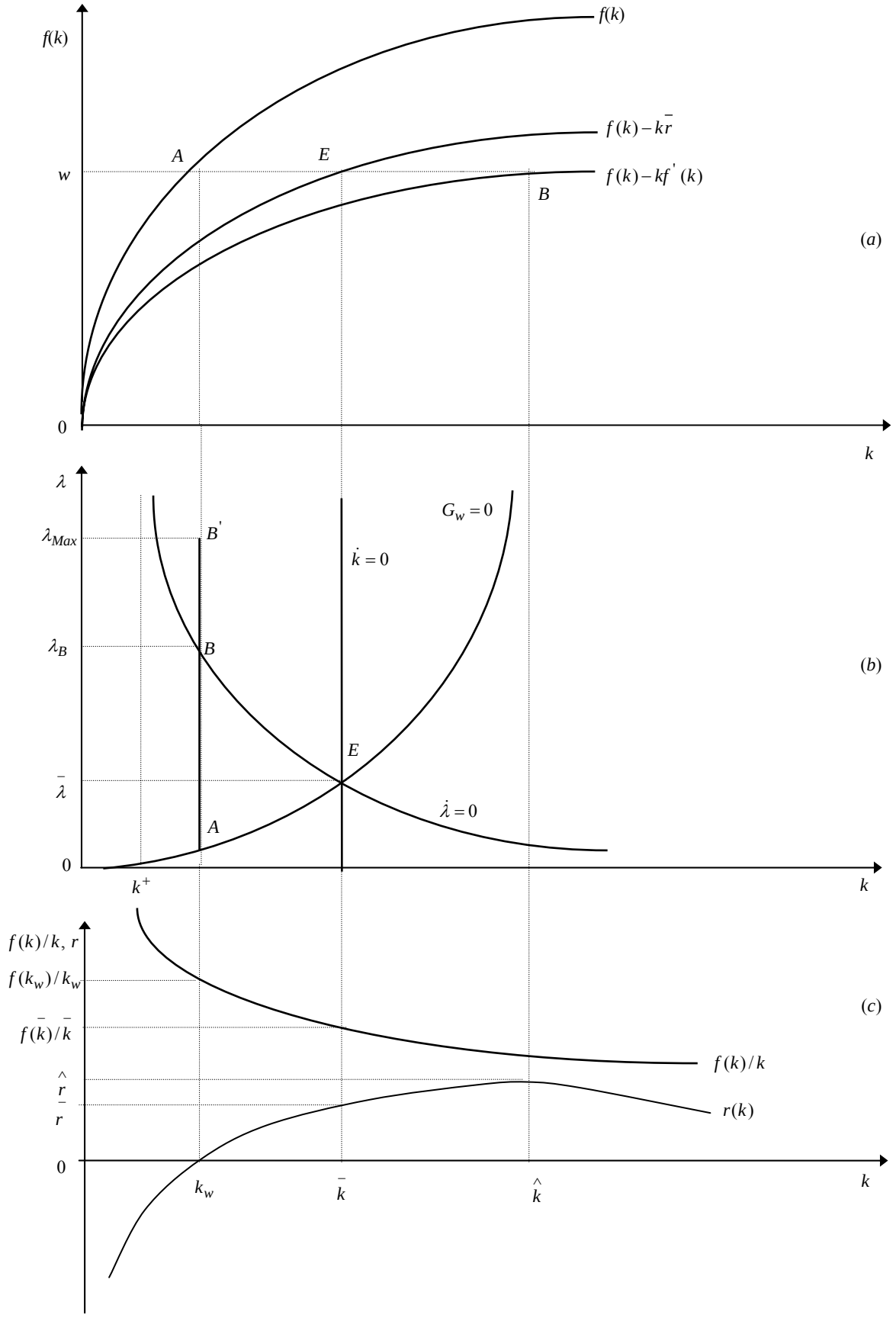


Figure 7. The endogenous growth dynamics of the revenue maximizing firm

If the firm must respect a minimum acceptable return on capital constraint of the form:

$$\frac{\pi}{K} \geq r_0 \text{, for all } t \in [0,\infty] \tag{122}$$

where $\pi = Q - wL$ is the instantaneous profit of the firm, then we must modify the optimal control problem to take account of this. As the constraint can also be expressed in the form:

$$Q - wL - r_0 K \geq 0 \tag{123}$$

the program of the revenue maximizing firm becomes:

Maximize $\int_0^{\infty} Q(K, L)e^{-\rho t}dt$

subject to

$$\begin{aligned} &\dot{K} = \alpha\left[Q - wL - r_0 K\right] \\ &0 \leq \alpha \leq 1 \\ &Q - wL - r_0 K \geq 0 \end{aligned} \tag{124}$$

and $K(0)=K_0$, K_∞ free, K_0, w, r_0 given.

The current value Hamiltonian must, in this case, be extended to form a Lagrangian function in current value, which takes account of the new constraint (123). This Lagrangian is:

$$\Im_c = Q + \lambda\alpha(Q - wL - r_0 K) + \gamma(Q - wL - r_0 K) \tag{125}$$

where λ and γ are two dynamic Lagrangian multipliers expressed in current values. As $\Im_c$ is concave in K and L, and it is linear with respect to α, the following conditions of the maximum principle are necessary and sufficient for the solution of the program (124):

$$\frac{\partial \Im_c}{\partial L} = Q_L + \lambda\alpha(Q_L - w) + \gamma(Q_L - w) = 0 \tag{126}$$

$$\frac{\partial \Im_c}{\partial \gamma} = Q - wL - r_0 K \geq 0,\ \gamma \geq 0,\ \gamma(Q - wL - r_0 K) = 0 \tag{127}$$

$$\dot{K} = \frac{\partial \Im_c}{\partial \lambda} = \alpha(Q - wL - r_0 K) \tag{128}$$

$$\dot{\lambda} = -\frac{\partial \Im_c}{\partial K} + \rho\lambda = -Q_K - \lambda\alpha(Q_K - r_0) - \gamma(Q_K - r_0) + \rho\lambda \tag{129}$$

and in addition we must maximize $\Im_c$ with respect to α. These conditions can be expressed in the following per capita form:

$$G_w(k,\lambda) = f(k) - kf'(k) + (\lambda\alpha + \gamma)[f(k) - kf'(k) - w] = 0 \tag{130}$$

$$f(k) - w - r_0 K \geq 0,\ \gamma \geq 0,\ \gamma[f(k) - w - r_0 k] = 0 \tag{131}$$

$$\dot{k} + \frac{\dot{L}}{L}k = \alpha[f(k) - w - r_0 k] \tag{132}$$

$$\dot{\lambda} = -f'(k) - (\lambda\alpha + \gamma)[f'(k) - r_0] + \rho\lambda \tag{133}$$

If the minimum acceptable return on capital constraint is not binding, from (131) we have γ=0. Furthermore, since $\Im_c$ is linear with respect to α, its value is maximized for α=1. In this case, the equations (130), (132) and (133) respectively reduce to:

$$G_w(k,\lambda) = f(k) - kf'(k) + \lambda[f(k) - kf'(k) - w] = 0 \tag{134}$$

$$\dot{k} + \frac{\dot{L}}{L}k = f(k) - w - r_0 k > 0 \tag{135}$$

$$\dot{\lambda} = -f'(k) - \lambda[f'(k) - r_0] + \rho\lambda \tag{136}$$

From (134) we obtain:

$$\lambda = \frac{f(k) - kf'(k)}{w - [f(k) - kf'(k)]} \tag{137}$$

which is the same as (77). So, all observations just referred to the $G_w(k, \lambda) = 0$ curve also apply to (134).

As to $\dot{\lambda} = 0$ curve, it now depends not only on the discount rate ρ, but also on the minimum return on capital constraint rate r_0. In order to stress this dependence, from now on this curve will be indicated by the symbol $\dot{\lambda}_r = 0$. From (136), setting the stationarity condition, we obtain:

$$\dot{\lambda}_r = -f'(k) + \lambda[(\rho + r_0) - f'(k)] = 0 \tag{138}$$

from which we derive:

$$\lambda\ (\dot{\lambda}_r = 0) = \frac{f'(k)}{(\rho + r_0) - f'(k)} \tag{139}$$

The only difference between this expression and (88) is that here we have $\rho + r_0$ in the denominator instead of only ρ. Hence, introducing a minimum acceptable return on capital rate r_0 is analogous to increasing the discount rate for the same amount.

Therefore, let k^{++} be the level of k which satisfies the equation $(\rho + r_0) - f'(k) = 0$, so that as $k \to k^{++}$ from the right we have $\lambda \to \infty$. Since $f'(k)$ decreases with k, for $k > k^{++}$ we have $(\rho + r_0) - f'(k) > 0$, which implies $\lambda > 0$. Furthermore, if we differentiate equation (138), we obtain:

$$\frac{d\lambda}{dk} = -\frac{\partial \dot{\lambda}_r / \partial k}{\partial \dot{\lambda}_r / \partial \lambda} = -\frac{-f''(k) - \lambda f''(k)}{(\rho + r_0) - f'(k)} = \frac{(1+\lambda)f''(k)}{(\rho + r_0) - f'(k)} < 0 \tag{140}$$

For $k^{++} < k$, this derivative takes a negative value because the denominator is positive, whereas $f''(k) < 0$. As $k \to k^{++}$ from the right, $\rho + r_0 - f'(k) \to 0$, so $d\lambda / dk \to -\infty$. Therefore, the $\dot{\lambda}_r$=0 curve is asymptotic for $k \to k^{++}$ from the right, it decreases when k increases and it takes a null

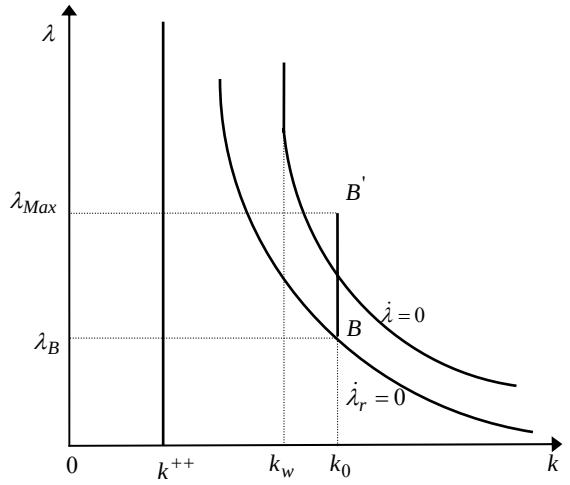


Figure 8: The mappings of $\dot{\lambda}_r = 0$ and $\dot{\lambda} = 0$ curves

value as $k \to \infty$, because $\lim_{k \to \infty} f'(k) = 0$. Furthermore, as we know from its definition, the value of k^{++} depends on the discount rate ρ and on the minimum return on capital rate r_0. Given r_0, for sufficiently high values of ρ, the ratio k^{++} is intermediate between zero and k_0, being k_0 that value of k which satisfies the minimum acceptable return on capital constraint with the equality sign, that is:

$$r_0 = \frac{f(k_0) - w}{k_0} \tag{141}$$

As it is shown in figure 8, the $\dot{\lambda}_r = 0$ curve is economically meaningful only for $k \geq k_0$, that is only for those capital-labor ratios which satisfy the minimum acceptable return on capital constraint. For $k > k_0$, the $\dot{\lambda}_r = 0$ curve is defined by (139), while for $k = k_0$ equation (133) stays true, from which we obtain:

$$\lambda = \frac{f'(k_0) + \gamma[f'(k_0) - r_0]}{\rho - \alpha[f'(k_0) - r_0]} \tag{142}$$

In this case, $\gamma > 0$ because of equation (131), whereas $0 \leq \alpha < 1$, so the maximum value of λ is given by:

$$\lambda_{Max} = \lim_{\alpha \to 0} \frac{f'(k_0) + \gamma[f'(k_0) - r_0]}{\rho - \alpha[f'(k_0) - r_0]} = \frac{f'(k_0) + \gamma[f'(k_0) - r_0]}{\rho} \tag{143}$$

The value of λ_B can be defined as the limit of equation (139) for $k \to k_0$ from the right, that is:

$$\lambda_B = \lim_{k \to k_0} \frac{f'(k)}{(\rho + r_0) - f'(k)} \tag{144}$$

Overlapping the $G_w(k,\lambda) = 0$ curve to the $\dot{\lambda}_r = 0$ curve, we obtain figure 9, which is similar to figure 5, but with the difference due to the minimum acceptable return on

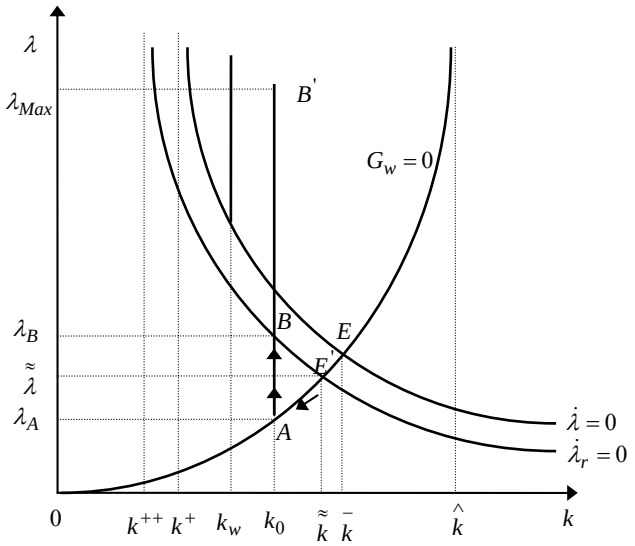


Figure 9. The dynamic equilibrium of a revenue maximizing firm (for $0 < k^{++} < k_0$) with a minimum acceptable return on capital constraint

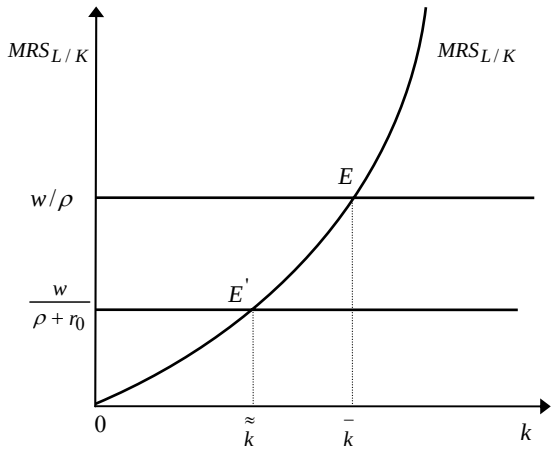


Figure 10. The endogenous growth equilibria of the revenue maximizing firm with and without a minimum acceptable return on capital constraint

capital constraint. We must point out now that the endogenous growth equilibrium is no more at point E, but at point E', where there is the stationarity of the capital-labor ratio at the level $\tilde{k} < \bar{k}$.

Taking the equations (137) and (139) together, we form a system whose solution gives the stationary value $\tilde{k}$, which is the equilibrium capital-labor ratio. Corresponding to this ratio, we have:

$$MRS_{L/K} = \frac{Q_L}{Q_K} = \frac{w}{\rho + r_0} \tag{145}$$

which is the counterpart of equation (96) in the case the firm must respect a minimum acceptable return on capital

constraint. From (145), it is clear that r_0 exerts the same effect as ρ on the equilibrium marginal rate of substitution. In figure 10, both equilibrium conditions, given respectively by (96) and (145), are depicted.

From (135), being $\tilde{\tilde{k}}$ a stationary level of k, at point E' we have $\dot{k}=0$. This means that, remembering α =1, from (132) we have:

$$\dot{k}=f(\tilde{\tilde{k}})-w-r_0\tilde{\tilde{k}}-\frac{\dot{L}}{L}\tilde{\tilde{k}}=0 \tag{146}$$

and the employment growth rate in E' is given by:

$$\frac{\dot{L}}{L}=\frac{f(\tilde{\tilde{k}})-w}{\tilde{\tilde{k}}}-r_0=\tilde{\tilde{r}}-r_0 \tag{147}$$

where $\tilde{\tilde{r}}=\left[f(\tilde{\tilde{k}})-w\right]\Big/\tilde{\tilde{k}}$ measures the actual return on capital rate corresponding to the equilibrium point E', that is to the capital-labor ratio $\tilde{\tilde{k}}$. Equation (147) says that in a dynamic equilibrium the employment of the firm grows at a rate equal to the difference between the real return on capital rate and the minimum acceptable return on capital rate. If we indicate with r^* this difference, taking account of the linear homogeneity assumption on the production function, we have:

$$r^*=\frac{\dot{L}}{L}=\frac{\dot{K}}{K}=\frac{\dot{Q}}{Q}=\tilde{\tilde{r}}-r_0 \tag{148}$$

For r_0 = 0, the dynamic equilibrium of the firm is at point E of figure 9. In this case, $r^*=\tilde{\tilde{r}}=\bar{r}$. For r_0>0, the dynamic equilibrium of the firm is at point E', corresponding to a stationary capital-labor ratio equal to $\tilde{\tilde{k}}<\bar{k}$. When r_0 increases, the point E' moves along the $G_w(k,\lambda)=0$ curve from E to A. For $r_0=\tilde{\tilde{r}}$, when the firm gives out all profits to the stockholders, its dynamic equilibrium is at point A, where the capital accumulation rate and the growth rate are both zero.

We are led to the same conclusion if the minimum return on capital constraint is binding. In this case, the relations (142), (143) and (144) hold, whereas the value of λ_A in figure 9 is given by:

$$\lambda_A=\lim_{k\to k_0}\frac{f(k)-kf'(k)}{w-[f(k)-kf'(k)]} \tag{149}$$

where the limit is calculated from equation (137) for $k\to k_0$ from the right. Therefore, we have multiple stationary equilibria, all corresponding to the capital-labor ratio k_0, which determines an actual return on capital rate equal to r_0, that is:

$$\tilde{\tilde{r}}=\frac{f(k_0)-w}{k_0}=r_0 \tag{150}$$

In this case, we have $r^*=\tilde{\tilde{r}}-r_0=0$, and the firm is in a stationary equilibrium where all profits are given out to the stockholders to satisfy the minimum acceptable return on capital constraint. As a consequence, nothing remains to the firm to finance the accumulation of capital and its growth process.

Finally, according to the respective positions of points A, B and B', we can do the same reasoning as in figure 5. So, let ρ_0 be the value of ρ for which the $\dot{\lambda}=0$ curve passes over point A in figure 9; in order that points A and B coincide, and from equation (149) we have:

$$\lambda_A=\lambda_B=\frac{f(k_0)-k_0f'(k_0)}{w-[f(k_0)-k_0f'(k_0)]}=\frac{f'(k_0)}{(\rho_0+r_0)-f'(k_0)} \tag{151}$$

Substituting the value of $w-f(k_0)$ calculated from equation (150), we obtain:

$$\lambda_A=\lambda_B=\frac{f(k_0)-k_0f'(k_0)}{k_0[f'(k_0)-r_0]}=\frac{Q_L}{k_0(Q_K-r_0)} \tag{152}$$

where Q_L and Q_K are the marginal productivities, respectively, of labor and capital measured for k=k_0. The equation (152) can be put in the form:

$$MRS_{L/K}(r_0)=\frac{Q_L}{Q_K-r_0}=k_0\lambda_A=k_0\lambda_B \tag{153}$$

which corresponds to condition (66) just found in the static analysis. As we know, this condition says that in equilibrium the shadow value of per worker capital must equal the net marginal rate of substitution between labor and capital, where the latter, which is indicated by $MRS_{L/K}(r_0)$, is given by the ratio of marginal productivity of labor over the marginal productivity of capital, net of the minimum rate of return on capital r_0.

4. Concluding Remarks

In conclusion, we have shown that there exist three values of the equilibrium capital-labor ratio that can be found both in a static and in a dynamic analysis of the theory of the firm. They are the ratio $\hat{k}$ where the firm maximizes profits, the ratio k_w where the firm maximizes revenue, and the ratio k_0 where the firm maximizes revenue subject to a minimum acceptable return on capital constraint, if the latter is binding.

Furthermore, we have found two values of the equilibrium capital-labor ratio that exist only in the dynamic analysis of a revenue maximizing firm. They are the ratio $\bar{k}$ where the firm is in an endogenous growth equilibrium,

and the ratio $\tilde{\tilde{k}}$ where the firm, not only is in an endogenous growth equilibrium, but it also satisfies a minimum acceptable return on capital constraint. Obviously, when the firm is in equilibrium in $\tilde{\tilde{k}}$, it grows at a smaller rate than that guaranteed in $\bar{k}$, as the difference is exactly the minimum return on capital rate given out to stockholders to satisfy the constraint.

REFERENCES

[1] Anderson W. L., *Profit Maximizing versus Revenue Maximizing Firms? Only time Will Tell*, mimeo (http://www.mises.org/asc/2003/asc9anderson.pdf), 2003.

[2] Baumol W., *Business Behavior, Value, and Growth*, New York, Macmillan, 1959.

[3] Baumol W., "On the Theory of the Expansion of the Firm", *American Economic Review*, 52, 1962, pp.1078-87.

[4] Blinder A. S. , "International Perspective: Trading with Japan: Why the US Loses even on a Level Playing Field", *Business Economics*, 27, 1992, pp.25-29.

[5] Blinder A. S., "A Simple Note on the Japanese Firm", *Journal of the Japanese and International Economies*, 7, 1993, pp.238-55.

[6] Cheung S. N. S., "The Contractual Nature of the Firm", *Journal of Law and Economics*, 26, 1983, pp.1-21.

[7] Chiang A. C., *Elements of Dynamic Optimization*, New York, McGraw-Hill, 1992.

[8] Coase R. H., "The Nature of the Firm", *Economica*, 4, 1937, pp.386-485.

[9] Conner K. R., "A Historical Comparison of Resource-based Theory of five Schools of Thought within Industrial Organization Economics: Do We Have a New Theory of the Firm?", *Journal of Management*, 17, 1991, pp.121-154.

[10] Conner K. R.-Prahalad C. K., "A Resource-based Theory of the Firm: Knowledge versus Opportunism", *Organization Science*, 7, 1996, pp.477-501.

[11] Cyert R. M.-Hedrick C. L., "Theory of the Firm: Past, Present, and Future; An Interpretation", *Journal of Economic Literature*, 10, 1972, pp.398-412.

[12] Cyert R. M.-March J. G., *A Behavioral Theory of the Firm*, New Jersey, Prentice Hall, 1963.

[13] Fama E., "Agency Problems and the Theory of the Firm", *Journal of Political Economy*, 88, 1980, pp.288-307.

[14] Foss N. J., "Theories of the Firm: Contractual and Competence Perspectives", *Journal of Evolutionary Economics*, 3, 1993, pp.127-44.

[15] Foss N. J., "The Theory of the Firm: The Austrians as Precursors and Critics of Contemporary Theory", *Review of Austrian Economics*, 7, 1994, pp.31-65.

[16] Foss N. J., "Firms, Incomplete Contracts, and Organizational Learning", *Human Systems Management*, 15, 1996*a*, pp.17-26.

[17] Foss N. J., "Knowledge-based Approaches to the Theory of the Firm: Some Critical Comments", *Organization Science*, 7, 1996*b*, pp. 470-6.

[18] Foss N. J., "Austrian Insights and the Theory of the Firm", *Advances in Austrian Economics*, 4, 1997, pp.175-98.

[19] Galbraith J.K., *The New Industrial State*, Boston, Houghton-Mifflin, 1967.

[20] Grant R. M., "Toward a Knowledge-based Theory of the Firm", *Strategic Management Journal*, 17, 1996, pp.109-22.

[21] Jensen M. C., *A Theory of the Firm: Governance, Residual Claims, and Organizational Forms*, Boston, Harvard University Press, 2003.

[22] Jensen M. C.-Meckling W. H., "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, 3, 1976, pp.305-60.

[23] Kagono T.-Nonaka I.-Sakakibara K.-Okumura A., *Strategic vs. Evolutionary Management: A US-Japan Comparison of Strategy and Organization*, Amsterdam, Elsevier, 1985.

[24] Klein B.-Crawford R. G.-Alchian A. A., "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process", *Journal of Law and Economics*, 21, 1978, pp.297-326.

[25] Komiya R., "Japanese Firms, Chinese Firms: Problems for Economic Reform in China", *Journal of the Japanese and International Economies*, 1, 1987, pp.31-61.

[26] Komiya R., *Business Enterprise in Japan*, Boston, The MIT Press, 1992.

[27] Larner R. J., "Ownership and Control in the 200 Largest Non-financial Corporations, 1929 and 1963", *American Economic Review*, 56, 1966, pp.777-87.

[28] Leland H. E., "The Dynamics of a Revenue Maximizing Firm", *International Economic Review*, 13, 1972, pp.376-85.

[29] Nakamura T., *The Postwar Japanese Economy, its Development and Structure*, Tokyo, University of Tokyo Press, 1981.

[30] Nelson R. R.-Winter S., *An Evolutionary Theory of Economic Change*, Cambridge, MA, Harvard University Press, 1982.

[31] Rothbard M. N., *Man, Economy, and State*, Auburn (Alabama), L. von Mises Institute, 1993 ed, 1962.

[32] Tabeta N.-Wang R., "Relative Revenue-maximizing Strategy Under Cournot Duopolistic Competition: the Case of US-Japan Bilateral auto trade", *SABRE Working Papers Series*, n. 1996*a*, pp.8-96.

[33] Tabeta N.-Wang R., *Who Will Be the Winner when a Profit-Maximizer Meets a Revenue-Maximizer?-The case of US-Japan Auto Trade*, mimeo (http://www.ntu.edu.sg/nbs/sabre/ working_papers/19-96.pdf), 1996*b*.

[34] Uekusa M.-Caves R. E., *Industrial Organization in Japan*, Wasinghton DC, the Brooking Institution, 1976.

[35] Van Hilton O.-Kort P.-van Loon P.J.J.M., *Dynamic Policies of the Firm. An Optimal Control Approach*, Berlin, Springer-Verlag, 1993.

[36] Williamson J., "Profit, Growth, and Sales Maximization", *Economica*, 33, 1966, pp.1-16.

[37] Williamson O. E., "Transaction-cost Economics: The

Governance of Contractual Relations", *Journal of Law and Economics*, 22, 1979, pp.233-61.

[38] Williamson O. E., "The Logic of Economic Organization", in O. E. Williamson- S. G. Winter (Eds.), *The Nature of the Firm – Origins, Evolution, and Development*, Oxford, University Press, 1991.

[39] Winter S. G., "Economic Natural Selection and the Theory of the Firm", *Yale Economic Essays*, 4, 1964, pp.225-72.

[40] Winter S. G., "On Coase, Competence, and Corporation", in O. E. Williamson- S. G. Winter (Eds.), *The Nature of the Firm – Origins, Evolution, and Development*, Oxford, University Press, 1991.

[41] Witt U., "Do Entrepreneurs Need Firms? A Contribution to a Missing Chapter in Austrian Economics", *Review of Austrian Economics*, 11, 1999, pp.99-109.

J. Serv. Sci. & Management, **2008, 1: 193-199**
Published Online September 18, 2008 in SciRes (www.SRPublishing.org/journal/jssm)

A Preliminary Examination of Risk in the Pharmaceutical Supply Chain (PSC) in the National Health Service (NHS) (UK)

Liz Breen

Bradford University School of Management, Emm Lane, Bradford, West Yorkshire, BD9 4JL
E-mail:l.breen@bradford.ac.uk

ABSTRACT

The effective management of pharmaceuticals in the National Health Service (NHS) is critical to patient welfare thus any risks attached to this must be identified and controlled. At a very basic level, risks in the pharmaceutical supply chain are associated with product discontinuity, product shortages, poor performance, patient safety/dispensing errors, and technological errors (causing stock shortages in pharmacies) to name but a few, all of which incur risk through disruption to the system. Current indications suggest that the pharmaceutical industry and NHS practitioners alike have their concerns as to the use of generic supply chain strategies in association with what is perceived to be a specialist product (pharmaceuticals).

The aim of the study undertaken was to gain a more realistic understanding of the nature and prevalence of risk in the Pharmaceutical Supply Chain (PSC) to be used as a basis for a more rigorous research project incorporating investigation in the UK, Europe and USA. Data was collected via a workshop forum held in November 2005. The outputs of the workshop indicated that there were thirty-five prevalent risks. The risks were rated using risk assessment categories such as impact, occurrence and controllability. The findings indicated that the risks identified are similar to those prevalent in industrial supply chains, regardless of the idiosyncrasies of pharmaceuticals. However, the group consensus was that caution must be applied in how such risks are addressed, as there are aspects of the product that highlight its uniqueness e.g. criticality.

Keywords: *pharmaceutical supply chain, risk management*

1. Introduction

According to Khan *et al* "*Risk is an ever-present aspect of organisational life"* [16] and *" there is a need to devise robust and well-grounded models of supply chain risk management, which incorporate risk management tools and techniques from other disciplines of research* [16]*"*. Whilst this may be the case, there has been no research to date, which has investigated risks within the total PSC as pertinent to NHS pharmacy. It is vital to do so because pharmaceuticals are a core input into healthcare treatment and are critical products. The authors above would support this sentiment arguing that risk management has only recently been seen as an issue that urgently needs to be addressed in supply chain management.

Academic researchers and practitioners believe that "*pharmaceuticals are different; they cannot be treated like other commodities"* [24]. The reasons for this sentiment were the high cost and long duration for research and development and the repercussions of the product not being available, hence again its criticality. Other unsupported perception-based factors that appear to make this supply chain distinctive include; the level of regulation in the product production, storage, distribution, consumption and the complexity of the fabric of this supply chain [17]. Research conducted by Blackhurst *et al* [3] developed a risk framework based on that formulated by Chopra and Sodhi [4], but it had to be customised further to incorporate industry idiosyncrasies. This may also be an option for the PSC in the NHS (UK).

It would appear, purely from the research findings, that there is concern as to the growing number of risks within the PSC, and the lack of co-ordinated effort in assessing and managing them. This is worrying as the purpose of the PSC is to source pharmaceuticals and materials that can be moved through the supply chain to provide treatment to the end-user. Risks have been identified by this research, which would negatively affect the performance of the total supply chain (from raw material sourcing through to dispensation of medication).

This paper aims to highlight this issue, identifying the nature and prevalence of risk, as determined by supply chain users. The paper will conclude by stating that there needs to be a more co-ordinated approach to and

grounded empirical research conducted into risk management in the total supply chain within NHS hospital pharmacy, as opposed to the current pockets of expertise evident (e.g. Alldred, 2006; Armitage *et al*, 2007; Khan *et al*, 2007; Wills and Stephens, 2007;) [1, 2, 16, 29].

2. Literature Review

The subject of risk is not a new one, but appears to have had a revival or resurgence in supply chain literature [16, 27]. Khan *et al* [16] in their research present an overview of writers investigating varying aspects of risk and supply chain management. Other authors choose to focus on the concept of risk in application [10, 6, 14], sources of risk [27], risk assessment [3, 32, 11], risk mitigation [5, 9], risk modelling [9], flexibility [28], risk management [18, 23, 4], and associated economic repercussions [20, 15].

Supply risk is defined as the potential occurrence of an incident or failure to seize opportunities with inbound supply, the outcome of which results in a financial loss for the firm [32]. The argument has also been put forward that operational risk is a product of governance and regulations [25]. This is particularly relevant to the supply chain under discussion as the PSC fabric is complex (See Figure 1) and is regulated and monitored by parties such as the Department of Health (DOH), National Institute for Health and Clinical Effectiveness (NICE), Medicine and Healthcare Products Regulatory Agency (MHRA), Medicines Controls Agency and the NHS Purchasing and Supply Agency (NHS PASA) to name but a few.

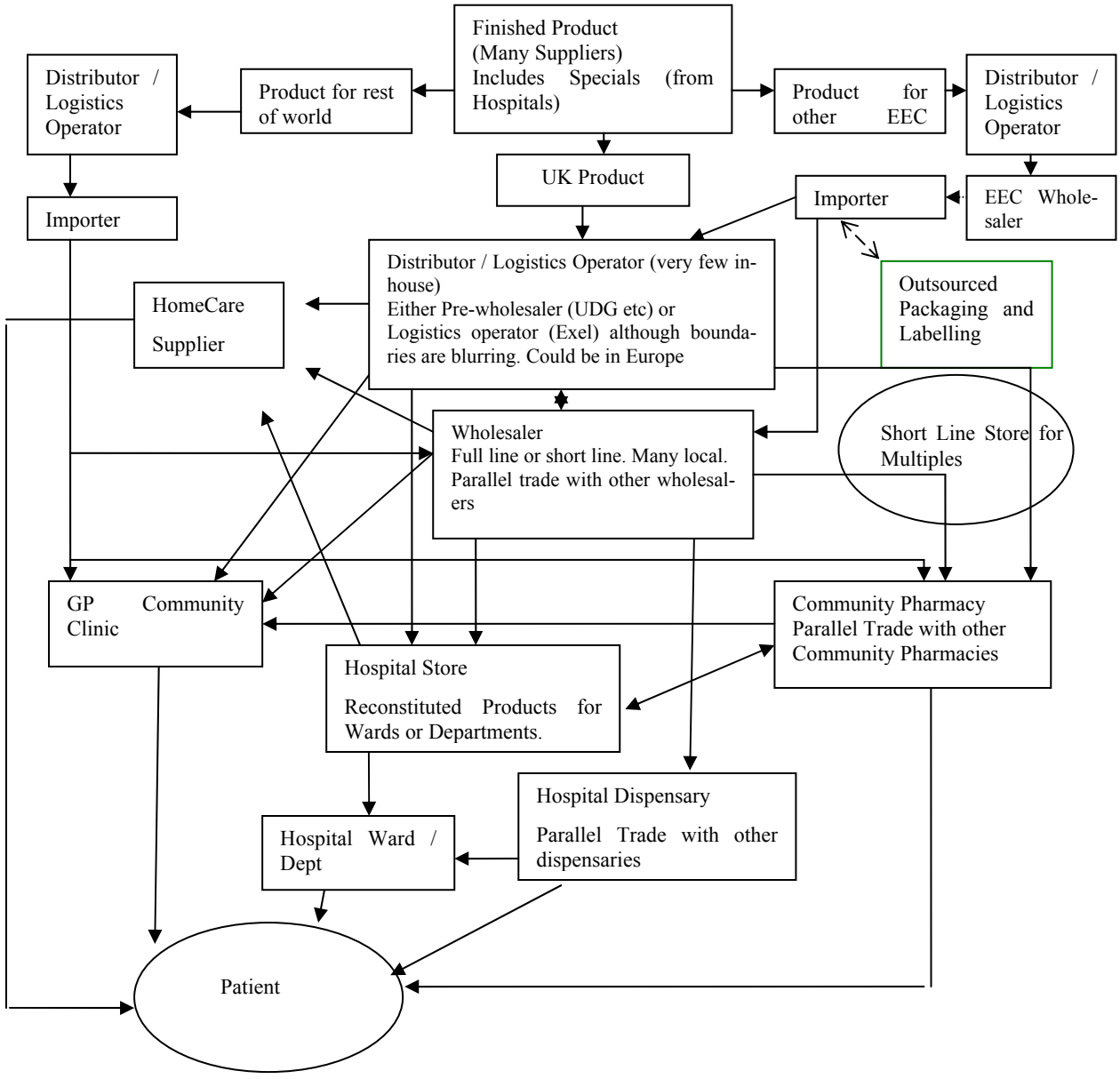


Figure 1. Map of the pharmaceutical supply chain as identified by the risk management workshop (2005)

Cousins *et al* [6] purport that supply chain risk comes in two forms: technological risk and strategic risk. Both types focus on an over-reliance and hence increased potential risk from product, process and technology and suppliers. At a very basic level, risks in the PSC are associated with product discontinuity, product shortages [17], poor performance, patient safety/dispensing errors, technological errors (causing stock shortages in pharmacies), internet pharmacies and counterfeit drugs; all of which cause delays in the system and cause anguish to the final users, the patients. All of these fall into the categories proposed by Cousins *et al* [6]. Costs associated with mismanagement of risks are also the focus of many authors e.g. Papadakis, 2006 and Hendricks and Singhal, 2003. Within the PSC, medication errors alone equate to £200-400 million per year in the UK, and to this must be added the unknown cost of errors in primary care and litigation [19].

Tang and Tomlin [28] assert that due to the spread and complexity of supply chains they are usually slow to respond to environmental changes and thus are less robust when faced with business disruptions. Like such businesses, there is evidence to suggest that the convoluted nature of the PSC in the NHS (UK) is similar to pharmaceutical supply chains in other countries, and similar issues exist such as counterfeit medications, product shortages etc. [30]. Research has indicated that in Europe medicines can travel through as many as 20-30 pairs of hands before it finally reaches the patient [7]. The supply chain has now become more fragmented with effectively 25 pharmaceutical markets in Europe, which has led to a decrease in transparency in the supply chain [8].

The structure of the PSC is such that an examination focusing on risk needs to encompass the complete supply chain and composite network of buyers and suppliers. In which case the total supply chain needs to be the subject of assessment as opposed to individual entities or parties e.g. risks attached to a supplier or purely to patients. Adopting a broad and encompassing view on this issue and not focusing on individual entities is critical in examining this area [16].

3. Methodology

The aim of this research study was to gain a more realistic understanding of the nature and prevalence of risk in the PSC. In order to do this a workshop was held in November 2005 focussing on risk identification within the PSC. This was attended by twenty key PSC stakeholders including pharmaceutical suppliers and wholesalers, NHS professionals and government bodies.

The attendees were split into two distinct groups which contained a mix of pharmaceutical manufacturers and wholesalers and NHS personnel. The two groups were facilitated by NHS pharmacy procurement specialists. The groups were asked to deliver key objectives; these being to identify risks in the PSC, to rate their criticality and to produce an agreed map of the structure of the PSC. Both groups worked individually and were brought back together for interim group sessions (IGS) to deliver their outputs. A consensual set of data was then compiled by the researcher in the presence of the total group, based on their views and comments. At each IGS discussion was generated in order to validate or explain outputs.

After a comprehensive list of risks had been identified the individual groups reconvened to apply a criticality rating. The criteria used to develop such ratings were standard of those currently used within the PSC i.e. impact, control and occurrence. Again, during the IGS there was some discussion and movement on the ratings, and a comprehensive list was compiled. As this study was a preliminary investigation into this area, further extrapolation and development of the risk ratings was not conducted. Further examination using tools such as modelling or Analytic Hierarchy Process (AHP) could be applied using the ratings identified to assist in more structured decision making [12]; however as this was a preliminary investigation they were not employed at this stage in the study.

Further discussion took place concerning the structure of the PSC. Due to time constraints the final map was completed after the workshop by an NHS pharmacy procurement specialist, but circulated to the workshop group for further comments and amendments. However, no amendments were requested. It was felt that a workshop was the most appropriate approach to adopt in collating the necessary data as there was a lack of published data on this subject, and the data produced would be accurate and timely. As a cross-section of companies were involved (17 companies/organisations represented) then a tentative assumption could be made that the findings would be more representative of opinion on this subject within this field than not.

4. Results

The outputs of the workshop indicated that there were 35 prevalent risks (See Table I for summary); with varying levels of criticality and that the structure of the PSC map was product-dependent, so taking a specific product into consideration the map profile could be substantially different. It was agreed that the top 10 risks identified were representative of the state of play in the PSC. The top rated risks included fragmentation of the supply chain (multiple channels leading to poor communication); lack of visibility concerning placement and availability of stock, inappropriate forecasting conducted by the customer and a general inability of capacity to meet demand. An outlying risk that was debated on the day was that of counterfeit medication. One member of the group was

Table 1. Risks identified and associated ratings

Risk	Rating	Risk	Rating
Fragmentation of SC – no single source, multiple channels, no communication, unilateral decisions	10	Too much information	6.5
Lack of visibility of stock	9	Short term SC planning	6.5
Unexpected increase in demand	8.5	Operational in/efficiencies e.g. systems operating properly	6
Demand versus Capacity	8.5	Non standard practice – customised policies per hospital. Lack of common codes etc	6
Information flow or lack of information	8.5	Counterfeiting	6
Lack of forecasting – customer side	8.5	Increase in demand due to NICE approval, patient involvement, press	6
Availability of raw material – true and commercially induced. Regulatory issues – manufacturing licensing/change of standards/drug recalls	8	Rationalisation of range	5.5
Demand/economics – not able to respond to demand	8	Cash flow/cash management – threat associated with small companies and hospitals	5.5
Inadequate buffer stock – JIT/lean	8	Storage/cold chain	5.5
Contracting treated as a commodity – big contracts equals big risk. Drive competitors out of market	8	Reimbursement policies not consistent	5.5
Transportation – unavailability of fuel, congestion, weather, illness	7.5	Response of industry to shortages – communication	5.5
Manufacturer defence tactics	7.5	Loss of expertise – unsophisticated purchasing/practice?	5
Diversion of manufacturing capacity	7.5	Risk of litigation – influence on market	5
External influences – disaster recovery	7.5	Lack of knowledge regarding manufacturing process or source of supply	4.5
Stock holding – more concentrated	7	Procurement Hubs – introduce more complexity	4.5
Exploitation	6.5	Theft	4.5
Dispensing/picking error – medication/packaging, prescription management	6.5	Prioritisation – conflict between patients/profits	4
Decrease in capacity linked to profit	6.5		

adamant that the risk attached to this activity was greater than the rest of the group perceived it to be. This risk was placed as 23 out of 35.

The findings indicated that the risks identified are similar to those prevalent in industrial supply chains crossing various categories e.g. financial, legal and operational, regardless of the idiosyncrasies of pharmaceuticals. However, the group consensus was that caution must be applied in how such risks are addressed, as there are aspects of the product that highlight its uniqueness e.g. criticality. Similar analysis was conducted by Blackhurst *et al* [3] into the automotive industry, the focus however being on supplier risk assessment and monitoring. An outcome of this study was the acknowledgement that a standard risk framework did not fit that company in that industry and therefore a standard tool had to be customised for the purpose of the research.

5. Discussion

The ratings produced were based on the following criteria; control, occurrence and impact. The risks identified fall into typical risk categories such as legal, operational, financial and strategic, all of which need to be addressed to achieve effective risk management [26]. A high rating indicated a greater potential for a more detrimental impact (financial or other) on the PSC thus needing more structured recovery mechanisms.

The findings can be discussed in three distinct sections: Supply chain structure - The highest ranked rating was due to the fragmentation of the supply chain (10/10). The group felt that there was a lack of uniformity in decision-making within the PSC which led to such problems, and affected the efficacy of the complete supply chain. This being the case, this was a risk that needed to be addressed urgently as it affected all parties and could result in financial loss. This view further supports the concerns within the industry of the increasing involvement of suppliers,

manufacturers, parallel importers, generics, and wholesalers to name a few.

As there are numerous players and communications within the PSC, the supply chain would benefit from having a coordinating body that is responsible for setting targets and meeting deadlines, and implementing strategy. This body needs to recognise the interconnectivity between members of this network and aim to support and nurture this, whilst ensuring that risk is not passed from one party to another [13]. Attention must also be paid to instigating mitigation strategies that positively impact on the total supply chain (all entities) and not exacerbate risk as a repercussion of the initial activity. In the case of the PSC, mitigating a risk attached to a product line should not increase the risk associated with another product line, or the relationship between the buyer and supplier.

Controllability – On looking at the risks presented it is clear that a certain number of risks proposed are operational and functional and are therefore within the control of the industry e.g. visibility of stock, communication channels, capacity management issues and information flow. This being the case, there is a need for more effective management of these risks. The industry is less capable of controlling other risks e.g. counterfeiting. This was rated as a 6/10, which is still high risk but was not considered high priority due to the low frequency in its occurrence. This is such a present threat that a new group (European Alliance for Access to Safe Medicines) was launched late 2007 "to campaign for the exclusion of counterfeit medicines from the supply chain" (Pharma Anti-Counterfeiting, 1:2008). Governmental influences are also less controllable e.g. the conflict between patients and profits, NICE, and the introduction of Procurement Hubs (Collaborative Procurement Organisations) whilst others are totally uncontrollable e.g. unavailability of fuel and natural disasters and illness. In reviewing the risks produced, it can be assumed that some of the risks can effectively be addressed through better co-ordination and management of the PSC and some through effective mitigation strategies as and when they arise.

Strategy – The findings of this research indicate that there are clearly identifiable risks in the NHS PSC (UK) as determined by industry practitioners. The practitioners were also in agreement that although there is an emphasis on risk from the various agencies, companies and the NHS itself, there is no co-ordinated strategy which governs risk across all the supply chain parties. Anecdotal evidence would suggest that even though practitioners know that there is risk attached to core activities within the PSC e.g. procurement contracting, there is no actual risk assessment conducted on suppliers to ascertain the level of risk in a proposed contractual relationship (a key risk being the potential to disrupt supply). Decision-making appears to reside with experienced members of staff; pharmaceutical experience not risk management. According to Blackhurst *et al* [3], like issues, e.g. supplier disruptions, led to their investigation into supplier risk assessment and monitoring and development of a risk factor framework.

Disruptions within the supply chain are a major source of risk. The findings from the risk management workshop indicated that supply disruptions could be seen in the form of availability of raw materials, transportation, disaster recovery, lack of knowledge regarding the source of supply, rationalisation of product range and theft. Any of these could restrict or stop the flow of products through the supply chain, increasing the risk to the patient.

6. Conclusions

The aim of this research study was to gain a more realistic understanding of the nature and prevalence of risk in the PSC as a preliminary research exercise. The approach adopted was qualitative and exploratory in nature. The aim was realised through the collation of data from a risk management workshop. Risk analysis in the PSC in the NHS (UK) is of key importance and has a valuable input into both practice and policy; therefore research into better understanding and management of this is justified. It affects not only the practitioners and policy makers, but also the public, as current and future users of this service. Further investigation conducted within this area could yield the following benefits; a grounded empirical research study, greater visibility of pharmaceutical supply chain activities and players, identification and rating of risks, more structured planning ability to strengthen practice and systems, informed contingency/recovery planning, more effective management of the impact of high severity events on recipients, e.g. unknown disasters, and reduce the impact of low severity events such as a breakdown in buyer-supplier relationships.

However, whilst some of the industry seems to be very active in pursuing this subject, others do not seem to join in their enthusiasm. The twenty contributors to the risk management workshop were representative of a cross-section of this supply chain, thus their views are representative of general feeling and concern regarding the level of risk attached to key issues. The views produced were consensual, the key issues and ratings discussed and agreed by the whole group. This is evidence that despite the complexity of the PSC there is a willingness to come together as a joint body to discuss topical matters. This activity was co-ordinated by a neutral party (University of Bradford School of Management) who were impartial in the proceedings.

From the research findings recommendations can be made to develop policy and practice within this area. These include adopting a structured approach to understanding the nature of risk in the total pharmaceutical supply chain in order to effectively manage it. This would involve detailed analysis of the various party and agency

activities concerning risk management, mitigation activity and successes. At present the researcher could find little evidence of this being done. Consultations should be performed on a continuous basis with PSC participants, representative of the total supply chain. Adequate training packages also need to be developed and disseminated to decision-makers within the PSC concerning the presence of risk and its mitigation strategies. This would involve strategists, policy makers, procurement bodies/staff and key pharmaceutical personnel e.g. pharmacy specialists. Finally, decision-making should be assisted by risk management specialists.

There is evidence within NHS literature and practice that risk assessments and risk management practice is conducted as a matter of expected and good practice. This analysis did not attempt to examine risk assessment and management in healthcare from a clinical perspective, but cannot ignore the fact that risk management practice appears to be well developed in this area, possibly because risk is perceived to be greater the closer to the patient, or because there is high level visibility and public accountability at the coal face so to speak.

According to Khan *et al* [16], risks cannot be understood simply at an individual buyer and supplier level. The authors believe that an examination of supply chain risk must incorporate empirical research into how risk is managed. Within the PSC in the NHS, there is no evidence however that there is a ongoing research into identifying and managing risk across the total pharmaceutical supply chain in the NHS, which is astonishing considering the impact of supply disruptions. There is a gap in current research in this area which needs to be addressed, as identified in part by this preliminary study. Further research aims to redress this. When we think of the most basic risk in the supply chain, i.e. supply disruption, we are reminded of the fact that if pharmaceutical products or associated equipment are delayed in reaching hospitals or pharmacies, patients could die. This is not being melodramatic, this is stating a fact.

7. Future Research

In order to build on the findings presented a future research agenda has been identified. The aim of this project will be to develop a typology of risk in the PSC as applied to NHS hospital pharmacy and provide appropriate and realistic mitigation strategies to manage risk more effectively. The research will encompass data collection from within the UK, USA and Europe. The outputs of which will indicate areas of success or centres of excellence in practice and outcome e.g. in product sourcing, supplier management etc., which should be developed further and shared nationally and internationally.

REFERENCES

[1] Alldred, A. Purchasing safer medicines – a national risk assessment strategy. *Hospital Pharmacist*, January, 3, 2006, pp.17-19.

[2] Armitage, G., Newell, R. and Wright, J. Reporting drug errors in a British acute hospital trust. *Clinical Governance: An International Journal,* 12 (2), 2007, pp.102-114.

[3] Blackhurst, J.V., Scheibe, K.P. and Johnson, D.J. Supplier risk assessment and monitoring for the automotive industry. *International Journal of Physical Distribution and Logistics Management,* 38 (2), 2008, pp.143-165.

[4] Chopra, S. and Sodhi, M.S. Managing risk to avoid supply-chain breakdown*. MIT Sloan Management Review,* 46 (1), 2004, pp.53-61.

[5] Christopher, M, and Lee, H. Mitigating supply chain risk through improved confidence. *International Journal of Physical Distribution and Logistics Management,* 34 (5), 388-396.

[6] Cousins, P., Lamming, R.C. and Bowen, F. (2004) The role of risk in environment-related initiatives. *International Journal of Operations and Production Management*, 24 (6), 2004, pp.554-565.

[7] Haigh, J. IMS Global Consulting, quoted in Parallel Trade in Medicines, *Social Market Foundation*, 2004.

[8] European Federation of Pharmaceutical Industries and Associations *Pharmaceutical Supply Chain Evolution*, EFPIA Position Paper, 2005.

[9] Faisal, M.N., Banwet, D.K. and Shankar, R. Supply chain risk mitigation: modelling the enablers. *Business Process Management Journal*, 12 (4), 2006, pp.535-552.

[10] Fraser, I. and Henry, W. Embedding risk management: structures and approaches. *Managerial Auditing Journal*, 22 (4), 392-409.

[11] Frosdick, S. (1997) The techniques of risk analysis are insufficient in themselves. *Disaster Prevention and Management,* 6 (3), 2007, pp.165-177.

[12] Gaudenzi, B. and Borghesi, A. Managing risks in the supply chain using the AHP method. *The International Journal of Logistics Management*, 17 (1), 2006, pp.114-136.

[13] Hallikas, J., Karvonen, I., Pulkkinen, U., Virolainen, V.M. and Tuominen, M. Risk management processes in supplier networks. *International Journal of Production Economics*, 90, 2004, pp.47-58.

[14] Harland, C., Brenchley, R. and Walker, H. Risk in supply networks. *Journal of Purchasing and Supply Management*, 9, 2003, pp. 51-62.

[15] Hendricks, K.B. and Singhal, V.B. An empirical analysis of the effects of supply chain disruption on long-run stock price performance and equity risk of the firm. *Production and Operations Management,* 14 (1), 2005, pp.35-52.

[16] Khan, O., Burnes, B. and Christopher, M. Risk and supply chain management: creating a research agenda. *The International Journal of Logistics Management,* 18 (2), 2007, pp.197-216.

[17] Knight, L.K. *Roles, Relationships and Dynamics of the Pharma Network. A discussion document on opportunities for supply research,* Draft internal report to the NHS Purchasing and Supply Agency 2005.

[18] Manuj, I and Mentzer, J.T. Global supply chain risk management strategies. *International Journal of Physical Distribution and Logistics Management,* 38 (3), 2008, pp.192-223.

[19] Matthew, L. and Bain, T. Bridging the Gap. Patient Safety: the role of the National Patient Safety Agency (NPSA) in helping pharmacists to improve patient safety. *Pharmacy Management*, 22 (3), 2006, pp.2-8.

[20] Papadakis, I. Financial performance or supply chains after disruptions: an event study. *Supply Chain Management: An International Journal*, 11 (1), 2006, pp.25-33.

[21] Pharma Anti-Counterfeiting News *Launch of a New European Safe Medicines Group*, (3), 2008, pp.1-4.

[22] Peck, H., Drivers of supply chain vulnerability: an integrated framework. *International Journal of Physical Distribution and Logistics Management,* 35 (4), 2005, pp.210-232.

[23] Ritchie, B. and Brindley, C. Supply chain risk management and performance. A guiding framework for future development. *International Journal of Operations and Production Management,* 27 (3), 2007, pp.303-322.

[24] Savage, C.J., Roberts, K.J., and Wang, X.Z. A holistic analysis of pharmaceutical manufacturing and distribution: are conventional supply chain techniques appropriate? *Pharmaceutical Engineering* July/August, 2006, pp.10-18.

[25] Smallman, C. Knowledge management as Risk management: A Need for Open Governance? *Risk Management, an International Journal,* 1 (4), 1999a, pp.7-20.

[26] Smallman, C. The Risk Factor *Financial Times Mastering Management Review*, 30, 1999b, pp.42-45, December.

[27] Smallman, C. Risk and organisational behaviour: a research model. Disaster Prevention and Management, 5 (2), 1996, pp.12-26.

[28] Tang, C. and Tomlin, B. The power of flexibility for mitigating supply chain risks. *International Journal of Production Economics*, doi:10.1016/j.ijpe.2008.07.008.

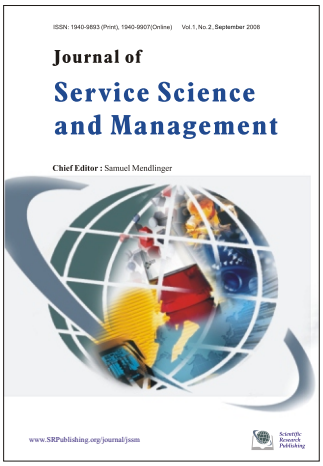
[29] Wills, S. and Stephens, M. How safe is information about medicines? A risk assessment framework. *Clinical Governance: An International Journal*, 12 (1), 2007, pp.2-37.

[30] World Health Organisation *Counterfeit medicines*. Accessed: 12/1/07, 2006.

[31] http://www.who.int/mediacentre/factsheets/fs275/es/.

[32] Zsidisin, G., Melnyk, S.A., Ragatz, G.L. and Burns, L.A. Preliminary findings from a supply risk audit instrument. *Proceedings to IPSERA Conference*, Paper 43, 2006.

Call for Papers



Journal of Service Science & Management (JSSM)

ISSN 1940-9893 (Print), 1940-9907 (Online)
www.srpublishing.org/journal/jssm/

JSSM is an international multidisciplinary journal with the emphasis laid on the service innovation in the global economy and entrepreneurship, the latest management technologies. It also explores the contributions of knowledge discovery and information management research and applications. The goal of this journal is to keep a record of the state-of-the-art research and promote the fast moving service science and management technologies.

Editor-in-Chief

Editorial Board (According to Alphabet)

Subject Coverage

All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed in hard copy. The journal publishes the highest quality, original papers included but not limited to the fields:

- Service Science
- Business Intelligence
- Operational Research
- Computational Economics
- Financial Engineering
- Decision Support System
- Business Process Re-engineering
- Data Mining and Knowledge Discovery
- Innovation and Entrepreneurship
- Risk Management
- Quality Management
- Project Management
- Supply Chain Management
- Software Engineering Management
- Environment and Energy Management
- Knowledge Management and Semantic Web
- Information System Management
- Customer Capital Management
- Human Resources Management

We are also interested in:
Short reports—Discussion corner of the journal:
2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data.
Case studies—Rather than present a classical paper.
Book reviews—Comments and critiques.

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Contact Us:

E-Mail: jssm@srpublishing.org

J. Serv. Sci. & Management, 2008, 1
Published Online September 2008 in SciRes (www.SRPublishing.org/journal/jssm)

CONTENTS

Vol. 1 No. 2 **September 2008**