THREE ESSAYS ON THE ECONOMICS OF INFORMATION TECHNOLOGY INNOVATION

A Thesis
Presented to
The Academic Faculty

by

Zhe Qu

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in the
College of Management

Georgia Institute of Technology

August 2008

THREE ESSAYS ON THE ECONOMICS OF INFORMATION TECHNOLOGY INNOVATION

Approved by:

Dr. Sridhar Narasimhan, Advisor
College of Management
Georgia Institute of Technology

Dr. Marie Thursby
College of Management
Georgia Institute of Technology

Dr. Haizheng Li
School of Economics
Georgia Institute of Technology

Dr. Sandra Slaughter
College of Management
Georgia Institute of Technology

Dr. Han Zhang
College of Management
Georgia Institute of Technology

Date Approved: 15 May 2008

ACKNOWLEDGEMENTS

I thank my dissertation advisor, Dr. Sridhar Narasimhan, for his enormous guidance, encouragement and support throughout my doctoral research and study. I also thank Dr. Marie Thursby, Dr. Sandra Slaughter, Dr. Han Zhang and Dr. Haizheng Li for serving on my committee and contributing invaluable insight and suggestions for this research. I am grateful to the faculty who helped me in the journey of writing the dissertation – Dr. Jerry Thursby, Dr. Chris Forman, Dr. Stuart Graham, Dr. D. J. Wu, Dr. Saby Mitra, Dr. Eugene Comiskey, and Dr. Eric Overby. I thank my fellow students - Fang Zhong, Jifeng Luo, Sam Ransbotham, Anne Fuller, and Yi Liu. I am also appreciative of the assistance of the TI:GER program and the Alan & Mildred Peterson Foundation.

I thank Swasti Gupta-Mukherjee for always having confidence in me, understanding my frustrations and dilemmas and helping lift my spirits along the journey. I also thank my friends Sandy Dai and Yi He for being my cheering squad. I am grateful to my parents, Lianzhong Qu and Ling Wang, and my sister, Na Na, for their immense love and support. And finally, I thank my husband, Jun Li, for providing an ear to listen and a shoulder to cry on, and bringing in fun and hope to this supposedly long and sometimes lonely journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

There are three essays on the economics of information technology innovation in my dissertation:

1. Procurement contracting strategies in a hierarchical supply network
2. R&D offshoring and technology learning in emerging economies – firm level evidence from the information technology industry
3. Software design strategies in markets with open source competitors

The first essay addresses the impact of an information technology enabled hierarchical supply structure on a firm's procurement strategies.

The second essay investigates information technology hardware innovation. I examine R&D offshoring of information technology hardware firms and its impact on R&D effort of firms in host countries.

The third essay focuses on software innovation. I investigate open source software and its impact on the design of proprietary software in terms of number of features bundled in the software.

CHAPTER 1

INTRODUCTION TO THESIS

Information technology (IT) innovation is the key driving force of the development of IT industry; successful application of IT innovation has fundamentally transformed the business practice of numerous industries and firms. The first essay in the thesis investigates the impact of advances in IT on firm procurement contracting strategies. The second essay studies the R&D offshoring strategies of multinationals IT firms and the impact of multinational R&D offshoring on the innovation incentive of host country firms. In the third essay, I explore software design strategy of proprietary software vendors in the presence of open source competitors.

The second Chapter of the thesis investigates procurement contracting strategy in a hierarchical supply network. Much of the existing research assumes a flat distribution structure where each supplier of an individual good is an independent entity in a pool of suppliers. Recent advances in IT have been instrumental in creating, enabling and supporting a hierarchical distribution structure where entities across distribution levels not only exhibit different characteristics (e.g. price, capacity), but also a strong interrelationship. This creates complexity in evaluating contracting decisions for a firm. This essay presents an optimization model that formalizes this problem. The two research questions are: How does the existence of a hierarchical structure of supplier distribution network affect the optimal procurement decisions of a buyer? Under what industry conditions should a buyer undertake a consideration of the full hierarchical procurement structure? The computational study shows that considering the hierarchical structure impacts buyer procurement contracting strategy. In general, buyers can achieve near optimal results by contracting with one level of the suppliers within the hierarchy; i.e., pure strategy is near optimal. However, no single pure strategy dominates the other pure strategies. The level of suppliers with which buyers should contract is strongly dependent on buyer type and the major uncertainty buyer faces. The solutions also show that the loss of not considering the hierarchy and selecting the wrong level of suppliers could be substantial. This study contributes to the literature of procurement strategy in a complex supplier network. The findings strongly indicate the importance of taking into consideration the existence of hierarchical distribution structure at the supplier side. This should raise the attention of academic researchers in their study of economic environment that exhibits the hierarchical structure. Additionally, the results also have strong practical implications: We provide buyers with guidance to manage their contracting and procurement costs taking into consideration the existence of hierarchical supply chain distribution structures.

The third Chapter studies R&D offshoring of multinational IT firms and technological learning in host emerging economies. Multinational R&D offshoring to emerging economies has become a prominent phenomenon in recent years. Understanding the response of host country competitors is not only critical if multinational enterprises (MNEs) are to choose the right R&D offshoring strategies but also important if policy makers in MNEs' home countries are to successfully regulate the MNE R&D offshoring to emerging economies. The research question in this essay is: How does R&D offshoring of MNEs affect the R&D investments of host country firms? A two-stage non-cooperative game is developed to analyze the strategic interaction between multinational and host country enterprises engaged in R&D investment. An empirical analysis of 12,309 manufacturing firms in the IT industry in China provides evidence that R&D intensity of host country firms in the IT industry in China is generally positively affected by the R&D intensity of MNEs. However, the magnitude of the impact depends on how easily host country firms are to learn from MNEs. The R&D intensity of a host country firm with relatively low technological and management capability is primarily influenced by the R&D intensity of only those MNEs whose capabilities are lower than those of the host country firms and less influenced by the R&D intensity of more advanced MNEs. For host country firms with relatively high technological and management skills, their R&D decisions are mainly affected by only those MNEs whose capabilities are higher than those of the host country firms and less affected by MNEs with lower technological and management capabilities. The policy implications to host countries of MNE R&D offshoring is that governments of emerging economies should be cautious about the R&D offshoring of MNEs since its positive impact on the R&D investment decisions of a domestic firm would be only conditional.

The fourth Chapter studies software design strategy of a proprietary software vendor (PRV) in markets with open source competitors. There is a growing body of literature investigating the strategic interaction between PRVs and their open source counterparts. Most prior studies focus on software market where the revenue of PRVs comes mostly from software license. Nevertheless, open source software is prominent in software market where the product exhibits evident post-sale service externality (that is, demand for service contracts increases with larger software transactions); consequently, both software license and post-sale service contract (for the purpose of maintenance, technical support, upgrade, etc) constitute some considerable portion of some PRVs' revenue. This Chapter aims to underline the difference between software characterized by high demand for service and software characterized by low demand for service in studying how open source software affects the design strategy of PRV. The analysis of an economic model implies a *magnetic phenomenon* in the design of proprietary software (PRS) with high demand for service that is not observed in the design of software with low demand for service. That is, when the number of features in competing OSS is below a threshold value, the optimal number of features in PRS with high demand for service is higher than the number of features in the competing OSS; the higher the number of features in the OSS, the lower the number of features in the PRS. This situation resembles one in which the position of PRV is *attracted* to that of the OSS. If the position of OSS exceeds the threshold, the design of PRS is placed at the maximal distance away from the position of OSS as if the two *repel* one another. An important implication is that in software markets characterized by high demand for service, when OSS is not very sophisticated PRV could introduce more basic version of their software in response to the growth of the OS counterparts.

CHAPTER 2

PROCUREMENT CONTRACTING STRATEGY IN A HIERARCHICAL SUPPLY DISTRIBUTION NETWORK

2.1 Introduction

Much of the existing research on procurement strategy assumes a flat distribution structure in which each supplier is an independent entity in a pool of suppliers. However, in practice, distribution networks often exhibit a hierarchical structure. The travel lodging industry, for example, has hotel franchise groups at the top of the hierarchy, hotel brands in the middle, and hotel properties at the bottom. There are multiple hotel properties under each brand and multiple brands often participate in the same franchise group based on multi-year contracts. As a buyer, a firm can participate in procurement contracts with and purchase from any or all of the three levels: the franchisor, the brand or the individual hotel property. Similar structures can be found in the insurance market and intermediate goods markets where distribution is through multiple levels of agencies.

The fundamental difference between a flat structure and a hierarchy structure is that not only are entities at different distribution levels in the hierarchy structure heterogeneous, but they also exhibit significant interconnection. In the travel lodging industry example, entities at different distribution levels in the hierarchy structure are heterogeneous in terms of their location demand coverage, in their requirements for discount availability and in contracting cost. These differences, however, are interrelated for suppliers across the levels in the hierarchy. For example, suppliers at higher levels in the hierarchy provide better location demand coverage than those at lower levels and, in return, they ask for larger order volume to make discounts available. Figure 2.1 illustrates the hierarchical industry structure in the travel lodging industry.

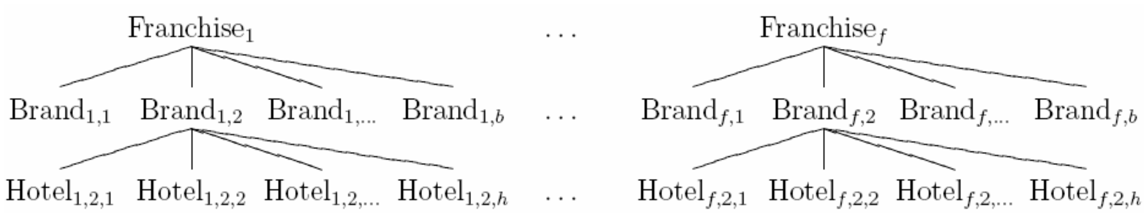

Figure 2.1: Lodging Industry Structure

The challenges that a hierarchical distribution structure bring to researchers are two-fold. First, the interrelationship embedded in the hierarchy creates complexities in procurement decisions for buyers. Second, in situations when a hierarchy structure exists but is considered as flat, the information inherently embedded in the hierarchy is lost and may lead to suboptimal outcomes. While researchers have extensive experience with traditional flat distribution structures, the procurement problem in a hierarchy supplier network remains unexplored.

Studying the procurement problem within a hierarchical supply network has strong practical implication as well. Recent advances in Information Technology (IT) have been instrumental in creating, enabling and supporting these complex modern distribution and procurement structures. From the perspective of a buyer, IT reduces the cost of searching for suppliers at lower level in a hierarchy that has not been considered in procurement previously, and hence provides buyers with a more complicated supply network with more levels of suppliers from which to choose. From the perspective of a seller, IT eases the management of complicated organizational structure hence is an enabler of the hierarchy. Yet, not only does IT support these structures, but it has also enabled the potentially deleterious ability for buyers to undertake these complex procurement processes. Using the travel lodging industry example again, according to Business Travel News (Davis, 2002), the average annual lodging expenditure by the 100 largest spenders in the U.S. on business travel (CT100) is large both in monetary value (an average of $46 million in 2001) and in average annual number of hotel rooms (352,000 in 2001). Domestically, average hotel spending by the CT100 (Cohen, 2004) is also large ($32 million in 2001; $30 million in 2002; $27 million in 2003). After 2003, the travel demand of the CT100 started rising (Cohen, 2004; Meyer, 2005) and keep rising through 2005 and 2006 (Meyer, 2006). Overall, travel represents one of the largest controllable costs for a firm. As such, the temptation to optimize procurement is quite real as IT has enabled compelx relationships, particularly with respect to service operations (Roth and Menor, 2003).

Despite significant research on supply chain optimization methods and techniques using a flat distribution structure, the literature provides little guidance on either the optimization of hierarchical distribution structures or, more importantly, the conditions under which such complex optimization may be worthwhile. Specifically, we explore the importance (quantified as the potential monetary loss to a buyer) of not considering the hierarchical structure when it does exist in the procurement environment. Our research questions are:

How does the existence of a hierarchical structure of supplier distribution network affect the optimal procurement decisions of a buyer?

Under what industry conditions should a buyer undertake a consideration of the full hierarchical procurement structure?

To study those problems, we consider the following procurement scenario. Suppose a firm (buyer) in time period t needs to purchase services (with no inventory necessary) through a period of time T_f. These purchases are based on substantial information gathering efforts at time T_g and subsequent analysis at time T_m. At the beginning of the purchasing horizon (T_0), the buyer establishes contracts with suppliers. Then the buyer purchases service whenever needed through T_f according to the terms in the established contract. The buyer needs to decide at T_0 with which levels in the hierarchy to contract, and among a pool of suppliers at that level, with whom to contract. Figure 2.2 illustrates the timeline of procurement decision making. Our focus is the impact of the hierarchical structure on buyer procurement contracting decision and the condition under which this structure makes a significant difference.

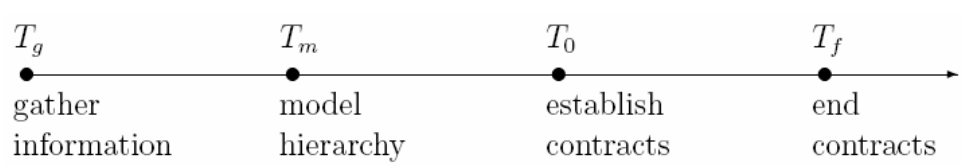

Figure 2.2: Timeline of Procurement Decision Making

We examine buyer procurement contracting decisions in the context of the travel lodging industry. Common corporate practice to reduce lodging cost is to either engage a travel management company and have them negotiate with hoteliers, or negotiate with hoteliers (at one or more levels) directly. The resultant procurement constracts specify corporate purchasing rates. The rates are determined by some discount scheme in negotiation. Examples are quantity discount schemes (e.g., 5% off regular price if the number of rooms purchased exceeds 500); market share discount (e.g., 10% if 60% of lodging demand is provided by one supplier); dollar sale discount (e.g., 5% if dollar sales exceeds $5000). For our analysis, we consider a quantity discount scheme because it is commonly adopted by practitioners (e.g. Serlen, 2002, 2004).

We first model a general hierarchical procurement structure and then link the formulation to prior models. Next, we examine 2,040 instances of the model under a range of parameters to evaluate the industry conditions under which consideration of the full model is justified. Finally, we distill the input conditions into heuristics which guide the usage of such models and discuss further issues requiring investigation. With this research, we contribute to both the nascent understanding of hierarchical procurement and the growing literature on service operations supply chain relationships. Addressing significant gaps noted in Kouvelis et al. (2006, pp. 464-465), the heuristics we offer from the complex models are (a) accessible to supply chain management professionals thereby providing guidance for knowing when to turn to research models; (b) inherently multi-party rather than focused on analytically tractable single party models; (c) service industry focused; and (d) oriented towards future relationships to show ``how the economic, technological, social, and political forces will serve to shape business and [supply chain] practices".

Section 2.2 describes literature related to this problem. In Section 2.3, we formulate a model to represent this problem and describe the relationship of the model representation to prior models. Section 2.4 outlines our framework for evaluation and presents computational results and sensitivity analysis. Finally, Section 2.5 summarizes our results, their implications, and outlines future research needed.

2.2 Related Work

Procurement strategy has been studied extensively in the literature of auction and supply contracting. Elmaghraby (2000) groups relevant literature into a matrix according to the number of rounds the buyer can select a supplier (single or multiple supplier selection periods) and the number of suppliers the buyer procures goods from (sole sourcing or multiple sourcing). Within this framework, procurement strategies are investigated from different angles under various circumstances. Winner selection mechanism (e.g. Deng and Elmaghraby, 2005; Tempelmeier, 2002; Kumar et al., 2006; Chen et al., 2005), sole sourcing versus multi-sourcing (e.g. Riordan and Sappington, 1987; Seshadri et al., 1991), and optimal menu of procurement contracts (e.g. Corbett and de Groote, 2000; Cachon and Zhang, 2006; Wu and Kleindorfer, 2005) have been studied under circumstances such as complete or incomplete information, information asymmetry, single contacts between buyers and sellers or multiple chance of sourcing, etc.

In general, our model falls into the category of supplier selection in a supply chain with multiple sourcing in single selection period. Several approaches have been applied to study this supplier selection problem. Degraeve et al. (2000) classify those models into two groups (single item and multiple items) using the number of items considered in the models. Within each group, they further categorize the models using research method applied in the study: rating, total cost approaches and mathematical programming. Our model fits into the category of one item mathematical programming. Among studies in this category, two studies are most pertinent to our research. Chaudhry et al. (1993) investigate supplier selection with price breaks. They summarize the types of price breaks in practice and discuss the complexity of the problem raised by price breaks. There are noncumulative price breaks and cumulative price breaks. The former refers to the case when an incremental price break applies only to those units purchased in excess of the quantity where a price break occurs. The latter occurs when price break applies to all units once the break satisfies quantity requirement. We adopt the policy of cumulative price breaks as it is commonly observed in practice (Chaudhry et al., 1993). Weber and Current (1993) present a multi-objective approach for the supplier selection problem with the constraints of supplier capacity and minimum order of selected suppliers. Although this research is in the same literature category as ours, it does not study the hierarchy structure in a complex supply network. Furthermore, their study does not incorporates neither the location dimension of demand nor transactional cost of contracting as we do in our model.

Other research methods are also used in the analysis of supplier selection in supply chains. For example, Snir and Hitt (2004) apply a game-theoretical approach and propose a two-stage contract that can help buyers select high-quality vendors in settings where vendor quality is uncertain. Tam and Hui (2001) empirically test the impacts vendor characteristics (product variety, brand name, average price, and network externalities) have in the selection of computer vendors. Overall most of the prior related studies (both analytical and mathematical) assume independent suppliers (Kouvelis et al., 2006) and do not examine the complex supplier network.

However, one stream of research that does consider the complex supplier network is collaborative planning, which refers to ``supply planning and demand fulfillment decision-making among all the suppliers belonging to a company's supply network" (Poundarikapuram and Veeramani, 2004, p. 111). A good example of collaborative planning is the global supply chain of personal computer vendors where each tier of the supply chain represents components suppliers. Most of the studies within this line of research concentrate on developing centralized decision-making models (e.g. Arntzen et al., 1995; Cohen and Lee, 1988). Poundarikapuram and Veeramani (2004) are among the first to study a distribution network with multiple tiers of suppliers using a distributed decision-making algorithm. Although collaborative planning resembles our model, it is fundamentally different since collaborative planning decisions are inherently distributed in nature due to the multiple stages involved in the process, while our model has a centralized aspect as the procurement decision is made in one stage. To be more specific, the suppliers in different tiers do not show the hierarchical structure that appears in our model; rather they are different in terms of the product (service) they are offering. For example, one tier of suppliers consists of raw material suppliers only, a second tier consists of component suppliers only, and a third tier consists of product suppliers only.

The other line of research that considers complex supplier network is multi-attribute procurement, which means that a vendor is selected by both price and other attributes (Parkes and Kalagnanam, 2005). Transportation costs have been incorporated into auctions in a complex supply network (Chen et al., 2005) by modeling the procurement decisions of a single buyer that has quantity requirements for a certain component at a set of geographically diverse locations. However, while this study focuses on auction mechanism design, we study procurement strategy. Also, they model the problem from the perspective of a third party auctioneer, while we focus on the problem from the perspective of the buyer. Furthermore, they do not consider the hierarchical structure of the supply network as we do.

Degraeve et al. (2004) do include a hierarchical structure in a supply network. In their model, there are two layers of suppliers with loose relationship between entities in one layer and those in another layer. Our work is different from theirs in that the goal of their study is to model a two-layer relationship formally, while our main purpose is to discover the significance of a hierarchical supply network in procurement decision making. Consequently we apply a different solution approach from their work. We propose heuristics of flat policy and benchmark their performance with the optimal solution that considers the full hierarchy.

2.3 Model Formulation

Our model assumes a centralized procurement mechanism. We assume a planning horizon of one year. At the beginning of the planning horizon, a buyer estimates its demand across locations and provides the information to potential suppliers at one or more of the three levels (franchise, brand and hotel property level). Suppliers who decide to compete for contract submit bids. For our research, a valid bid contains the average unit price, the number of discount intervals, the lower bound and upper bound of each discount interval, and the discount offered in each discount interval. Based on the bids, the buying firm estimates contracting cost and makes its procurement decision--- the set of suppliers to contract with and the procurement allocation among those suppliers. The objective is to find an optimal strategy to minimize its expected lodging procurement cost. Our model does not consider large meetings, so capacity constraint is not considered.

We consider sets of franchise hotels F, branded hotels B, and hotel properties H. A hotel property/branded hotel belongs to one and only one branded hotel/franchise hotel or is independent. The set of locations a buyer is interested in is denoted as A. The discount scheme of each entity (a franchise, a brand or a property) is modeled with the lower bound (L), upper bound (U) and discount (D) for each discount interval in the scheme. Bounds and discounts are specified in terms of number of rooms the buying firm expects to purchase. The following notation is used in the problem formulation:

A	set of locations for which the buying firm has potential demand
F	set of franchise hotel groups
B	set of brand hotels
H	set of hotel properties
I	set of all entities with which the buyer can contract, i.e., $F \cup B \cup H$
i	set index representing entity levels, $i \in \{F, B, H\}$
I_i	set of entities in level i, $I_i \subseteq I$
$I_{i,a}$	set of level i entities that have properties at location $a \in A$, $I_{i,a} \subseteq I_i$
e	index of individual entity (a franchise, a brand or a hotel property)
M_a	expected total demand at location $a \in A$
K_e	set of discount intervals offered by entities e

$D_{e,k}$ discount offered for the k^{th} interval ($k \in K_e$) by entity e

$U_{e,k}$ upper bound on rooms to purchase from entity e for its k^{th} discount interval ($k \in K_e$)

$L_{e,k}$ lower bound on rooms to purchase from entity e for its k^{th} discount interval ($k \in K_e$)

p_e average unit price for rooms of entity e

N_e contracting cost at entity e

Q a sufficiently large number

The decision variables are as follows:

$$r_e = \begin{cases} 1\text{, if the buyer contracts with entities } e \ \forall e \in I \\ 0\text{, otherwise} \end{cases}$$

$$s_{e,k} = \begin{cases} 1\text{, if the buyer purchases at discount interval } k \text{ offered by entities } e,\ k \in K_e, \forall e \in I \\ 0\text{, otherwise} \end{cases}$$

$u_{e,k}$ = integer, rooms to purchase at entity e at discount interval $k \in K_e \ \forall e \in I$

The objective is to minimize the procurement cost for a buyer at three levels (Equation 1). The summation across set index $i = F$, B and H represents the sum of expected costs at franchise level, brand level and hotel property level respectively. At each level, the expected costs are composed of the lodging costs and the contracting costs. Lodging costs are incurred whenever a buyer actually uses a discount $D_{e,k}$ from entity e to purchase some positive number of hotel rooms that belongs to entity e, i.e., $u_{e,k} > 0$. The contracting costs are incurred whenever a buyer contracts with an entity e, i.e., $r_e = 1$.

$$\text{Minimize } Z = \sum_{i \in \{F,B,H\}} \sum_{e \in I_i} \sum_{k \in K_e} \left[p_e u_{e,k} \left(1 - D_{e,k}\right) \right] + \sum_{i \in \{F,B,H\}} \sum_{e \in I_i} \left[N_e r_e \right] \tag{1}$$

The minimization is bounded by several constraints. First, Equation 2 enforces the demand constraint; i.e., total number of rooms a buyer purchases at location a should satisfy the expected demand at location a. Equation 3 requires that the buyer does not contract with a supplier if total number of rooms to purchase from that supplier is zero. It also enforces the requirement that a buyer cannot purchase using the discount from a supplier if it does not contract with that supplier. Equation 4, 5, and 6 enforce that a buyer can purchase rooms at a certain discount only when the number of rooms a buyer purchases falls within the range required for that discount and that only one discount interval within the discount scheme offered by an individual supplier can be used in the purchase.

$$\sum_{i\in\{F,B,H\}}\sum_{e\in I_{i,a}}\sum_{k\in K_e} u_{e,k} = M_a \quad \forall a \in A \tag{2}$$

$$r_e Q \geq \sum_{k\in K_e} u_{e,k}, \forall e \in I_i, I_i \in I \tag{3}$$

$$u_{e,k} \geq s_{e,k} L_{e,k}, \forall k \in K_e, e \in I_i, I_i \in I \tag{4}$$

$$u_{e,k} \leq s_{e,k} U_{e,k}, \forall k \in K_e, e \in I_i, I_i \in I \tag{5}$$

$$\sum_{k\in K_e} s_{e,k} = r_e, \forall e \in I_i, I_i \in I \tag{6}$$

Equation 7, 8, and 9 impose integrity and non-negativity conditions on the decision variables.

$$r_e \in \{0,1\} \qquad \forall e \in I_i, I_i \in I \tag{7}$$

$$s_{e,k} \in \{0,1\} \qquad \forall k \in K_e, e \in I_i, I_i \in I \tag{8}$$

$$u_{e,k} \geq 0 \qquad \forall k \in K_e, e \in I_i, I_i \in I \tag{9}$$

Our methodology of mathematical programming is well justified. First, it is difficult to solve the problem analytically given the complex hierarchical structure. Like many complex, service management problems, our problem is not conductive to analytical modeling (Roth and Menor, 2003). Second, the various discount schemes and prices offered by hoteliers across multiple locations complicates the problem when the number of hoteliers available becomes large. This is determined by the nature of lodging industry where both hotel brands and hotel properties have relatively strong power in managing room prices and discounts even when they are under the same franchise hotel (Boyd and Bilegan, 2003). For example, as a leading global hospitality group, InterContinental Hotels Group currently has 11 brands with more than 3,500 hotel properties across nearly 100 countries and territories. According to U.S. Census Bureau 2002, there are 60,870 hotel establishments in U.S. alone. The set of hotels a buying firm can contract with becomes even larger when employees travel globally. Individual procurement decision making is straight forward; however, the problem size and interaction of decision variables create a complex problem. Hence, it is not possible for a buyer to find an exactly optimal procurement strategy using a simple heuristic.

As formulated, this problem is related to the concentrator location problems in telecommunications or the facility location problems in logistics. From this perspective, the problem hierarchical supply chain procurement is to find the least cost path from the corporate customer to the hotel property, recognizing that brand and franchise intermediaries exist and can work like network concentrators to aggregate demand efficiently. But, benefiting from these intermediaries requires set up costs. Currently, several heuristics and efficient algorithms are available for the general problem which could potentially applied to our formulation (e.g. Pirkul,1987; Labbe and Yaman, 2006; Aardal et al., 1995; Labbe et al., 2005), particularly in a hierarchical arrangement (Narasimhan and Pirkul, 1992).

However, we do not focus on solution time of heuristics for finding optimal or close to optimal solutions. First, the establishment of hotel contracts is an infrequent activity, often reviewed only on an annual basis. Second, the calculation of the optimal contracting arrangement is done relatively few times during the contract review process.

2.4 Evaluation Framework and Computational Results

2.4.1 A Framework for Contracting Policy Analysis

Recall that our goal is to identify the level in the hierarchy from which a buyer selects suppliers to contract with at the beginning of the planning horizon, normally one year. Since this problem is not a real time decision problem, speed of solution is not the most important concern here. Instead, the key requirement of heuristic is that it must be understandable and executable by the user prior to initiating expensive and time consuming information gathering and model setup. Keeping this requirement in mind, we propose three heuristics of contracting strategy: franchise contracting only, brand contracting only, and hotel contracting only. Intuitively, franchise only strategy refers to the strategy of only contracting with suppliers at franchise level. Similarly, the strategies of brand and hotel contracting only suggest contracting only with suppliers at brand level or suppliers at hotel level respectively. For notational convenience, we also use pure strategies to denote the three heuristics. The pure strategies are benchmarked against using the full hierarchy. Consideration of the full hierarchy will always perform equal or better than a pure strategy since it subsumes every pure strategy. The advantage of focusing on these heuristics is that, first, they are straightforward and easy to evaluate by users. Second, they are consistent with current industry practice of procurement for hotel services (Davis, 2001; Serlen, 2005; Meyer, 2006), hence benchmarking the performance of the three pure strategies with the optimal solution that considers the full hierarchy provides interesting implications regarding whether or not current industry practice is justified.

We propose a two-dimensional framework for the analysis of contracting polices. We capture buyer heterogeneity along one dimension and the uncertainty the buyer faces in the other dimension. In terms of buyer heterogeneity, we use the expected demand of the buyer as the indicator of buying firm size and expected service price at locations to which the buying firm travels as the indicator of location characteristics of corporate travel. For example, high expected demand in general implies large buyers that have high lodging demand across multiple locations. High expected price implies that the travel demand of the buyer is mainly in metropolitan or pricey areas. The three-by-three matrix in Table 2.1 demonstrates the nine types of buyers we consider in our analysis which provides a broad coverage of corporate buyers of lodging service.

Table 2.1: Framework of Buyer Heterogeneity for Contracting Policy Analysis

	High mean price	Medium mean price	Low mean price
High mean demand	Case 1 = (D_h, P_h)	Case 2 = (D_h, P_m)	Case 3 = (D_h, P_l)
Medium mean demand	Case 4 = (D_m, P_h)	Case 5 = (D_m, P_m)	Case 6 = (D_m, P_l)
Low mean demand	Case 7 = (D_l, P_h)	Case 8 = (D_l, P_m)	Case 9 = (D_l, P_l)

In terms of uncertainty, we consider three key uncertain factors buyer may face at the contracting stage: location demand, location service price and contracting cost. We focus on uncertainty in our analysis because contract theory has shown that the main reason for buyers to contract is to prepare for the risk of unforeseen circumstances that may be encountered later. This implies that uncertainty is a critical factor in buyer contracting decision making. Hence, in our study of contracting strategy, we underline the importance of uncertainty in decision making. The three key uncertain factors are selected based on a simple framework proposed by Tullous and Utecht (1992) in their study of decision making regarding multiple or single sourcing. Through literature review and interviews with purchasing agents, Tullous and Utecht (1992) develop three representative uncertainty purchasing situations: need uncertainty, market uncertainty and transaction uncertainty. We map their framework to our problem domain by defining a) need uncertainty to be buyer uncertainty about his demand for lodging service through the year, b) market uncertainty to be buyer uncertainty about the dynamic price of lodging service across locations (Elmaghraby and Keskinocak, 2003), and c) transaction uncertainty to be buyer uncertainty about the potential cost of contracting with each supplier (Serlen, 2005). Each of the uncertain factors we consider has directly implications on practice. For example, at the contracting stage, a buyer always faces the uncertain demand and price. Buyers often face highly uncertain contracting cost when they contract with a supplier for the first time. Overall, our framework allows us to study contracting strategy in a hierarchical supply network under a broad range of scenarios that covers a large pool of buyers.

2.4.2 Computational Evaluation

The objective of computational evaluation is to examine the effectiveness of the three pure strategies by benchmarking the solution of the pure strategies with the optimal solution generated using the full hierarchy. The model is an integer linear program. Its optimal solution and the solution of the heuristics are obtained using ILOG Cplex 10.0.1.0 (Concert Technology) with Microsoft Visual C++ 2005. Random data is generated according to the framework presented in Section 2.4.1. The hierarchical relationship was generated using 3 franchises, 10 brands, and 10,000 hotel properties. This is generated for illustration purposes. A summary of the hierarchical structure of top 50 franchises in the U.S. can be found at:

http://www.ahla.com/products_info_center_top50.asp. The hotel properties were randomly distributed to 1,000 locations. The location can be thought as a region, a state, or a city. The nine types of buyers were generated following industry practice. Buyers with high, medium and low demand are designed as having an expected demand of 300, 150 and 50 rooms in each location respectively (Davis, 2002; Baker et al., 2007). Locations with high, medium and low service prices are designed as having an expected room price of 300, 150 and 50 U.S. dollars respectively (Baker et al., 2007; Anonymous, 2006). This leads to nine types of buyers as discussed in Section 2.4.1.

Three categories of data sets were generated for each type of buyers. In the first category, we vary location demand uncertainty and fix other uncertain variables at their mean value. We generated 10 data points for location demand standard deviation. The range of data points generated is dependent on the mean. This way, we generated a reasonable but wide range of data. In the second and third categories, we vary location service price uncertainty and contract cost uncertainty fixing other uncertain variables at their mean value respectively. Altogether we generated 27 data sets. The goal of computational evaluation is to compare the performance of the heuristics with the optimal solutions in each data set.

Recall that the difficulty of the problem we study here lies in the fact that suppliers in different levels of the hierarchy are interrelated. We incorporate the interrelationship in data generation. Specifically, suppliers at franchise level, brand level and hotel level have different and related expected contracting cost and requirements on the quantity threshold for discount to be available. Our experiences with industry practice show that in general the expected contracting cost with suppliers at higher level of the hierarchy is higher than those at lower level; the quantity discount threshold of suppliers at higher level is in general higher than that at lower levels. Although helpful, this information is still vague. To further reduce the subjective estimates of the parameters, we investigated the brand and hotel property information of the top 50 hotel companies in the directory of Hotel & Lodging Companies managed by the American Hotel & Lodging Association. We found that the average number of brands among the 50 hotel companies is 7, the average number of hotel properties under each brand is 110 and 193 for domestic properties and for all the properties (including both domestic and non-domestic properties), respectively. We did not calculate the contracting cost and the purchase quantity threshold of higher level in the hierarchy by multiplying the contracting cost at lower level with the number of entities in the lower level because contracting cost in general exhibits economy of scale to some extent; e.g., contracting cost with a franchise of 7 brands is in general less than the sum of contracting cost with each of the brands. With this in mind, we estimated the mean contract cost for suppliers at franchise level to be 36,000, for suppliers at brand level to be 9000 and for suppliers at hotel level to be 900.

The purchase quantity thresholds to validate discount at franchise, brand and hotel levels are estimated to be 75, 37 and 10 respectively. To further reduce subjective effects from modeling assumptions on our results, we conduct sensitivity analysis in Section 2.4.4 for mean contract cost and quantity discount threshold requirement.

2.4.3 Findings

The heuristic performance is determined by the difference between the heuristic solution and the optimal solution. The results of data set 1 (based on demand uncertainty) are presented in Figure 2.3.

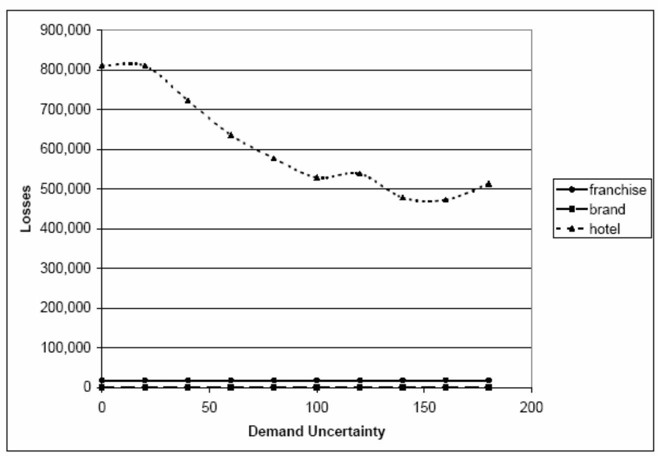

Figure 2.3: Loss of Heuristics with Demand Uncertainty

Although not shown, data set 2 (based on price uncertainty) shows a similar pattern. Figure 2.3 shows that given the discount structure, the brand only pure strategy is the robust optimal strategy among the three heuristics when fixing contracting cost at the mean value and varying the mean and standard deviation of location demand (or room price) one at a time. In both data sets, the full hierarchy is not necessary; a near optimal result can be achieved by contracting at brand level only. In addition, the strategies of contracting only at franchise level dominate the hotel only contracting strategy.

The rationale for brand only strategy to be optimal is that in general a brand only contracting strategy offers the advantage of satisfying demand at a relatively low contracting cost when compared with franchise contracting and hotel contracting. Contracting at franchise level provides best location demand coverage at the cost of potentially high contracting cost; contracting at hotel level provides poor location coverage at relatively low contracting cost. Given the discount structure, the result is robust across large range of standard deviation of location demand and room price, which is also intuitive. Our goal is to compare the three pure strategies with the full hierarchy strategy. Because neither demand nor price directly varies across franchise, brand and hotel level (demand varies across locations and price varies across hotel properties), they will affect the cost savings but not the relative performance of the four strategies.

Since both mean and standard deviation of contracting cost naturally varies across franchise, brand and hotel property, the optimal strategy among the four heuristics will be affected by contracting cost distribution. Our results show that this truly is the case. We summarize our findings in Table 2.2.

Overall, when contract cost is uncertain, pure strategy performance can be as good as using the full hierarchy, but there are cases when the full hierarchy performs slightly better. We only present optimal pure strategies in Table 2.2 because a full hierarchy strategy only performs slightly better in small portion of our solutions. Table

2.2 shows that none of the three pure strategies dominates other pure strategies. This is markedly different from the results in Figure 2.3.

Table 2.2: Optimal Pure Strategies under Contract Cost Uncertainty

Demand	Contract Cost σ	Price		
		High $\mu_P = 300$	Medium $\mu_P = 150$	Low $\mu_P = 50$
High $\mu_D = 300$	Low Medium High	**Case 1** Brand Hotel Hotel	**Case 2** Brand Hotel Hotel	**Case 3** Brand Franchise Hotel
Medium $\mu_D = 150$	Low Medium High	**Case 4** Brand Franchise Hotel	**Case 5** Brand Franchise Hotel	**Case 6** Brand Franchise Hotel
Low $\mu_D = 50$	Low Medium High	**Case 7** Brand Hotel Hotel	**Case 8** Franchise Hotel Hotel	**Case 9** Franchise Hotel Hotel

Taking a closer look at the results, we find that overall the results provide good heuristics for corporate customer contracting strategy taking into consideration the uncertainty of contract cost. Essentially a corporate customer can benefit from contract cost uncertainty to an extent under some conditions. Whether and when a corporate customer should exploit contract cost uncertainty is moderated by relative cost between contracting and room purchasing. The rationale is that the objective function consists of contract cost and room purchasing cost where the former is determined by contract cost standard deviation (when we fix mean contract cost) and the latter by the product of price and demand. We use these findings in Table 2.2 to develop the following propositions.

Proposition 1 When contract cost standard deviation is low, whether buyer should seek risk in exploiting uncertain contract cost is dependent on the relative cost of contract and room.

a. *When room cost is relatively high compared with contract cost, a buyer should focus more on managing room cost than on managing contract cost. It is not worth contracting at the franchise level for some chance of low contract cost. Since contracting at hotel property level does not provide good coverage of room demand, contracting at brand level naturally becomes the best option (cases 1 through 7).*

b. *When room cost is relatively low (cases 8 and 9), buyers should focus more on managing contract cost than on room cost. It is better for a buyer to exploit the uncertain contract cost (risk seeking behavior) by contracting at franchise level. Note that although risk seeking behavior is better in this case, it does not mean the buyer bears much risk as the standard deviation of contract cost is not high. As a result, contracting at franchise level is better.*

Proposition 2 When contract cost standard deviation is sufficiently high, the chance for buyer to run into a large contract cost is also high. Buyer should not risk contracting at franchise or brand level. In this case, regardless of the relative cost between contract and room, the optimal strategy for buyer is to contract only at the level that provides the lowest mean contract cost, i.e., at hotel property level (cases 1 through 9).

Proposition 3 When contract cost standard deviation is in between the situation discussed in propositions 1 and 2, whether the buyer should seek risk in exploiting uncertain contract cost is dependent on the relative cost of contract and room.

a. *When room cost is high compared with contracting cost (cases 1 and 2), the buyer should focus more on managing room cost than on managing contracting cost. It is not worth contracting at the franchise level for some chance of low contract cost. This is the same as the situation in proposition 1(a). However, since the contract cost standard deviation in this case is higher than in proposition 1(a), the risk of incurring high contracting cost by contracting with the brand level is also higher in this case, which pushes the optimal strategy from brand contracting in proposition 1(a) to contracting with hotel property level.*

b. *When room cost is relatively low compared with contract cost (cases 7 through 9), the buyer should not exploit the uncertain contract cost as in proposition 1(b) because the risk of incurring high contracting cost by contracting with the franchise and brand levels is higher in this situation than in proposition 1(b). Moreover, since demand is low in these cases, the need for high demand coverage is not large. As a result, contracting with hotel properties is the optimal strategy.*

c. *When room cost is in between a and b (cases 3 through 6), our computational evaluation shows that the contracting cost reduction benefit of risk seeking exceeds the risk of incurring high contracting cost by contracting at franchise level.*

Although proposition 1 to 3 can be used to guide buyers in choosing the right procurement contracting strategy, for corporate buyers to easily come up with some ideas about the level of suppliers with which to contract, we need to develop some rules-of-thumb. To do so, we relax the requirement of best heuristic as we did in understanding the rationale and only look at the heuristics that produce reasonably good result under all the scenarios considered above. For example, in the low mean demand and medium mean price case, when contracting cost uncertainty is sufficiently low, contracting at franchise level appears to be the optimal heuristic, but contracting at brand level performs quite well as well and the difference in their performance is small. Overall, we find that there does exist some threshold in contract cost uncertainty, below and above which the reasonably good level of suppliers with which to contract changes. Specifically, we find that the first threshold of contract cost uncertainty is when the standard deviation of contract cost is at the mean contract cost and the second threshold is at one and a third of the mean contract cost. For example, suppose the mean contract cost is 900. When the contract cost standard deviation is below 900, contracting at either brand or franchise level perform well. When it is above 900 but below 1200, contracting at either franchise level or hotel level performs well. When it is above 1200, buyer should only contract at hotel level. This is the case in three out of the nine types of buyers in our results. Similar pattern in terms of the threshold of contract cost uncertainty can be found in other types of buyers as well. As is shown in the next section, this pattern also holds in our sensitivity analysis.

To summarize, although the performance of heuristics with respect to contract cost uncertainty appears random at the first look, the rationale analysis provides some intuition about the selection of supplier levels in contracting and the rules-of-thumb provide buyers with information regarding the extent to which their selection still holds.

2.4.4 Sensitivity Analysis

The results in Section 2.4.3 are based on fixed expected contract cost and quantity threshold requirements. To check whether the findings are robust when adjusting either of them, we generated three additional sets of data for the three uncertain factors respectively and further analyze the heuristics. We begin with the analysis by adjusting discount threshold requirement. Then we adjust the mean contract cost in the analysis. Since our intention here is not to derive additional complete sets of solutions, but check the sensitivity of our findings, we do not repeat the analysis as we did in Section 2.4.3; rather we choose one scenario out of the nine types of buyers in our analysis. Without loss of generality, we settle on the scenario of medium mean demand and mean price (the most general case).

2.4.4.1 Sensitivity analysis of Discount Thresholds

Since it is the relative discount threshold suppliers at different levels of the hierarchy offer matters, we adjusted the discount threshold requirement of suppliers at franchise level and hotel level in our analysis and left the discount threshold of suppliers at brand level the same as in Section 2.4.3. Specifically, we analyze four cases: increasing the threshold requirement of suppliers at franchise level to 150, then 225 and increasing the threshold requirement of suppliers at hotel level to 20 then 25. Our analysis shows that the findings regarding the heuristic performance with respect to demand and price uncertainty are robust when we adjust the discount threshold requirement. This implies that if the key uncertainty a buyer faces at the contracting stage is demand or the dynamic service prices at locations to which travel is needed, the buyer can follow a similar contracting strategy in terms of level of suppliers to contract with when suppliers offer discount thresholds that are different from our analysis in Section 2.4.3. Another practical implication of our sensitivity analysis is regarding corporate travel strategy of whether to enforce the use of corporate contract in employee daily lodging reservation. It is a common practice for corporate buyers of lodging service to provide their employees with the flexibility to either reserve using corporate negotiated rates from preferred suppliers (purchase through contract) or find good deals on the Internet (purchase in the spot market)(e.g. Cohen, 2000; McCartney, 2006). Our sensitivity analysis shows that the heuristics perform robustly across the two strategies: enforcing or not enforcing the use of corporate contract in reservation for lodging service. The result is shown in Figure 2.4. The results for price uncertainty are not included; however, they are similar to the results for demand uncertainty.

Then we examined the heuristic performance with respect to contract cost uncertainty when we adjust the discount threshold requirement of suppliers at different levels. Our analysis shows that the findings are robust when the discount threshold of suppliers at hotel level is sufficiently lower than the threshold offer of suppliers at brand level, but sensitive when the discount threshold of suppliers at hotel level moves closer toward that of suppliers at brand level. This implies that if the key uncertainty buyer faces at the contracting stage is contract cost, he can follow a similar contracting strategy in terms of level of suppliers with which to contract when the discount threshold of suppliers at hotel level is sufficiently lower than the discount threshold of suppliers at brand level. Otherwise, he should be cautious in selecting the level of suppliers in the

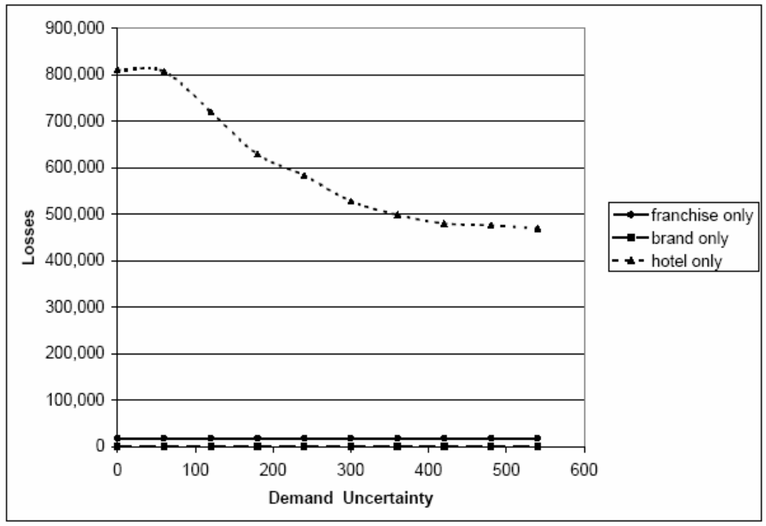

Figure 2.4: Sensitivity of Heuristic Losses to Changes in Discount Thresholds

hierarchy with which to contract when the discount threshold offered by suppliers is different from the ones used in Section 2.4.3. We also find that the heuristic performance is sensitive to discount threshold of suppliers at franchise level. When we increase the discount threshold of suppliers at franchise level, the best heuristic changes and the loss difference between using the new and the old optimal heuristic is large, indicating a potential large loss from not considering the sensitivity of heuristics performance to discount threshold of suppliers at franchise level.

2.4.4.2 Sensitivity Analysis of Mean Contract Cost

Since it is the relative mean contract cost suppliers at different levels of the hierarchy incur that matters, we adjusted the mean contract cost of suppliers at franchise level and hotel level in our analysis and left the mean contract cost of suppliers at brand level the same as in Section 2.4.3. Specifically, we analyze four cases: increasing the mean contract cost to 45,000 and decreasing it to 27,000 for suppliers at franchise level, and increasing it to 1,800 and decreasing it to 450 for suppliers at hotel level. We begin with the analysis of demand and price uncertainty. Overall the performance of the heuristics is robust to adjustment of the mean contract cost of suppliers at different levels in the hierarchy.

Then we examined the heuristic performance with respect to contract cost uncertainty by varying supplier mean contract cost. Our analysis shows that the findings are robust when the mean contract cost of suppliers at franchise level moves toward the mean contract cost of suppliers at brand level, but sensitive when the mean contract cost of suppliers at franchise level is much higher than that of suppliers at brand level. This implies that if the key uncertainty buyer faces at the contracting stage is contract cost, he can follow a similar contracting strategy in terms of level of suppliers with which to contract when the mean contract cost with suppliers at franchise level is sufficiently close to the mean contract cost with suppliers at brand level. Otherwise, he should be careful applying our findings in Section 2.4.3 directly. We also find that the heuristic performance is sensitive to mean contract cost with suppliers at hotel level. When we increase and decrease the mean contract cost of suppliers at hotel level, the best heuristic changes and the loss difference between using the new and the old optimal heuristic is large, implying a potential large loss from not considering the sensitivity of heuristics performance to mean contract cost of suppliers at hotel level.

2.5 Conclusion and Future Work

This paper presents an optimization model that formalizes a procurement contracting problem in hierarchical supply distribution network. The solutions show that considering the hierarchical structure impacts buyer procurement contracting strategy. In general, buyers can achieve near optimal results by contracting with one level of the suppliers within the hierarchy; i.e., pure strategy is near optimal. However, no single pure strategy dominates the other pure strategies. The level of suppliers with which buyers should contract is strongly dependent on buyer type and the major uncertainty buyer faces. Our solutions also show that the loss of not considering the hierarchy and selecting the wrong level of suppliers could be substantial. Therefore, we identify some rules-of-thumb for selecting the right level of contracting suppliers. Our findings strongly indicate the importance of taking into consideration the existence of hierarchical distribution structure at the supplier side. This should raise the attention of academic researchers in their study of economic environment that exhibits the hierarchical structure. Additionally, our results also have strong practical implications: Although simple heuristic rules are easy to be implemented, they can yield significantly suboptimal results. Our results give guidance to practitioners of the potential savings obtainable before undertaking costly and time consuming information gathering processes.

We use publicly available statistics and randomly generated data to obtain our solutions; as such, solutions would likely be somewhat different with real industry data. This will be one direction to explore in future work. Additionally, for tractability, the current model assumes homogeneous hotel rooms. To formulate the model closer to reality, this assumption can be relaxed in future research by allowing for heterogeneous hotel rooms where hotel ranks, reputation and other characteristics are considered.

Further, our paper focuses on lodging procurement problems in a centralized environment. Future work can be done in a decentralized environment where there are multiple sourcing divisions that are responsible for mutually exclusive geographic lodging demand.

CHAPTER 3

R&D OFFSHORING AND TECHNOLOGY LEARNING IN EMERGING ECONOMIES: FIRM-LEVEL EVIDENCE FROM THE IT INDUSTRY[1]

3.1 Introduction

Geographical boundaries have become increasingly blurred with regard to the R&D activities of multinational enterprises (MNEs).[2] Rising R&D costs, the increasing risk and complexity of technological development, and intense competition in domestic and global markets have compelled firms to locate their R&D activities outside the borders of their home countries (Stembridge, 2007). The primary destinations of the R&D offshoring of MNEs appear to be emerging economies[3]. A survey of over 200 MNEs in the United States and Western Europe in 2005 revealed that about 70 percent of the respondents expected an increase in R&D employment in China in the next three years, and slightly more than 40 percent anticipated an expansion in India (Thursby and Thursby, 2006).

---

[1] Zhe Qu acknowledges financial support from the Alan & Mildred Peterson Foundation.
[2] "MNEs" and "foreign firms" are used interchangeably to refer to firms that operate not only in their home countries but also in foreign countries. "Domestic firms," "local firms," and "host country firms" are used interchangeably to refer to firms that operate only in countries where MNEs offshore their R&D.
[3] Emerging economies are defined as "low-income, rapid-growth countries using economic liberalization as their primary engine of growth" (Hoskisson et al. 2000; Arnold and Quelch 1998). China is commonly identified as an emerging economy (Hoskisson et al. 2000).

The escalating R&D expansion of MNEs to emerging economies has drawn considerable attention in their home countries. Policy makers in their home countries are worried that R&D offshoring could lead to a loss of crucial intellectual property to overseas competitors (Hemphill, 2005), which could create an immediate threat to their competitive advantage (Lieberman, 2004; Bardhan and Jaffee, 2005). Moreover, the possible shedding of well-paid R&D jobs and the downward pressure on the wages of engineers and research scientists in the home countries could bear serious consequences because their students would be discouraged from pursuing engineering or science careers, which further exacerbates the problem of losing competitive advantage in the long run[4].

Managers of MNEs are challenged by an emerging "networked innovation system" based on their R&D offshoring actions (Engardio, 2006; Bardhan, 2006). The Boston Consulting Group (2005) argued that decisions such as where to locate a new R&D center, what roles to assign it, and how to integrate it into the existing innovation infrastructure of the MNE would directly impact the outcome of investments and thus demand systematic attention.

Responding to the concerns and questions about R&D offshoring, researchers have investigated factors that motivate or inhibit R&D offshoring (e.g., Thursby and Thursby, 2006; Ambos, 2005; Kuemmerle, 1999; Le Bas and Sierra, 2002) and the conditions in which R&D offshoring could benefit MNEs and their home countries (e.g., Gersbach and Schmutzler, 2006). Thursby and Thursby (2006) found that market growth potential and R&D personnel quality are the top two factors that drive MNEs to offshore R&D to emerging economies, while the quality of intellectual property protection inhibits

---

[4] Despite the potential disadvantages accompanying R&D offshoring, researchers have found that it might improve welfare in home countries (Gersbach and Schmutzler, 2006).

MNE R&D offshoring to these economies. In terms of the impact of MNE R&D offshoring on home countries, Gersbach and Schmutzler (2006) found that it usually increases their welfare[5].

This paper investigates the impact of MNE R&D offshoring on the R&D investment of host country firms in emerging economies. We believe this impact could have a long-term effect on both the MNEs and their home countries. This is because MNE R&D offshoring disturbs the equilibrium of R&D investment competition between foreign firms and domestic firms in host countries and forces the domestic firms to respond. That is, since domestic firms have much to learn from their foreign counterparts, they would strengthen their commitment to R&D so that they could compete with the foreign entrants at a more advanced level. Furthermore, MNE R&D investment in host countries increases the R&D capital stock in these countries, manifested as scientific instruments, advanced machinery, and sophisticated facilities. This capital stock could generate a demand for advanced facilities and machinery in the host countries and stimulate the development of related high-tech industries. MNE R&D investment could also create research-oriented jobs, which would motivate host countries to expand their higher education system, train more scientists and engineers, and nurture the development of high caliber human resources. Overall, the effects of MNE R&D offshoring on host countries could affect the competitive edge of both the MNEs and their home countries in the long run. To succeed in overseas markets, the MNEs and the policy makers of their home countries

[5] There is a large body of endogenous growth literature that investigate the impact of R&D on economic growth (e.g., Romer, 1990; Aghion and Howitt, 1992; Jones, 1995). This line of literature is not thoroughly reviewed because the focus in this paper is the impact of R&D offshoring.

must anticipate the possible reactions of host country firms before they engage in any R&D activity. Otherwise, they could be positioned at a disadvantage in the long run.

Host country governments, particularly those in emerging economies, have endeavored to facilitate technology spillover from foreign direct investment (FDI) (Lieberman, 2004) because they believe successful technology spillover can help domestic firms build indigenous technological capabilities and eventually move up in the value chain. However, evidence has shown that the presence of foreign capital does not guarantee technology spillover. For example, Aitken and Harrison (1999) found that foreign investment negatively affected the productivity of domestically-owned plants in Venezuela because joint ventures gained market share at the expense of domestically-owned firms and forced them to produce less at a higher average cost. We argue that it is difficult for a domestic firm to achieve technology spillover and improve its indigenous innovation capabilities without its own adequate R&D investment, even with an influx of foreign capital. Therefore, it is important that policy makers and business managers in host countries recognize the impact of MNE R&D offshoring on the R&D investment decisions of the domestic firms.

To explain how MNE R&D offshoring shapes R&D investment decisions of domestic firms, we have developed a two-stage non-cooperative game. In the first stage of the game, innovating MNEs and innovating domestic firms engage in cost reducing R&D. In the second stage, all firms participate in Cournot competition in a homogeneous product market. The model implies that MNE R&D offshoring could positively affect the R&D effort of a domestic firm if it is sufficiently easy for the firm to learn from the MNE.

The empirical analysis of 12,309 manufacturing firms in the electronic and telecommunications (IT) industry in China shows that MNE R&D offshoring has a positive effect on the R&D intensity of domestic firms in the IT industry in China. In particular, R&D intensity of domestic firms is mainly influenced by MNEs with higher technological and management capabilities, implying relatively strong learning from MNEs with higher capabilities. When we separate domestic firms into low- and high-capability groups in our analysis, we find that R&D intensity of domestic firms with relatively low technological and management capabilities is primarily influenced by MNEs whose capabilities are lower than theirs and less influenced by MNEs whose capabilities are superior. By contrast, for domestic firms with relatively high capabilities, their R&D intensity is mainly influenced by MNEs whose capabilities are superior to theirs and less influenced by less capable MNEs. We also find that the positive effect of foreign R&D investment diminishes as the geographic distance between the MNEs and the domestic firms increases because it is difficult for domestic firms to learn from MNEs that are located far away from them.

The remainder of this paper is organized as follows. Section 3.2 provides the framework used to explain the rationale behind the R&D investment decisions of domestic firms in response to MNE R&D offshoring. Section 3.3 presents empirical evidence, Section 3.4 presents findings and related policy implications, and Section 3.5 concludes.

3.2 A Theoretical Framework

3.2.1 Literature Review

As reviewed by Crespo and Fontoura (2007), the literature on FDI suggests that positive spillover could occur through five channels: demonstration of foreign enterprises, the mobility of labor, the positive impact of MNE on the export capacity of domestic firms, the competition induced by MNE in the domestic economy and backward and forward linkage with domestic firms[6]. However, a number of studies presented contrasting findings that FDI could negatively affect the productivity of host country firms. For example, using the data of Venezuelan firms, Aitken and Harrison (1999) proposed the market stealing hypothesis to account for negative FDI spillovers. They contested that competition from foreign entrants forces domestic firms to reduce their output, which indirectly increases their average production costs. Consequently, the presence of foreign capital lowers the productivity of the domestic firms. In their study of Moroccan firms, Haddad and Harrison (1993) similarly rejected the hypothesis that the presence of foreign firms accelerated the productivity growth of domestic firms. In spite of the rigor of these analyses using plant level data, they received some criticism. Javorcik (2004) argued that studies by Haddad and Harrison (1993) and Aitken and Harrison (1999) failed to find positive FDI spillovers because they focused merely on intra-industry, or horizontal spillovers. Javorcik (2004) measured vertical (inter-industry) spillover in a study of Lithuanian manufacturing firms and found positive productivity spillovers between foreign affiliates and their local suppliers in upstream sectors.

Another stream of research, which emphasized the impact of FDI on the R&D investment of host country firms, provided a different explanation for the empirical find-

[6] Technology spillover in FDI literature generally refers to the positive effects of FDI on the productivity of host country firms. A detailed review of the FDI literature can also be found in Blomstrom and Kokko (1998, 2001), Saggi (2002) and Marin and Bell (2006).

ing of negative FDI spillover, which we refer to as the "R&D impact" hypothesis. It is known that the R&D investment of domestic firms not only directly contributes to their productivity growth (Potterie and Lichtenberg, 2001; Basant and Fikkert, 1996) but also indirectly influences their growth by facilitating technology spillover from FDI (Kokko, 1994; Alvarez and Molero, 2005). Therefore, the productivity of domestic firms could be positively affected by FDI if the FDI had a positive effect on domestic R&D investment; otherwise, it could be negatively affected. In fact, the linkage revealed by the "R&D impact" hypothesis is one important component of the effect of FDI on the R&D investment of domestic firms, and then on their productivity growth.

The conclusions of most previous theoretical work pertaining to the "R&D impact" hypothesis have been ambiguous. Sanna-Randaccio (2002) studied a two-country, two-firm problem in which each firm decides "the mode of foreign expansion; how much to invest in R&D; and how much to sell in each market." The equilibrium results proved that FDI has an ambiguous effect on domestic R&D. The author argued that the intensity of spillover, the technological characteristics of the firms, and the characteristics of the sector together shape the effect of FDI on domestic R&D.

Similarly, Haller (2004) investigated how the modes of entry of MNEs influenced the R&D decision of host country firms. The author found that the entry of a more efficient foreign firm will lead to lower domestic R&D investment but higher total R&D investment in an industry. However, the author ruled out the possibility of spillover from the MNEs to the domestic firms in the theoretical analysis, which is the central argument in this paper.

Petit and Sanna-Randaccio (2000) investigated the interaction between the mode of foreign expansion and the R&D investment decision of two competing firms in two countries. They found a positive relationship between multinational expansion and R&D investment and that R&D investment increases the likelihood of multinational expansion. Their model and the model we presented in the next section share some similarity in the model setup. Overall the findings of our model are consistent with theirs. That is, the impact of foreign R&D on the R&D level of host country competitors is not strictly increasing or decreasing, but dependent on the R&D spillover coefficients. However, their model is applicable to developed countries and consequently some key parameters in the model (e.g., innovation spillover coefficient) are set to be equal for competing firms. In our model, host country is an emerging economy. Naturally the spillover coefficients are different for host country firms and for multinationals. We discuss in more details the difference between our findings and theirs in the next Section.

Belderbos, Lykogianni, and Veugelers (2005) analyzed the strategic interaction in R&D internationalization decisions by two MNEs, a technology leader and a technology laggard. The authors investigated the role of local inter-firm R&D spillovers and international intra-firm transfer of knowledge on the two firms' R&D location decisions, that is, the share of a firm's R&D resources allocated to that firm's foreign subsidiary. They found that R&D localization of the two firms is strategic substitute. The reason is that increasing R&D at home improves a firm's absorptive capacity, hence allows it to benefit more from the rival's increasing local R&D level. Furthermore, when inter-firm knowledge spillover increases, technology leader allocates a larger share of R&D at home while technology laggard allocates a larger share of R&D abroad. When intra-firm international knowledge transfers are more efficient both technology leader and laggard allocate a larger share of R&D abroad.

Empirical studies on the impact of FDI on the R&D investment of domestic firms by and large supported the theoretical propositions. Veugelers and Houte (1990) argued that MNEs in a host country may stimulate the innovative activities of domestic firms because of the potential knowledge spillovers. However, they also emphasized that competition from the MNEs may limit their production scale and thus reduce domestic R&D expenditures. Their empirical analysis pointed to a significant negative effect of MNEs on the R&D expenditures of domestic firms in Belgium. Using industry level data of the United Kingdom, Driffield (2001) examined how the R&D investments of foreign and domestic firms affect the productivity growth of the domestic firms. The author found that the R&D investment of domestic firms has a significant positive effect on their productivity growth but the R&D investment of foreign firms has an insignificant effect. Moreover, the author found that the R&D investment of foreign firms negatively affects the R&D investment of their domestic counterparts.

Different from the above literature, this study focuses on the impact of MNE R&D effort on host country firms in *emerging economies*, whose technological capabilities are generally inferior to those of firms in developed economies.[7] The intensity of MNE technology spillover and the influence of MNEs on the R&D intensity of domestic firms could differ in emerging economies, compared with those in advanced countries

[7] Although Belderbos, Lykogianni and Veugelers (2005) differentiated technology leader and laggard in their study, they used a different approach to model the two types of firms. In their paper, leader has larger overall R&D resources than the laggard does, while in our case, the two types of firms have different R&D spillover coefficients. We did not follow their modeling approach because we do not possess the data to compare the total R&D resources of the two types of firms.

such as Belgium and the UK. Moreover, economists have generally considered FDI as the flow of homogeneous capital to host countries and studied their overall impact on the R&D investment of domestic firms without differentiating the diverse types of FDI. However, FDI that involves production or service supposedly influences local firms differently from FDI that involves R&D activities. For example, so-called technology-exploiting MNEs might employ mainly local low-skilled workers in their manufacturing plants; in contrast, technology-seeking MNEs may hire primarily highly-educated personnel such as engineers and scientists in their R&D centers[8]. Accordingly, the mechanism and magnitude of technology spillover as well as their impact on the R&D investment of local firms in these two types of MNEs might also differ.[9] We view R&D investment of foreign firms in emerging economies as an advanced form of FDI and disentangle its effect on the R&D effort of domestic firms from the overall impact of FDI.

Because of the important distinctions between this research and prior literature, we did not draw on the existing FDI spillover literature, which largely focuses on FDI in general rather than FDI in R&D in particular. Instead, we build a simple model in the next section to develop a theoretical foundation for our analysis. Our intention is not to develop a theory, but to illustrate the underlying economic rationale within the context of this research.

---

[8] According to Chung and Alcacer (2002), technology exploiting and technology seeking are two major motives of FDI. The former means that an MNE internalizes its unique capabilities and utilizes them in foreign countries; the latter means that an MNE expands abroad in search of capabilities that are not available in its home country.

[9] In surveys of literature characterizing the mechanisms through which spillover takes place, Blomstrom and Kokko (1998) and Saggi (2002) suggest that the labor mobility effect is one of the channels through which spillover may be realized. The other three channels are the demonstration effect, the competition effect, and the R&D stimulation effect.

3.2.2 A Model of Foreign and Domestic R&D Investment

Consider m innovating domestic firms, n non-innovating domestic firms, f innovating MNEs (MNEs that innovate in the host country) and g non-innovating MNEs (MNEs that do not innovate in the host country) that operate in a homogeneous product market in a host country. Following Spencer and Brander (1983), the decision making process of the firms is modeled as a two-stage non-cooperative game (with no collusion). In the first stage, the innovating firms simultaneously engage in cost-reducing R&D activities[10]. In the second stage, all the firms participate in a Cournot competition[11].

The marginal cost of production of both innovating and non-innovating MNEs is $\hat{c}_F$ before the innovating MNEs invest in R&D in the host country. Similarly, the marginal cost of production of both innovating and non-innovating domestic firms is $\hat{c}_D$[12]. In the first stage, the innovating domestic firms and the MNEs simultaneously determine their cost-reducing R&D effort x_i where $i = 1, 2, \ldots, m$ and x_y where $y = 1, 2, \ldots, f$. In addition to the cost reduction resulting from own R&D, an innovating domestic firm can reduce its production costs through absorbing R&D spillover from the innovating MNEs and other innovating domestic firms. As a result, an innovating domestic firm's marginal production cost at the beginning of the second stage is $c_i(x_i, X_D, X_F) = \hat{c}_D - x_i - \alpha\,(X_D - x_i)$ -

[10] We study cost-reducing R&D because many MNEs offshore their business to China in order to reduce production cost, though MNEs could offshore R&D for the purpose of new product development as well. We need to assume a heterogeneous product market in the model if our focus is product innovation.

[11] Cournot competition is a reasonable assumption in our setting because it is fairly difficult to adjust capacity in the IT manufacturing industry. For instance, semi-conductor sector is known to have long lead-time for capacity expansion (Karabuk and Wu, 2003).

[12] In the rest of the paper, we use subscripts F and D to denote the MNE and the domestic firms, respectively.

$\alpha_F X_F$. X_D is the aggregate R&D effort of innovating domestic firms, and $X_D = \sum_{i=1 \text{ to } m} x_i$. X_F is the aggregate R&D effort of innovating MNEs, and $X_F = \sum_{y=1 \text{ to } f} x_y$. $\alpha_F < \alpha < 1$. α represents how easily an innovating domestic firm can absorb R&D spillover from other innovating domestic firms; in other words, it represents how easily an innovating domestic firm can learn from the R&D activities of other innovating domestic firms. Similarly, α_F represents how easily an innovating domestic firm can learn from the R&D activities of innovating MNEs. A non-innovating domestic firm's marginal production cost at the beginning of the second stage is $c_j (X_D, X_F) = \hat{c}_D - \beta X_D - \beta_F X_F$, $j = 1, 2, \ldots, n$ and $\beta < \alpha$, $\beta_F < \alpha_F$, $\beta_F < \beta$. β represents how easily a non-innovating domestic firm can learn from the R&D activities of innovating domestic firms. β_F represents how easily a non-innovating domestic firm can learn from the R&D activities of innovating MNEs. The MNEs can take in local technological expertise as well. The marginal production cost of an innovating MNE at the beginning of the second stage is $c_y (x_y, X_D, X_F) = \hat{c}_F - x_y - \gamma (X_F - x_y) - \gamma_D X_D$[13]; $\gamma_D < \gamma < 1$. γ represents how easily an innovating MNE can absorb R&D spillover from other innovating MNEs. γ_D represents its capability of exploiting local technological know-how. The marginal production cost of a non-innovating MNE at the beginning of the second stage is $c_z (X_D, X_F) = \hat{c}_F - \delta X_F - \delta_D X_D$, z $= 1, 2, \ldots, g$ and $\delta_D < \delta < 1$, $\delta < \gamma$, and $\delta_D < \gamma_D$. δ represents how easily a non-innovating MNE can absorb R&D spillover from innovating MNEs. δ_D represents its capability of exploiting local technological know-how.

[13] We consider the R&D investment of domestic firms as the only source of local technological expertise in order to simplify the analysis, though the MNE can exploit local technological expertise in other ways, such as collaboration with local universities and research institutes (Thursby and Thursby 2006).

Following D'Aspremont and Jacquemin (1988), we assume the costs of R&D to be quadratic and model them as $\theta x_i^2/2$ and $\theta x_y^2/2$ for an innovating domestic firm and an innovating MNE, respectively. We use the same parameter θ for the MNE and the domestic firm because they employ the same pool of resources for their R&D activities (D'Aspremont and Jacquemin, 1988)[14].

In the second stage of the game, all the firms engage in a Cournot competition in which each firm determines its production quantity q conditional on x_i where $i = 1, 2, \ldots, m$ and x_y where $y = 1, 2, \ldots, f$. The inverse demand function is modeled as $p = a - bQ$, where $Q = \sum_{i=1 \text{ to } m} q_i + \sum_{j=1 \text{ to } n} q_j + \sum_{y=1 \text{ to } f} q_y + \sum_{z=1 \text{ to } g} q_z$ and p denote the market clear price. The model is solved backwards so that the subgame perfect equilibria are obtained.

3.2.3 Nash-Cournot Equilibrium

In the second stage, the profit of a firm is

$$(1)\ \pi(q,Q,x) = (a - bQ)q - c(x)q - \theta\frac{x^2}{2}.^{15}$$

The Nash-Cournot equilibrium is computed as $q_i^* = \dfrac{a - bQ^* - c_i(x_i, X_D, X_F)}{b}$,

$q_j^* = \dfrac{a - bQ^* - c_j(X_D, X_F)}{b}$, $q_y^* = \dfrac{a - bQ^* - c_y(x_y, X_D, X_F)}{b}$, and

$q_z^* = \dfrac{a - bQ^* - c_z(X_D, X_F)}{b}$ for an innovating domestic firm, non-innovating domestic

[14] We do not assume that the MNE is more efficient in R&D than the domestic firm, i.e., $\theta_y < \theta_i$, because it is not clear whether this is the case in the IT sector in China. Moreover, incorporating this constraint into the model would not change the main findings. To keep it simple, we choose to stay with $\theta_y = \theta_i = \theta$.

[15] The fixed cost of production is normalized to zero. The subscripts of q, c and x are omitted and represent the corresponding variables for each firm among the four types of firms. As a result, equation (1) represents $m + n + f + g$ equations.

firm, innovating MNE and non-innovating MNE respectively

where $Q^* = \dfrac{a(m+n+f+g) - (\sum_{i=1}^{m} c_i + \sum_{j=1}^{n} c_j + \sum_{y=1}^{f} c_y + \sum_{z=1}^{g} c_z)}{b(m+n+f+g+1)}$.

We assume no exit option in the game, i.e., $q^* > 0$, because we are interested in the relationship between the decision of the MNEs and that of the domestic firms.

3.2.4 Equilibrium of R&D Effort

Going back to the first stage, the objective functions for the innovating domestic firms and MNEs are

(2) $\pi_i(q_i^*, x_i) = b(q_i^*)^2 - \theta \dfrac{x_i^2}{2}$, *i = 1 to m*

(3) $\pi_y(q_y^*, x_y) = b(q_y^*)^2 - \theta \dfrac{x_y^2}{2}$, *y = 1 to f.*

Assuming symmetry in innovating domestic firms, there are unique optimal solutions for x_i when α is sufficiently large[16]. The equilibrium solution for x_i satisfies $\dfrac{\partial \pi_i}{\partial x_i} = 0$, which is equivalent to

(4) $\dfrac{A}{b}(B + EX_D + GX_F) - Hx_i^* = 0$, *i = 1 to m where*

[16] α satisfies $\alpha\ (m-1) > m + n + f + g - \beta n - f\gamma_D - g\delta_D - (m + n + f + g + 1)\sqrt{\dfrac{\theta b}{2}}$

$$A = \frac{2(m+n+f+g-\alpha m+\alpha-n\beta-f\gamma_D-g\delta_D)}{(m+n+f+g+1)}, B = \frac{a+(f+g)\overline{c}_F-(f+g+1)\overline{c}_D}{m+n+f+g+1},$$

$$E = \frac{\alpha(n+f+g+2)-1-n\beta-f\gamma_D-g\delta_D}{m+n+f+g+1},$$

$$G = \frac{\alpha_F(n+f+g+1)-n\beta_F-\gamma f+\gamma+g\delta-1}{m+n+f+g+1} \text{ and } H = \theta-\frac{A}{b}+\frac{\alpha A}{b}.$$

Again, assuming symmetry in innovating MNEs, there are unique optimal solutions for x_y when γ is sufficiently large[17]. The unique optimal x_y satisfies $\frac{\partial \pi_y}{\partial x_y} = 0$, which is equivalent to

(5) $\frac{I}{b}(J+KX_F+LX_D)-Rx_y^* = 0$, *y = 1 to f, where*

$$I = \frac{2(m+n+f+g-m\alpha_F-n\beta_F-\gamma(f-1)-g\delta)}{m+n+f+g+1}, J = (B+\overline{c}_D-\overline{c}_F), K = G-\alpha_F+\gamma,$$

$L = E-\alpha+\gamma_D$, and $R = \theta-\frac{I}{b}+\frac{\gamma I}{b}$.

X_F and X_D can be calculated by solving the following two equations simultaneously:

$\frac{mA}{b}(B+GX_F) = (H-\frac{mA}{b}E)X_D$, and $\frac{fI}{b}(J+LX_D) = (R-\frac{fI}{b}K)X_F$[18].

3.2.5 Interaction between R&D Effort of MNEs and Host Country Firms

We derive the relationship between x_y and x_i by differentiating the left-hand side of equation (4) with respect to x_y:

---

[17] γ satisfies $\gamma\,(f-1) > m+n+f+g-m\alpha_F-n\beta_F-g\delta-(m+n+f+g+1)\sqrt{\frac{\theta b}{2}}$

[18] $X_D = \frac{bmABR-mABfIK+fIJmAG}{b^2RH-RmAEb-fIKHb+fIKmAE}$, $X_F = \frac{fI(J+LX_D)}{Rb-fIK}$.

(6) $\frac{dx_i}{dx_y} = -\frac{\partial^2 \pi_i / \partial x_i \partial x_y}{\partial^2 \pi_i / \partial x_i^2}$, *i =1 to m, y = 1 to f.*

The sign of equation (6) is determined by the nominator, which is positive when $\alpha_F > \frac{n\beta_F + 1 + \gamma f - \gamma + g\delta}{n + f + g + 1}$ and negative when $\alpha_F < \frac{n\beta_F + 1 + \gamma f - \gamma + g\delta}{n + f + g + 1}$. This implies that the R&D efforts of the innovating MNEs have a positive effect on the R&D efforts of the innovating domestic firms if it is sufficiently easy for the domestic firm to learn from the MNEs. However, the R&D effort of the MNEs has a negative effect if it is difficult for the domestic firms to learn from the MNEs[19]. This is consistent with the findings in Petit and Sanna-Randaccio (2000). Furthermore, the threshold of spillover coefficient is dependent on the ease for non-innovating domestic firms to learn from foreign R&D investment and the ease for foreign firms to learn from innovating MNEs, suggesting the importance of considering the difference in spillover coefficients among firms. This is different from the findings in Petit and Sanna-Randaccio (2000)[20].

---

[19] This is different from the findings of Belderbos, Lykogianni and Veugelers (2005). Belderbos, Lykogianni and Veugelers (2005) found that host country firms lower their R&D level if foreign competitors increase their R&D level in the host country. The difference is due to the model setup. In our model, the competition effect may offset the spillover effect, leading to a threshold in the spillover coefficient of host country firms, below which the competition effect dominates, and above which the spillover effect dominates. Belderbos, Lykogianni and Veugelers (2005) also considered product market competition, but competition in their model refers to global competition, which is the overall competition in the two countries. While in our case, we study a one-country problem where host country firms do not expand abroad. As a result, the increased competition due to the entry of foreign competitors hurts the market share of host country firms, which lowers their incentive to innovate.

[20] The spillover coefficient is the same for competing firms in Petit and Sanna-Randaccio (2000).

We examine the rationale behind these results by inspecting how the R&D efforts of the MNE affect the output of the innovating domestic firms[21].

$$(7)\ \frac{\partial q_i}{\partial x_y} = \frac{\partial q_i}{\partial c_i}\frac{\partial c_i}{\partial x_y} + \sum_{l \neq i}\frac{\partial q_i}{\partial c_l}\frac{\partial c_l}{\partial x_y}.$$

The first multiplicative expression in equation (7) is always positive and represents a spillover effect: Because an innovating domestic firm absorbs the R&D spillover from the MNEs, it is able to reduce its cost and raise its output. The second expression is always negative and represents a competition effect: Although an increase in the R&D efforts of a MNE cuts its own production cost, it indirectly raises the costs of the domestic firms, leading to a decrease in its output. The two processes compete with each other, and the observed result is their net effect.

To summarize, when it is sufficiently easy for the domestic firms to learn from the MNEs, that is, when absorbing R&D spillover from the MNEs is easy, the spillover effect dominates the competition effect, leading to a net positive effect of the MNE R&D offshoring on the R&D effort of the domestic firms. When it is difficult for the domestic firms to learn from the MNEs, the competition effect dominates the spillover effect, leading to a net negative effect of the MNE R&D offshoring. In general, how easily the domestic firms learn from the MNEs determines the net impact of the two competing factors.

3.3 Empirical Evidence from the IT Sector

3.3.1 Data

[21] Because $\frac{\partial q_i^*}{\partial x_i}$ is always positive, the sign of $\frac{dx_i}{dx_y}$ is determined by $\frac{\partial q_i}{\partial x_y}$.

The dataset used in this study was constructed by the National Bureau of Statistics of China on a yearly basis from 2001 to 2005. The cleaned dataset includes 4,506 firms in the 2001 data, 4,834 in the 2002 data, 5,565 in the 2003 data, and 8,525 in the 2005 data[22], totaling 12,309 manufacturing firms in the two-digit electronic and telecommunication (IT) sector. Each firm is assigned an invariant code in the dataset so that we can match the observations of each firm across the four-year observation period. The dataset contains more than 50 statistical indicators, including firm input, output, R&D expenditures, capital composition, employment, geographical location, the sector in which a firm operates (at four-digit sector level), ownership status, and assets and liabilities.

The IT industry in China is an excellent example for our empirical research. First, China's IT sector has developed rapidly in recent years. The main driver of the development of the Chinese IT sector has been the rapid development of domestic firms and the inflow of foreign investment to this industry. China's trade volume of IT products increased from 35 billion US Dollars in 1996 to 180 billion US Dollars in 2004, with an average annual growth rate of 38 percent (OECD, 2005). China is now the sixth strongest global IT market. Meng and Li (2002) documented that 47 percent of electronic products manufactured at the end of the 1990's in China were produced by foreign firms or joint ventures. Having transferred their manufacturing branches to China to reduce production costs, a majority of MNEs in the IT sector imported the critical components normally developed in advanced countries, assembled the products in China, and exported the final

[22]Data is cleaned by deleting the observations with negative industrial output value, employee number, sales volume, industrial added value, and R&D investment value. The data in 2004 are not included in this analysis because of the lack of R&D investment value.

products to overseas market.[23] Because of this so-called "processing trade", China became the biggest exporter of IT goods in the world in 2004 (OECD, 2005; Katsuno, 2005).

Another reason why the IT industry is an excellent example for this research is the considerable competition and interaction between foreign firms and local competitors in the Chinese IT industry. In spite of their relatively limited size and technological know-how, Chinese domestic IT firms are rapidly developing their production, export, and R&D capabilities. After all, they compete effectively with MNEs in various product markets within China, and several of them, active in overseas markets, have even emerged as global players[24].

3.3.2 Econometric Framework

To test how the R&D effort of MNEs affects local firms, we adopt the following general econometric framework in the empirical analysis.

(8) $(R\&D\ intensity\ of\ domestic\ firm)_{jit} = f\{(R\&D\ intensity\ of\ MNEs)_{jt},\ (control\ variables)_{jit}\}$,

where *i*, *j*, *t* represent a firm, a four-digit level industry, and time, respectively. The data for the variable of the R&D intensity of a domestic firm and the control variables are at the firm level, but those for R&D intensity of MNEs are at a four-digit industry level.

[23] Research by Lemoine and Unal-Kesenci (2004), China's National Bureau of Statistics (2005), and Fung (2005) attributed the recent expansion of China's exports in machinery, electrical equipment, and electronic products, in large part to processing trade and the global division of labor, especially in East Asia.
[24] Examples include Lenovo, which acquired the IBM personal computer business in 2004 and became the world's third largest personal computer producer, TCL, which acquired the television business of Thomson in France in late 2003, and Huawei and ZTE, which are active players in the worldwide telecommunications equipment market.

In this framework, we use the R&D intensity of a domestic firm, i.e., the ratio of R&D expenditures to sales value, as the dependent variable. The average R&D intensity of foreign firms at the four-digit industry sector level enters the right-hand side of the function as a key independent variable[25].

Since the R&D intensity of a firm is influenced by not only its foreign competitors but also many other factors, we include these factors as control variables in the regression. We discuss the control variables in detail in Section 3.3.2.2.

The theoretical model in Section 3.2 implies that how the R&D effort of MNEs in a host country impacts the R&D decision of domestic firms depends on the ease of learning in the domestic firms, i.e., how difficult it is for spillover to take place. The difficulty of testing the propositions lies in measuring the ease of learning in the domestic firms and constructing the key independent variables. We discuss the measurement issue and our approach of constructing the key independent variables in Section 3.3.2.1.

3.3.2.1 Measurement of Ease of Learning and Construction of MNE R&D Intensity

Kaiser (2002) reviewed several different methodologies of constructing knowledge spillover variables. These include measuring the technological distance between firms by the patent activities (Jaffe, 1986) and share of scientists in a firm (Adams, 1990), and meas-

[25] R&D intensity has been commonly used in the literature to measure R&D effort of a firm (e.g., Veugelers and Vanden Houte (1990), and Cohen and Klepper (1992)). In our model presented in Section 3.2, we use x to denote R&D effort and we investigate the interaction between the R&D efforts of domestic firms and MNEs. Correspondingly, R&D intensity is equivalent to $\theta x^2/2q$ and the behavior of R&D intensity of domestic firms in response to R&D intensity of MNEs is determined by the sign of $d(x_i^2/q_i^*)/d(x_y^2/q_y^*)$. We obtain similar implications analyzing $d(x_i^2/q_i^*)/d(x_y^2/q_y^*)$. That is, there is a threshold value in α_F, above which the spillover effect dominates the competition effect, while below which the latter dominates the former. We present our analysis in Appendix A.

uring the Euclidean distance of firm characteristics (Inkmann and Pohlmeier, 1995). Ornaghi (2006) measured the ease of spillover by grouping firms according to their size, R&D expenditures, and geographical location.

In this paper, we measure ease of learning of a domestic firm using both the technological and geographical distance between the subsidiary company of an MNE in China and local firms. We use labor productivity to construct technological distance since we do not have firm-level data on patent applications and employed scientists in our dataset. Moreover, although we possess data about size and R&D expenditures of the subsidiaries of MNEs, we hesitate to use them as indicators of the technology level of the foreign firms because this information represents only the characteristics of their subsidiaries in China, rather than their overall technology capability. Labor productivity, defined as added value divided by number of employees, represents the technology and management skills of a manufacturing firm in general.

To test the effect of technological distance on the relationship between the R&D intensity of MNEs and domestic firms, we construct three variables—R&D intensity of same-group MNEs, R&D intensity of higher-level-group MNEs, and R&D intensity of lower-level-group MNEs — and include them as key independent variables in our analysis. We follow Ornaghi (2006) to divide the dataset into five subsets (groups) according to the average labor productivity of all firms during the observation period. Each of the groups contains one-fifth of the firms in the dataset[26]. Suppose a domestic firm is classi-

[26] In this study, a firm classified in a relatively high-labor productivity group would be equipped with more advanced machinery and equipment or have established a more efficient organizational and management structure.

fied in the group with the lowest labor productivity[27]. Then the average R&D intensity of the foreign firms in the same lowest labor productivity group and in the same four-digit sector of the domestic firm is taken as the value of the variable R&D intensity of same-group MNEs for this domestic firm. Similarly, we construct this variable for the domestic firms in the other four groups. The variable R&D intensity of higher-level-group MNEs is constructed by taking the average R&D intensity of the foreign firms in a group with higher labor productivity, compared with the group of the domestic firm, and are also in the same four-digit sector of the domestic firm. For example, suppose a domestic firm is in Group 1. For this domestic firm, the variable R&D intensity of higher-level-group MNEs is calculated by taking the average R&D intensity of the foreign firms in Group 2 and in the same four-digit sector of the domestic firm. For a domestic firm in Group 2, the average R&D intensity of the foreign firms in Group 3 and in the same four-digit sector of the domestic firm is taken as the value of the variable. Different from Ornaghi (2006), we take 1 as the value of the variable R&D intensity of higher-level- group MNEs for a domestic firm in Group 5, i.e., the group with the highest labor productivity, because 1 represents the upper bound of the R&D intensity of a firm. This numerical manipulation is justified by the analysis of the data presented in Table 3.1. Table 3.1 shows that foreign firms in the groups of higher labor productivity generally have higher R&D intensity. Similarly, we construct the variable R&D intensity of lower-level-group MNEs by entering the average R&D intensity of the foreign firms in the group directly below the group of the domestic firm. For example, for a domestic firm in Group 2, the average R&D intensity of the foreign firms in Group 1 and in the same four-digit sector of the

[27] The labor productivity of Groups 1, 2, 3, 4, 5 is in an ascending order, i.e., Group 1 has the lowest labor productivity, and Group 5 has the highest labor productivity.

domestic firm is taken as the value of the variable R&D intensity of lower-level-group MNEs. For a domestic firm in Group 1, i.e., the group of lowest labor productivity, 0 is taken as the value of the variable because foreign firms in the group of lower productivity tend to have lower R&D intensity (Table 3.1). The methodology of constructing the variables of R&D intensity of same-group MNEs, R&D intensity of higher-level-group MNEs, and R&D intensity of lower–level-group MNEs is summarized in Table 3.2.

Incorporating the three variables into our general econometric framework, we further identify our econometric function as:

(9) *(R&D intensity of domestic firm)* $_{jit}$ = *f {(R&D intensity of same-group MNEs)* $_{jt}$, *(R&D intensity of higher-level- group MNEs)* $_{jt}$, *(R&D intensity of lower-level-group MNEs)* $_{jt}$, *(control variables)* $_{jit}$*}*,

where *i*, *j*, *t* represent a firm, a four-digit level industry, and time, respectively.

By constructing the variables and estimating the coefficients of R&D intensity of same-group MNEs, R&D intensity of higher-level-group MNEs, and R&D intensity of lower-level-group MNEs, we can identify how foreign firms with different levels of technology and management capabilities affect the R&D effort of domestic firms. In other words, we can test how the technological distance between domestic and foreign firms determines the impact of foreign R&D intensity on domestic R&D decisions.

Table 3.1: Ranking of Average R&D Intensity of MNEs among Different Groups

	R&D Intensity					Ranking of R&D Intensity Among 5 Groups				
Four-digit Sector	Group 1 (Lowest labor productivity)	Group 2	Group 3	Group 4	Group 5 (Highest labor productivity)	Group 1 (Lowest labor productivity)	Group 2	Group 3	Group 4	Group 5 (Highest labor productivity)
4111	.0120	.0065	.0156	.0088	.0256	3	5	2	4	1
4112	.0104	.0019	.0019	.0135	.0203	3	5	4	2	1
4113	.0022	.0082	.0023	.0049	.0046	5	1	4	2	3
4119	.0041	.0018	.0016	.0060	.0062	3	4	5	2	1
4130	.0026	.0029	.0062	.0041	.0074	5	4	2	3	1
4141	.0024	.0000	.0002	.0020	.0009	1	5	4	2	3
4143	.0024	.0026	.0026	.0021	.0026	4	2	3	5	1
4151	.0010	.0032	.0022	.0009	.0010	3	1	2	5	4
4153	.0054	.0018	.0077	.0093	.0027	3	5	2	1	4
4155	.0055	.0048	.0003	.0048	.0081	2	3	5	4	1
4160	.0009	.0027	.0017	.0022	.0019	5	1	4	2	3
4171	.0015	.0008	.0015	.0023	.0091	3	5	4	2	1
4172	.0002	.0008	.0016	.0017	.0047	5	4	3	2	1
4182	.0002	.0000	.0010	.0113	.0368	4	5	3	2	1
4189	.0043	.0084	.0010	.0040	.0062	3	1	5	4	2
Sum of Ranks						52	51	52	42	28

Note:
1. The labor productivity of groups 1, 2, 3, 4, 5 is in an ascending order, i.e., Group 1 has the lowest labor productivity and Group 5 has the highest labor productivity.

Table 3.2: Methodology of Constructing R&D Intensity of Lower-Level-Group MNEs, R&D Intensity of Same-Group MNEs, and R&D Intensity of Higher-Level-Group MNEs

A domestic firm in different groups	Variables		
	R&D intensity of lower-level-group MNEs	R&D Intensity of same-group MNEs	R&D intensity of higher-level-group MNEs
A domestic firm in Group 1 (lowest labor productivity)	0	A	B
A domestic firm in Group 2	A	B	C
A domestic firm in Group 3	B	C	D
A domestic firm in Group 4	C	D	E
A domestic firm in Group 5 (highest labor productivity)	D	E	1

Note:
1. The labor productivity of groups 1, 2, 3, 4, 5 is in ascending order, i.e., Group 1 has the lowest labor productivity, and Group 5 has the highest labor productivity.
2. A: Average R&D investment intensity of the MNEs in Group 1, which are also in the same four-digit sector of the domestic firm
B: Average R&D investment intensity of the MNEs in Group 2, which are also in the same four-digit sector of the domestic firm
C: Average R&D investment intensity of the MNEs in Group 3, which are also in the same four-digit sector of the domestic firm
D: Average R&D investment intensity of the MNEs in Group 4, which are also in the same four-digit sector of the domestic firm
E: Average R&D investment intensity of the MNEs in Group 5, which are also in the same four-digit sector of the domestic firm

We are also interested in how this impact differs between domestic firms with high and low technological and management capabilities. Therefore, we analyze Equation 9 by running the regression on four sub-samples of domestic firms separately. The first sub-sample includes domestic firms of Groups 1 (the lowest labor productivity group), 2, 3, and 4, and excludes domestic firms of Group 5 (the highest labor productivity). The second sub-sample includes domestic firms of Groups 1, 2, and 3, and excludes domestic firms of Groups 4 and 5. Regressing on the two sub-samples separately allows us to focus on domestic firms with relatively low technology and management capabilities and to examine the effect of technology distance on the relationship between the R&D intensities of these domestic firms and those of MNEs. By the same token, we run regressions using the other two sub-samples that focus on domestic firms with relatively high technology and management capabilities. The third sub-sample includes domestic firms of Groups 2, 3, 4, and 5, and excludes the domestic firms of Group 1; the fourth sub-sample includes the domestic firms of Groups 3, 4, and 5 and excludes the domestic firms of Groups 1 and 2. The regression functions that differentiate the two types of domestic firms are presented in the following:

(10) *(R&D intensity of low-capability domestic firm)* $_{jit}$ *= f {(R&D intensity of same-group MNEs)* $_{jt}$*, (R&D intensity of higher-level-group MNEs)* $_{jt}$*, (R&D intensity of lower-level-group MNEs)* $_{jt}$*, (control variables)* $_{jit}$*}*,

and

(11) *(R&D intensity of high-capability domestic firm)* $_{jit}$ *= f {(R&D intensity of same-group MNEs)* $_{jt}$*, (R&D intensity of higher-level-group MNEs)* $_{jt}$*, (R&D intensity of lower-level-group MNEs)* $_{jt}$*, (control variables)* $_{jit}$*},*

where *i*, *j*, *t* represent a firm, a four-digit level industry, and time, respectively.

Because different industry sectors differ in their production function and labor productivity of firms in some industry sectors are intrinsically higher than in other sectors, grouping firms using labor productivity without considering this natural difference between industry sectors may generate noise in the analysis. For example, a firm with relatively high labor productivity in its industry sector may be placed in the group of low-capability because the sector it belongs to has low labor productivity and, a firm with relatively low labor productivity may be placed in the group of high-capability because it belongs to a sector with high labor productivity. Following the survey of current business by Bureau of Economic Analysis on research and development satellite account[28], we separate technology-intensive industry sectors from labor-intensive sectors and estimate the coefficients for the two types of industry sectors separately. We compare the estimation with the full sample results in Section 3.3.4.

To examine the robustness of the results obtained by classifying the firms into five groups, we also divide the firms into three groups, each of which contains one-third of all the firms, and then estimate the coefficients. We compare the estimation of the five- and three-group approaches in Section 3.3.4.

Another type of "distance" between domestic and foreign firms is geographical distance. We make use of the geographical code in the dataset to identify the geographical location of each firm. We construct three additional variables: R&D Intensity of MNEs within a City, R&D Intensity of MNEs within a Province, and R&D Intensity of MNEs within Neighboring Provinces. For a domestic firm, R&D Intensity of MNEs within a City is the average R&D intensity of foreign firms that are in the same four-digit sector and in the same city of the domestic firm. R&D Intensity of

[28] The complete survey can be downloaded at www.bea.gov.

MNEs within a Province is the average R&D intensity of foreign firms that are in the same four-digit sector, in the same province, but in a different city of the domestic firm. By the same token, R&D Intensity of MNEs within Neighboring Provinces is the average R&D intensity of foreign firms that are in the same four-digit sector, within the area of the same province plus neighboring provinces[29], but in a different city of the domestic firm. Foreign firms encompassed by the variable R&D Intensity of MNEs within Neighboring Provinces are spread throughout a wider area than those included in the calculation of variable R&D Intensity of MNEs within a Province. The geographical areas covered by these two variables overlap, but they are mutually exclusive to the area covered by the variable R&D Intensity of MNEs within a City. As a result, the variables R&D Intensity of MNEs within a Province and R&D Intensity of MNEs within Neighboring Provinces enter the regression separately. The regression functions are presented in the following:

(12) (R&D intensity of domestic firm)$_{jit}$ *= f {(R&D Intensity of MNEs within a City)*$_{jt}$*, (R&D Intensity of MNEs within a Province)*$_{jt}$*, (control variables)*$_{jit}$*},*

and

(13) (R&D intensity of domestic firm)$_{jit}$ *= f {(R&D Intensity of MNEs within a City)*$_{jt}$*, (R&D Intensity of MNEs within Neighboring Provinces)*$_{jt}$*, (control variables)*$_{jit}$*},*

where i, j, t represent a firm, a four-digit level industry and time, respectively.

Our dataset shows that among all the provinces, Guangdong province hosts the largest number of foreign and domestic IT firms. Shanghai and Jiangsu follow to host the second and third largest number of IT firms. However, the number of firms located in Guangdong is higher than that in Shanghai and Jiangsu combined. Given the importance of Guangdong province to the Chinese IT industry, we single out the data

[29] The list of neighboring provinces for each province in our data is presented in Table B.1 in Appendix B.

of the firms located in Guangdong and compare the results based on these data to the results obtained from the full sample.

3.3.2.2 Control Variables

Since more financial resources allow a firm greater opportunity to experiment and less stringent requirements for performance, they contribute to risky investments in R&D. That is, the more financial resources a firm has, the more likely it will invest in R&D or increase R&D investment. Following studies on slack resources by Greve (2003), Daniel et al. (2004) and Tan and Peng (2004), we include the ratio of administrative, financial, and selling expenses to sales value in the function to measure the financial resources of a firm.

Schumpeter (1950) was among the first to hypothesize that large firms in a mature capitalist economy generate a disproportionately large share of society's technological advances. Scholars who support this hypothesis have articulated that larger firms possess larger-scale, internally-generated funds, so they secure more resources with which to conduct risky R&D projects. Scale economies of R&D activity and return to R&D investment given a larger volume of sales also contribute to the advantage of larger firms. However, Cohen et al. (1987) argued that these points were flawed because of inadequate attention to the unit of analysis and to industry effects. They found that overall firm size has a very small, statistically insignificant effect on business unit R&D intensity. In a recent study, Lee and Sung (2005) contested that firm size does not directly affect R&D intensity, but it does exert influence by affecting firm-specific technological competence. Although no consensus was reached on the relationship between firm size and R&D intensity in previous literature, we include the number of employees of a firm divided by the average number of the employees of all firms in the database as a control variable in our analysis.

Before conducting R&D activities, a firm must first invest in sophisticated technological equipment (Del Canto and Gonzalez, 1999). Investment in R&D equipment, similar to investment in ordinary machinery and production, is typically classified as fixed asset investment. A firm investing more in fixed assets, i.e. having higher capital intensity, is more likely to expand its R&D investment. Therefore, we include the ratio of fixed asset investment to added value as an explanatory variable to control for the effect of capital intensity on R&D intensity.

Committed to fixed asset investment in R&D equipment, a firm also has to invest in human capital to build an efficient R&D team consisting of scientists and technicians with the proper qualifications, knowledge, and skills. Their experience and knowledge constitute the human resources of the firm, which contribute to the success of R&D projects. Following Del Canto and Gonzalez (1999), we use compensation per employee of the firm, which is the total compensation divided by the number of employees, as a proxy variable for the human resources of the firm.

The relationship between industry R&D intensity and market structure has been extensively investigated in the literature of industrial organization. According to Lee (2005), a number of studies found that the relationship between industry R&D intensity and market concentration can be described as an inverted-U shape. This indicates that moderately concentrated industries invest more in R&D activity than either highly competitive or highly concentrated industries. Bearing this in mind, we include in the regression a variable of the Herfindahl-Hirschman Industrial Concentration Index (HHI), calculated at the four-digit sector level, to account for the potential impact of market structure on a firm's R&D intensity.

We also include several ownership, industry and year dummy variables in the functions to control for ownership, industry and time effect, respectively. The definitions of all the dependent and independent variables are summarized in Table 3.3, and the descriptive statistics of the variables are displayed in Table 3.4.

3.3.3 Estimation Methodology

3.3.3.1 Censored Dependent Variable

More than half of the domestic firms in our dataset did not conduct R&D during the observation period, so the R&D intensity of these firms was zero. This type of dependent variable is known as a "censored dependent variable," i.e., the values of the variable in a certain range are all reported as a single value, e.g., zero. Conventional linear regression method is not able to distinguish between non-linear "zero" observations and continuous observations. Thus, we use the Tobit model as our baseline model. Moreover, although some time-invariant characteristics of domestic firms could also affect the dependent variable, their impact is not captured by the baseline Tobit model. Therefore, we run a panel Tobit model to control for these unobserved firm-specific factors.

3.3.3.2 Endogenous Independent Variables

The theoretical model presented in Section 3.2 clearly shows that the R&D intensity of foreign firms is endogenous in structural functions because of strategic interaction between a foreign firm and a domestic firm. To correct the bias caused by endogeneity of the explanatory variable, we instrument the R&D intensity of foreign firms with two variables: new-product intensity, defined as new-product value divided

Table 3.3: Variables

Variable Name	Definition and Note
Domestic Firm R&D Intensity (Dependent Variable)	R&D expenditure of a domestic firm / Sales value of a domestic firm
Domestic Firm R&D Intensity at Four-digit Sector Level (Independent Variable)	Sum of R&D expenditure of domestic firms in the same four-digit sector / Sum of sales value of domestic firms in the same four-digit sector
R&D intensity of same-group MNEs[1]	Sum of R&D expenditure of MNEs in the same four-digit sector / Sum of sales value of MNEs in the same four-digit Sector (The MNEs are classified into the same labor productivity group as the group of domestic firms)
R&D intensity of higher-level-group MNEs [1]	Sum of R&D expenditure of MNEs in the same four-digit sector / Sum of sales value of MNEs in the same four-digit Sector (The MNEs are classified into a higher labor productivity group compared with the group of domestic firms)
R&D intensity of lower-level-group MNEs [1]	Sum of R&D expenditure of MNEs in the same four-digit sector / Sum of sales value of MNEs in the same four-digit Sector (The MNEs are classified into a lower labor productivity group compared with the group of domestic firms)
Financial Resource	Administrative, financial and selling expenses / Sales
Firm Size	Number of employees / Average number of employees of firms in the sample
Fixed Asset Investment Intensity	Fixed asset investment /Added value
Human Resource	Employee compensation value / Number of employees
Herfindahl-Hirschman Industrial Concentration Index (HHI)	$\sum_{j=1}^{n}$ (Market share (Percentage) of j firm in a four-digit sector)^2
Ownership Dummy Variables	For domestic firms, the base group is "other domestic firms". The dummy variables represent state-owned, collective, employee shareholding cooperatives, joint operation enterprises, limited liability companies, stock companies, private enterprises.
Year Dummy Variables	The base group is 2001. The dummy variables represent 2002, 2003 and 2005.
Industry Dummy Variables	The base group is Audio and Video Equipment. The dummy variables represent Computer and peripheral equipment, Communications Equipment, Semiconductor and related device, Radio and Television Broadcasting Equipment, Electronic and Communications Equipment Repairing, and Other Electronic Component
MNE R&D Intensity within a City	Average R&D intensity of MNEs that are in the same four-digit sector and within the same city or county of domestic firms.
MNE R&D Intensity within a Province	Average R&D investment intensity of MNEs that are in the same four-digit sector, in the same province but in a different city or county of domestic firms.
MNE R&D Intensity within Neighboring Provinces	Average R&D intensity of MNEs that are in the same four-digit sector, within the area of the same province plus neighboring provinces but in a different city or county of domestic firms.
MNE new-product intensity	Sum of new product value of MNEs in the same four-digit sector / Sum of output value of MNEs in the same four-digit Sector
Ratio of the administrative, financial and selling expenses to sales value of MNE	Sum of administrative, financial, and selling expenses of MNEs in the same four-digit sector / Sum of sales value of MNEs in the same four-digit Sector

Note: 1. a detailed definition of the variable is presented in Section 3.3.2.1.

Table 3.4: Summary Statistics of Variables

Year	Number of Domestic Firms		Domestic Firm R&D Intensity (Dependent Variable)	Domestic Firm R&D Intensity at Four-digit Sector Level (Independent Variable)	R&D Intensity of Same-group MNEs	R&D Intensity of Higher-level Group MNEs	R&D Intensity of Lower-level-Group MNEs	Financial Resource	Firm Size	Fixed Asset Investment Intensity	Human Resource	Herfindahl-Hirschman Industrial Concentration Index (HHI)
2001	2294	Mean	.0065	.011	.0026	.16	.0020	.66	.82	20	13	.039
		Standard Deviation	.025	.017	.0045	.36	.0045	13	2.2	510	20	.041
		75^{th} Percentile	.0011	.0092	.0026	.0074	.0015	.27	.67	1.9	14	.058
2002	2486	Mean	.012	.013	.0028	.17	.0018	.34	.79	5.5	14	.042
		Standard Deviation	.20	.019	.0039	.37	.0032	2.0	2.4	72	12	.043
		75^{th} Percentile	.0021	.011	.0028	.0081	.0016	.26	.63	1.7	14	.062
2003	2727	Mean	.014	.012	.0035	.18	.0023	.34	.73	2.6	15	.034
		Standard Deviation	.31	.018	.0061	.38	.0051	6.2	2.3	21	14	.037
		75^{th} Percentile	.0022	.011	.0029	.0072	.0020	.22	.59	1.4	17	.046
2005	4046	Mean	.010	.014	.0053	.19	.0034	.17	.63	2.7	17	.032
		Standard Deviation	.031	.0091	.0085	.38	.0039	.57	2.3	32	14	.052
		75^{th} Percentile	.0014	.020	.0047	.013	.0044	.18	.50	1.2	19	.041

by the industrial output, and the ratio of administrative, financial, and selling expenses to sales value.

Legitimate instrumental variables need to be highly correlated with the endogenous variable but not correlated with the residual of the structural function. R&D activity leads to the development of new products, so R&D intensity is correlated with new product development. However, R&D intensity of a host country firm in a particular four-digit sector could not substantially impact the average new product intensity of all the MNEs in that sector. In this sense, new-product intensity is a qualified instrumental variable. In Section 3.3.2.2, we argued that a firm's financial resources contribute to its decision to invest in R&D. Therefore, R&D intensity is highly correlated with the variable measuring a firm's financial capability. Furthermore, R&D decision of a host country firm can hardly affect the average ratio of the administrative, financial, and selling expenses to the sales value of all the MNEs in the same four-digit industry sector. Therefore, the ratio of administrative, financial, and selling expenses to sales value is a valid instrument variable of the R&D intensity of foreign firms. The definition of the two instrumental variables is also summarized in Table 3.3.

3.3.3.3 Sample Selection

Our panels of firm-level data might suffer under selection bias. That is, we only observe R&D data for firms that choose to invest in R&D. If the missing R&D data (how much would firms that currently have zero R&D investment invest in R&D if they choose to do so) were missing at random, regular estimation methods would be fine. However, if the missing R&D data were not missing at random, that is, the disturbance in the equation of whether firms would invest in R&D is correlated with the disturbance in the equation of how much firms would investment in R&D if they choose to do so, estimation methods that ignore this correlation may lead to biased estimates of regression parameters.

We follow a two-stage Heckman procedure to correct the regression for sample selection. In the first stage, the selection equation is estimated for each of the observation periods. Following Veugelers and Cassiman (1999) we regress whether a firm has positive R&D intensity on the following independent variables: size, export intensity, financial resources, market structure, and ownership dummies. In the second stage, we correct the selection bias and estimate equation (9) using the selected sample for each of the observation periods[30].

3.3.4 Empirical Results

The estimation results of Function 9 are presented in Table 3.5. The coefficients of R&D intensity of higher-level-group MNEs are significant and positive in three of the four regression models. The coefficients of R&D intensity of lower-level-group MNEs and R&D intensity of same-group MNEs are significant and positive in one of the four regression models. To correct the endogeneity of explanatory variables, we instrument the two statistically significant variables in the Tobit model results, i.e. R&D intensity of higher-level-group MNEs and R&D intensity of lower-level-group MNEs in the three-group results, and report the instrumental variable regression results in Table 3.6. The table shows that R&D intensity of higher-level-group MNEs and R&D intensity of lower-level-group MNEs are all maintained to be statistically significant in instrumental variable Tobit models. The results of Heckman selection

[30] Heckman selection model is rarely used in a panel data environment in the existing empirical literature. Theoretical studies on how to obtain consistent estimator of a panel data sample selection model can be found in Wooldridge (1995), Kyriazidou (1997) and Semykina and Wooldridge (2005).

Table 3.5: Effect of MNE R&D Activity: Ease of Learning Measured by Labor Productivity

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables			
	Five Groups		Three Groups	
	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model
R&D Intensity of Same-group MNEs	.33(.48)	.19(.11)*	-.26(.44)	.14(.10)
R&D Intensity of Higher-level-group MNEs	.005(.008)	.005(.003)*	.03(.007)***	.01(.002)***
R&D Intensity of Lower-level-group MNEs	.73(.75)	.22(.18)	1.74(.69)***	.27(.17)
Average R&D Intensity of Domestic Firms in the Same Four-digit Sector (Independent Variable)	-.15(.23)	.007(.064)	-.03(.23)	.03(.06)
Financial Resource	.04(.0004)***	.0005(.0003)	.04(.0004)***	.0005(.0003)
Firm Size	.008(.001)***	.002(.0004)***	.008(.001)***	.002(.0004)***
Fixed Asset Investment Intensity	-.002(.00004)***	-.00004(.00002)*	-.002(.00004)***	-.00004(.00002)*
Human Resource	.002(.0002)***	.0006(.00005)***	.002(.0002)***	.0006(.00005)***
Herfindahl-Hirschman Industrial Concentration Index (HHI)	.04(.09)	.02(.02)	.01(.09)	.01(.02)
Ownership Dummy Variables that are Statistically Significant at 10 Percent	limited liability companies and stock companies	stock companies	limited liability companies and stock companies	stock companies
Year Dummy Variables that are Statistically Significant at 10 Percent	2002 and 2003	2002	2002, 2003	2002
Industry Dummy Variables that are Statistically Significant at 10 Percent	All industry sectors except other electronic component	All industry sectors except other electronic component	All industry sectors except other electronic component	All industry sectors except other electronic component
Number of Observations	11553	11553	11553	11553
Number of Groups for Panel Model		6456		6456
Likelihood-ratio Chi-squared Statistics for Tobit Model / Wald Chi-squared Statistics for Panel Tobit Model	2693.8***	731.5***	2717.71***	758.9***

Note: 1. The data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%, * denotes a significance level of 10%.

regression are reported in Table 3.6. The null hypothesis of sample selection is rejected for each of the observation periods. The corrected results show that R&D intensity of same-group MNEs and R&D intensity of higher-level-group MNEs are statistically significant in 2005 and 2001 respectively. The coefficients of R&D intensity of lower-level-group MNEs become insignificant during the observation periods.

Overall, the robust positive signs of R&D intensity of higher-level-group MNEs indicate that R&D offshoring of MNEs stimulated R&D effort of domestic firms in the IT industry in China. Moreover, the R&D effort of domestic firms is mainly influenced by foreign firms with higher technological and management capabilities, implying relatively strong learning from foreign firms with higher capabilities.

In addition, the results presented in Table 3.5 reveal that a host country firm with greater financial resources, a larger size, and more R&D human resources is more likely to invest more in R&D. However, contradictory to our prediction, a firm investing more in fixed assets tends to invest less in R&D. Neither the average R&D intensity of host country firms in the four-digit sector nor the industrial concentration index has statistically significant influence on R&D decisions of an individual firm. We also find that limited liability and stock companies tend to invest more in R&D than firms of other ownership status. The R&D effort level of domestic firms in 2001 exhibited a statistically significant difference from that of domestic firms in 2002, but it did not between 2001 and 2003 or between 2001 and 2005.

Since Tobit model estimates the effect of change in explanatory variables on the expected value of latent dependent variable holding other explanatory variables

Table 3.6: Instrumental Variable Tobit Model and Heckman Selection Model

Independent Variables	R&D intensity of higher-level-group MNEs is instrumented	R&D intensity of lower-level-group MNEs is instrumented	Heckman Correction
R&D Intensity of Same-group MNEs	-.37(.44)	-.84(.49)*	.39(.15)*** (year 2005)
R&D Intensity of Higher-level-group MNEs	.039(.007)***	.03(.007)***	.007(.004)* (year 2001)
R&D Intensity of Lower-level-group MNEs	1.72(.69)**	7.04(2.10)***	Insignificant in observation periods
Wald Test of Exogeneity (Chi-squared Statistics)/Wald Chi-squared Statistics	16.52 ***	7.83 ***	Significant at 1% level in the observation periods

Note:
1. The data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%l, * denotes a significance level of 10%.
2. To simplify, the coefficients of the controlling and dummy variables are not reported.

constant, marginal effect is calculated and reported in Table 3.7[31]. Table 3.7 shows that in the 5-Group analysis, an increase of .1 in the R&D intensity of foreign firms with higher capabilities will lead to an increase of .0001 in the R&D intensity of a domestic firm, holding other explanatory variables constant. In the 3-Group analysis, an increase of .1 in the R&D intensity of foreign firms with higher capabilities will cause an increase of .0003 in the R&D intensity of a domestic firm.

We present the regression results of Functions 10 and 11 in Tables 3.8 using both the five-group and the three-group approaches. Panel A in Table 3.8 shows that when we use the sub-sample of domestic firms with relatively low technology and management capability (domestic firms in Groups 1, 2, 3, and 4 or Groups 1, 2, and 3 in the five-group approach and domestic firms in Group 1 and 2 in the three-group approach) in the regression, the coefficients of R&D intensity of higher-level-group and same-group MNEs are insignificant; however, the coefficients of R&D intensity of lower-level-group MNEs are generally significant and positive in the regressions. When we use the sub-sample of domestic firms with relatively high technology and management capability (domestic firms in Groups 2, 3, 4 and 5 or Groups 3, 4, and 5 in the five-group approach and domestic firms in Group 2 and 3 in the three-group approach) in the regression (Panel B in Table 3.8), the coefficients of R&D intensity of lower-level-group and same-group MNEs are insignificant; however, the coefficients of R&D intensity of higher-level-group MNEs are generally significant and positive in the regressions.

To summarize, although the overall impact of MNE R&D offshoring on the R&D effort of domestic IT firms is positive, its influence differs between domestic firms with high and low technology and management capabilities. The R&D effort of

[31] The marginal effect is calculated at the mean of explanatory variables.

Table 3.7: Marginal Effect of MNE R&D Activity: Ease of Learning Measured by Labor Productivity

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables			
	Five Groups		Three Groups	
	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model
R&D Intensity of Same-group MNEs	.07(.11)	.05(.03)*	-.06(.10)	.03(.02)
R&D Intensity of Higher-level-group MNEs	.001(.002)	.001(.0006)*	.007(.002)***	.003(.0005)***
R&D Intensity of Lower-level-group MNEs	.16(.16)	.05(.04)	.38(.15)***	.06(.04)
Average R&D Intensity of Domestic Firms in the Same Four-digit Sector (Independent Variable)	-.03(.05)	.002(.02)	-.007(.05)	.007(.01)
Financial Resource	.008(.0001)***	.0001(.00007)	.008(.0001)***	.0001(.00007)
Firm Size	.002(.0002)***	.0005(.00009)***	.002(.0002)***	.0005(.00009)***
Fixed Asset Investment Intensity	-.0005(.00001)***	-8.68e-06(.00000)*	-.0005(.00001)***	-8.47e-06(.00000)*
Human Resource	.0004(.00004)***	.0002(.00001)***	.0003(.00004)***	.0001(.00001)***
Herfindahl-Hirschman Industrial Concentration Index (HHI)	.008(.02)	.004(.005)	.002(.02)	.003(.005)
Ownership Dummy Variables that are Statistically Significant at 10 Percent	limited liability companies and stock companies	stock companies	limited liability companies and stock companies	stock companies
Year Dummy Variables that are Statistically Significant at 10 Percent	2002 and 2003	2002	2002, 2003	2002
Industry Dummy Variables that are Statistically Significant at 10 Percent	All industry sectors except other electronic component	All industry sectors except other electronic component	All industry sectors except other electronic component	All industry sectors except other electronic component
Number of Observations	11553	11553	11553	11553
Number of Groups for Panel Model		6456		6456
Likelihood-ratio Chi-squared Statistics for Tobit Model / Wald Chi-squared Statistics for Panel Tobit Model	2693.8***	731.5***	2717.71***	758.9***

Note: 1. Data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%, * denotes a significance level of 10%.

Table 3.8: Effect of MNE R&D Activity on Low- and High-Capability Domestic Firms

Panel A: Low-Capability Domestic Firms

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables					
	Five Groups				Three Groups	
	Groups 1, 2, 3, 4		Groups 1, 2, 3		Groups 1, 2	
	Tobit	Panel Tobit	Tobit	Panel Tobit	Tobit	Panel Tobit
R&D Intensity of Same-group MNEs	-.21(.92)	.09(.20)	-.35(1.09)	.03(.24)	-1.35(.96)	-.25(.22)
R&D Intensity of Higher-level-group MNEs	-.12(.63)	.17(.14)	-.84(1.19)	-.17(.28)	.41(.67)	.11(.16)
R&D Intensity of Lower-level-group MNEs	1.28(.94)	.37(.20)*	2.22(1.38)	.75(.31)**	2.83(1.37)**	.82(.32)***

Panel B: High-Capability Domestic Firms

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables					
	Five Groups				Three Groups	
	Groups 2, 3, 4, 5		Groups 3, 4, 5		Groups 2, 3	
	Tobit	Panel Tobit	Tobit	Panel Tobit	Tobit	Panel Tobit
R&D Intensity of Same-group MNEs	.36(.19)**	.16(.12)	.43(.22)**	.18(.11)	.25(.19)	.17(.11)
R&D Intensity of Higher-level-group MNEs	.004(.003)	.006(.003)**	.001(.004)	.004(.003)*	.01(.003)***	.008(.002)***
R&D Intensity of Lower-level-group MNEs	.10(.30)	.006(.18)	-.05(.34)	-.24(.18)	.49(.30)*	.06(.18)

Note:

1. The data in parentheses are standard deviations. *** denotes a significance of 1%, ** denotes a significance of 5%, * denotes a significance of 10%.
2. The labor productivity of groups 1, 2, 3, 4, 5 is in ascending order, i.e., Group 1 has the lowest labor productivity, and Group 5 has the highest labor productivity. Similarly, when the sample is divided into 3 groups the labor productivity of groups 1, 2, 3 is in ascending order, i.e., Group 1 has the lowest labor productivity, and Group 3 has the highest labor productivity.
3. To simplify the tables, the coefficients of control variables and dummy variables are not reported.

low-capability domestic firms is only influenced by the R&D intensity of lower-level-group MNEs; and the R&D effort of high-capability domestic firms is only influenced by the R&D intensity of higher-level-group MNEs.

To examine whether the results still hold if we include in the analysis the difference in industrial labor productivity, we separate the analysis for technology-intensive and labor-intensive industry sectors and present the results in Table 3.9 through Table 3.11. Table 3.9 shows the estimation results of equation 9 for technology-intensive and labor-intensive industries. R&D intensity of higher-level-group MNEs has a significant and positive impact on the R&D effort of domestic firms in both technology-intensive and labor-intensive industry sectors. This is consistent with the results using the full sample. Moreover, R&D effort of domestic IT firms in labor-intensive industry sectors also increases with the R&D intensity of lower-level-group MNEs. The results after correcting the endogeneity of explanatory variables and the sample selection bias (Table 3.10) in general support the results presented in Table 3.9.

Table 3.11 reports the results of equations 10 and 11 for technology- and labor-intensive industry sectors. R&D effort of low-capability domestic firms in both types of industry sectors increases with the R&D intensity of lower-level-group MNEs; and R&D effort of high-capability domestic firms increases with the R&D intensity of higher-level-group MNEs. This is completely consistent with the results using full sample.

Overall the analysis of firms in technology- and labor-intensive industries proves that the findings using full sample are robust. That is, R&D effort of domestic IT firms increases with the R&D intensity of foreign firms. The R&D effort of high-

Table 3.9: Effect of MNE R&D Activity: Ease of Learning Measured by Labor Productivity for Technology- and Labor-Intensive Industry Sectors

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables			
	Technology-Intensive Industry Sectors		Labor-Intensive Industry Sectors	
	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model
R&D Intensity of Same-group MNEs	-.05(.59)	-.02(.42)	-.30(.26)	-.19(.26)
R&D Intensity of Higher-level-group MNEs	.03(.01)**	.03(.01)***	.01(.002)***	.01(.003)***
R&D Intensity of Lower-level-group MNEs	1.65(.97)*	1.07(.69)	.91(.30)***	.70(.31)**
Number of Observations	4638	4638	6915	6915
Number of Groups for Panel Model		2704		4036
Likelihood-ratio Chi-squared Statistics for Tobit Model / Wald Chi-squared Statistics for Panel Tobit Model	1641.24***	5238.88***	764.04***	423.28***

Note:
1. Data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%, * denotes a significance level of 10%.
2. To simplify the table, the coefficients of control variables and dummy variables are not reported.

Table 3.10: Instrumental Variable Tobit Model and Heckman Selection Model for Technology- and Labor-Intensive Industry Sectors

Panel A: Technology-Intensive Industry Sectors

Independent Variables	R&D intensity of higher-level-group MNEs is instrumented	R&D intensity of lower-level-group MNEs is instrumented	Heckman Correction
R&D Intensity of Same-group MNEs	-.16(.59)	-.52(.65)	4.00(2.40)* (year 2002) .43(.17)*** (year 2005)
R&D Intensity of Higher-level-group MNEs	.04(.01)***	.02(.01)*	.05(.03)* (year 2002)
R&D Intensity of Lower-level-group MNEs	1.6(.97)*	6.55(2.89)**	Insignificant in observation periods
Wald Test of Exogeneity (Chi-squared Statistics)/ Wald Chi-squared Statistics	9.35 ***	3.21*	Significant at 1% level in the observation periods

Panel B: Labor-Intensive Industry Sectors

Independent Variables	R&D intensity of higher-level-group MNEs is instrumented	R&D intensity of lower-level-group MNEs is instrumented	Heckman Correction
R&D Intensity of Same-group MNEs	-.33(.26)	-.44(.26)*	-1.83(.95)** (year 2002)
R&D Intensity of Higher-level-group MNEs	.01(.002)***	.01(.002)***	.01(.005)* (year 2005)
R&D Intensity of Lower-level-group MNEs	.90(.30)***	3.22(.65)***	-1.72(.97)* (year 2002) 3.32(1.68)** (year 2003)
Wald Test of Exogeneity (Chi-squared Statistics)/ Wald Chi-squared Statistics	4.18 **	16.35***	Significant at 1% level in the observation periods

Note:
1. Data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%l, * denotes a significance level of 10%.
2. To simplify the tables, the coefficients of control variables and dummy variables are not reported.

Table 3.11: Effect of MNE R&D Activity on Low- and High-Capability Domestic Firms for Technology- and Labor-Intensive Industry Sectors

Panel A: Technology-Intensive Industry Sectors

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables			
	Groups 1, 2		Groups 2, 3	
	Tobit	Panel Tobit	Tobit	Panel Tobit
R&D Intensity of Same-group MNEs	-2.07(1.71)	-2.16(1.80)	.21(.13)	.17(.13)
R&D Intensity of Higher-level-group MNEs	-.28(1.11)	-.52(1.16)	.002(.003)	.006(.004)*
R&D Intensity of Lower-level-group MNEs	6.35(2.66)**	5.68(2.83)**	.12(.23)	.0006(.23)

Panel B: Labor-Intensive Industry Sectors

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables			
	Groups 1, 2		Groups 2, 3	
	Tobit	Panel Tobit	Tobit	Panel Tobit
R&D Intensity of Same-group MNEs	-.28(.37)	-.15(.39)	-.03(.32)	.03(.32)
R&D Intensity of Higher-level-group MNEs	.41(.35)	.28(.35)	.007(.002)***	.008(.003)***
R&D Intensity of Lower-level-group MNEs	.80(.41)**	.71(.41)*	.24(.33)	.20(.34)

Note:
1. Data in parentheses are standard deviations. *** denotes a significance of 1%, ** denotes a significance of 5%, * denotes a significance of 10%.
2. Labor productivity of groups 1, 2, 3 is in ascending order, i.e., Group 1 has the lowest labor productivity, and Group 3 has the highest labor productivity.
3. To simplify the tables, the coefficients of control variables and dummy variables are not reported.

capability domestic firms is primarily influenced by the R&D intensity of higher-level-group MNEs; the R&D effort of low-capability domestic firms is mainly influenced by the R&D intensity of lower-level-group MNEs. The positive impact of higher-level group MNEs on high-capability domestic firms is intuitive and implies that high-capability domestic firms have the capability to learn much from MNEs with higher capability. The positive impact of lower-level group MNEs on low-capability domestic firms could be driven by potential threat of foreign firms in lower-level-group to a low-capability domestic firm[32].

The regression results of Functions 12 and 13 are presented in Table 3.12. The regression using the full sample reveals significant and positive effect of R&D intensity of MNEs within a city in the Tobit model but insignificant effect in the Panel Tobit model. The impacts of R&D intensity of MNEs within a province and R&D intensity of MNEs within neighboring provinces are insignificant. The regression analysis of host country firms in Guangdong province shows that both R&D intensity of MNEs within a province and R&D intensity of MNEs within neighboring provinces have significantly negative effects on the R&D intensity of domestic firms. The marginal effect is calculated to document the magnitude of the impact[33]. The

[32] High-capability domestic firms may also face the potential threat of foreign firms in lower-level group, but the empirical analysis shows that they choose to respond to the R&D effort of higher-level group MNEs because of the opportunity to learn from the more advanced foreign firms. Low-capability domestic firms, however, may find it difficult to learn from higher-level group MNEs, hence choose to respond to lower-level group MNEs to maintain the technological gap between them and the lower-level group MNEs. The impact of lower-group MNEs on the R&D intensity of low-capability domestic firms also reveals one limitation of the economic model presented in Section 3.2: the model did not differentiate low-capability domestic firms from high-capability domestic firms. We will need to model the two groups of domestic firms separately to provide theoretical explanation to the reactions of high- and low-capability domestic firms to foreign R&D effort.
[33] The marginal effect is calculated at the mean of explanatory variables.

Table 3.12: Effect of MNE R&D Activity: Ease of Learning Measured by Geographical Distance

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables							
	Overall Domestic Firms				Domestic Firms in Guangdong Province			
	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model
R&D Intensity of MNEs within a City	.52(.24)**	.07(.06)	.55(.24)**	.08(.06)	1.51(.57)***	.73(.61)	1.62(.58)***	.94(.62)
R&D Intensity of MNEs within a Province	-.22(.38)	-.15(.10)	-	-	-5.17(2.42)**	-8.78(2.83)***	-	-
R&D Intensity of MNEs within Neighboring Provinces	-	-	-.54(1.12)	-.25(.29)	-	-	-6.53(4.15)	-10.36(4.70)**

Note:

1. Data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%, * denotes a significance level of 10%.
2. To simplify the table, the coefficients of the control variables and dummy variables are not reported.

results are presented in Table 3.13. Table 3.13 shows that in Guangdong province, an increase of .1 in the R&D intensity of MNEs within a province will lead to a decrease of .19 in the R&D intensity of a domestic firm, holding other explanatory variables constant. An increase of .1 in the R&D intensity of MNEs within neighboring provinces will cause a decrease of .23 in the R&D intensity of a domestic firm, holding other explanatory variables constant.

To verify whether the negative impact of the R&D intensity of MNEs within a province and the R&D intensity of MNEs within neighboring provinces is not driven by location characteristics in Guangdong province [34], we collect data on the number of universities that have engineering and science majors for each city in Guangdong province and include this new variable in the analysis. The results are summarized in Table 3.14. The table shows that including number of universities in the analysis does not change the results. Furthermore number of universities does not have a significant impact on the R&D efforts of domestic firms in this sample, implying that university is not a major factor in Guangdong province for firms to choose location.

3.4 Discussion and Policy Implications

The empirical results based on technological distance between domestic firms and foreign firms imply that R&D intensity of host country firms in the IT industry in China generally increases with the R&D intensity of MNEs[35]. In particular, R&D intensity of domestic firms is mainly influenced by MNEs with higher technological and management capabilities, implying relatively strong learning from MNEs with

[34] For example, a firm that invests in R&D chooses to locate in a city because of the characteristics of the city that facilitates its R&D activities; a firm that does not invest in R&D chooses to not locate in this city because these characteristics are not of interest to it.

[35] It could be driven by some common technological opportunity, but our control of years in the empirical analysis partially removed that effect.

Table 3.13: Marginal Effect of MNE R&D Activity: Ease of Learning Measured by Geographical Distance

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables							
	Overall Domestic Firms				Domestic Firms in Guangdong Province			
	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model
R&D Intensity of MNEs within a City	.11(.05)**	.02(.02)	.12(.05)**	.02(.02)	.34(.13)***	.16(.13)	.36(.13)***	.20(.13)
R&D Intensity of MNEs within a Province	-.05(.08)	-.04(.02)	-	-	-1.16(.54)**	-1.90(.61)***	-	-
R&D Intensity of MNEs within Neighboring Provinces	-	-	-.12(.25)	-.06(.07)	-	-	-1.46(.93)	-2.25(1.02)**

Note:
1. Data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%, * denotes a significance level of 10%.
2. To simplify the table, the coefficients of the control variables and dummy variables are not reported.

Table 3.14: Effect of MNE R&D Activity: Ease of Learning Measured by Geographical Distance Controlling Location Characteristic

Independent Variable	The R&D Intensity of Domestic Firms as Dependent Variables			
	Domestic Firms in Guangdong Province			
	Tobit Model	Panel Tobit Model	Tobit Model	Panel Tobit Model
R&D Intensity of MNEs within a City	1.51(.57)***	.68(.62)	1.63(.58)***	.92(.62)
R&D Intensity of MNEs within a Province	-5.25(2.52)**	-9.48(2.95)***	-	-
R&D Intensity of MNEs within Neighboring Provinces	-	-	-6.43(4.27)	-10.89(4.82)**
University	.00009(.0008)	.0009(.001)	-.00008(.0008)	.0005(.0009)

Note:
1. Data in parentheses are standard deviations. *** denotes a significance level of 1%, ** denotes a significance level of 5%, * denotes a significance level of 10%.
2. To simplify the table, the coefficients of the control variables and dummy variables are not reported.

higher capabilities. The R&D intensity of a domestic firm with relatively low technological and management capability is primarily affected by the R&D intensity of MNEs whose capabilities are lower than those of the domestic firm and not by the R&D intensity of more advanced MNEs. For domestic firms with relatively high technological and management skills, we find that their R&D effort is influenced by MNEs whose capabilities are higher than those of the domestic firms and not by MNEs with lower technological and management capabilities.

The results based on geographical distance demonstrate that the R&D effort of domestic firms in the IT industry in China is not affected MNEs located outside of a city. Furthermore, the analysis of domestic firms in Guangdong province shows that R&D investment by MNEs located outside of a city, but within the same province and neighboring provinces, has a negative impact. These findings consistently reveal a geographical boundary for a positive impact of the R&D offshoring of MNEs in the Chinese IT industry.

Given the insights derived from the theoretical model and empirical evidence, we argue that through industry policy, governments in emerging economies could promote the productivity growth of domestic firms.[36] In fact, industry policy that encourages cooperation between domestic and foreign firms and facilitates learning within domestic firms through R&D or innovation subsidy is already in place in emerging economies. Such sound policy could ensure success, as argued by Laderman et al. (2003). Their study suggested that a host country could unleash a sequence of investments by success-

[36] Defined by Pack and Saggi (2006), an industry policy is a selective governmental intervention that promotes the economic growth of industry sectors in the way that would not occur in the absence of such intervention in the market equilibrium.

fully inducing FDI from one or two important firms. Our analyses shed new light on the relevance and implementation of industry policies conducive to technological learning in emerging economies.

First, governments in emerging economies should be cautious about encouraging MNE R&D offshoring since its positive impact on the R&D effort of a domestic firm would be only conditional. Second, our findings show that the limited technological and management capabilities of domestic firms would impede their benefit from spillover generated by the R&D investment of MNEs. If an industry in a developing economy is relatively developed, industry policies favorable to foreign R&D investment would provide domestic firms with opportunities to learn from the more advanced foreign competitors. The knowledge gained from this process is essential if emerging economies are to achieve successful technological upgrading at a later stage. However, if an industry falls into the category of an "infant" industry, government in a developing economy has reasons to be suspicious about the impact of the R&D investment of MNEs on the R&D effort of domestic firms. Even so, the R&D investment of MNEs could still exert a positive influence on the local economy such as increasing local employment opportunities.

3.5 Conclusions

MNE R&D offshoring to emerging economies has become a prominent phenomenon in recent years. In this paper, we investigate how the R&D offshoring of MNEs affects the R&D effort of host country firms. To investigate this issue, we have developed a two-stage non-cooperative game as a theoretical foundation and gathered empirical evidence from a large dataset of 12,309 manufacturing firms in the IT industry in China.

In general, we have found that domestic firms in the Chinese IT industry have increased R&D effort in response to the R&D offshoring decisions of MNEs. This implies that the positive spillover effects exceeded the negative competition effects of the R&D offshoring of MNEs on the R&D effort of domestic firms. Nevertheless, the magnitude of the impact is determined by the relative distance between the MNEs and the domestic firms. The R&D effort of high-capability domestic firms is primarily influenced by the R&D intensity of MNEs with higher technological capability; the R&D effort of low-capability domestic firms is mainly influenced by the R&D intensity of MNEs with lower technological capability. The positive impact of MNEs' R&D investment on domestic firms diminishes as the geographical distance between them increases.

Our findings provide important policy implications to both the host and the home countries of R&D offshoring. Understanding the response of host country competitors is not only critical if the MNEs are to choose the right R&D offshoring strategies but also important if policy makers in the MNEs' home countries are to successfully regulate the MNE R&D offshoring to emerging economies. For host country policy makers, our findings provide important implications regarding how emerging economies could leverage their FDI policy to facilitate technology spillover from MNEs and technology learning in domestic firms. The success of technology spillover and technology learning in host countries will greatly contribute to the long-term development of the indigenous innovation capabilities of domestic firms.

Our study is limited in scope because we investigate only intra-industry R&D spillovers, not R&D spillovers from MNEs in one industry to other industries in host country. As a result, our findings could underestimate the overall positive effect of

MNEs' R&D offshoring on host country firms. In addition to including other industries, such as upstream and downstream industry sectors, in the analysis, this study could be extended to industries other than IT, which will allow us to examine the generalization of our findings and the R&D spillover patterns in sectors other than IT. Furthermore, this research could examine whether the potential influence of MNEs' R&D offshoring on the R&D effort of domestic firms would result in changes in the productivity of domestic firms.

CHAPTER 4

SOFTWARE DESIGN STRATEGIES IN MARKETS WITH OPEN SOURCE COMPETITORS

4.1 Introduction

Despite all the critiques about open source software (OSS), OSS has made significant inroads in many areas of software development, from operating systems (Linux), programming languages (Perl) and Web browsers (Mozilla Firefox) to Web servers (Apache), database management systems (MySQL), mail management systems (Sendmail) and typesetting engines (Tex) (Economides and Katsamakas, 2006; Lerner and Tirole, 2002; Schmidt and Schnitzer, 2003; Bitzer and Schroder, 2006; Gaudeul, 2007; Sen, 2007; Lerner and Tirole, 2005). The growing market share and volume of installed base of some OSS force major proprietary software vendors (PRVs) to seriously consider the competitive position of OSS in the industry and employ a variety of strategies to counter the challenges of successful open source (OS) movement. For instance, Oracle stated in its analysis of key competition in 2007 that "The enterprise software industry is intensely competitive…open source alternatives such as MySQL AB in database, Red Hat, Inc. in middle ware, and SugarCRM Inc. in applications, are also impacting the competitive environment…in the sale of database software…our competitors include … the open source databases, MySQL and PostgresSQL…In the sale of middleware products…, our competitors include… open source vendors such as Red Hat, Apache and ObjectWeb. …Business intelligence competitors include…open source vendors Netezza Corporation…"[37] Sun Microsystems similarly stated in its annual report that "we are seeing increased competition and pricing pressures from competitors offering systems running Linux software and other open source software."[38] BEA Systems acknowledged that "competitive pressures and open source availability of functionally competitive software could require us to reduce the price of our products and related services, which could harm our business."[39]

Facing the challenge of OSS, commercial firms employ a number of strategies and build innovative business models to counter the challenges and to capitalize on OS movement (e.g., Lawton, 2007; Tam, 2004; Bank, 2003). For example, Red Hat uses a support business model where it commercially provides complementary services that are not supplied efficiently by the Linux community. Major PRVs such as Oracle, HP, Novell, IBM, SAP and Sun Microsystems are well-known advocates of OS movement (Vara, 2006; Tam, 2004; Stone, 2004; Kerstetter, 2004). For instance, Oracle started selling its own technical support for Red Hat's version of Linux in 2007; Sun Microsystems is also determined to offer free downloads of its Solaris software; In 2004, IBM submitted a lightweight database called Cloudscape to the OS community - Apache Software Foundation, an organization that manages the widely used Apache Web server.

Academic researchers have also made significant contribution examining the issues for competitive strategy in the landscape of OS movement (e. g., Bessen, 2006; Bitzer, 2004; Lerner and Tirole, 2002; Lerner and Tirole, 2005; Dahlander and Magnusson, 2006; Haruvy, Sethi and Zhou, 2005; Schmidt and Schnitzer, 2003; Economides and

[37] Source: Oracle annual report 2007.
[38] Source: Sun Microsystems annual report 2007.
[39] Source: BEA Systems annual report 2007.

Katsamakas, 2006a; Rossi and Bonaccorsi, 2006). Despite the growing body of literature investigating the strategic issues between PRVs and their OS counterparts, most of the prior pieces focus on software market where vendors' revenue mainly comes from software license. Nevertheless, OSS is also prominent in software markets where there is significant demand for maintenance and support service after the sale of software license. That is, both software license and post-sale service contract represent considerable portion of some PRVs' revenue. This is shown in Table 4.1.

Table 4.1 is a snapshot of influential OSS. It shows that OSS is not only prominent and known in software markets where users do not require much support and service (e.g., Web browser and typesetting), but also in software markets where a significant portion of PRVs' revenue comes from maintenance and support of software (e.g., infrastructure software[40] and database management systems[41]). For instance, as one of the world's largest database software providers, 33% of Oracle's total revenue in 2007 came from software license, and 46% from software maintenance and product support; maintenance and support also represented the highest margin business unit for Oracle[42]. Sun Microsystems, a major provider of infrastructure product, earned 28.56% of its total revenue in 2007 from support services consisting primarily of maintenance contract[43]. BEA Systems, Inc., a world leader in enterprise application and service infrastructure software, earned 40.9% of its revenue from license fees and 59.1% from maintenance and

---

[40] For example, Apache dominates the web server market; Sendmail and Postfix are popular OSS in E-mail server market; Linux is well known in operating systems market.
[41] MySQL, PostgresSQL and Firebird are popular OS databases.
[42] Source: Oracle annual report 2007.
[43] Source: Sun Microsystems annual report 2007.

Table 4.1: Major Open Source Projects

Program	Nature of program	Year of introduction	Competitors	Market penetration	For profit?	OS Business model
MySQL	DBMS	1995	Oracle, IBM, Microsoft	>1 million installations (2005)	Yes MySQL	Warranty, customer support, optionally additional functionality
Linux	Server operating system	1991	Microsoft Windows	25.7% of total shipments in 2008 (IDC prediction)	Yes Red Hat	Support and other complementary services of Linux
Sendmail	Internet mail transfer agent	1979	Microsoft Exchange	Handle ~80% of Internet e-mail traffic (2002)	Yes Sendmail Inc.	Service
Apache	Web server	1994	Microsoft IIS	~55% (February 2008)	No	N/A
Perl	System administration and programming language	1987	Java (Sun) VB, ActiveX (Microsoft)	Estimated to have 1 million users (2002)	No	N/A
Tex	Typesetting	1982	Microsoft Word	N/A	No	N/A
Mozilla Firefox	Web browser	2004	Microsoft IE	17.83% of the recorded usage share of Web browsers (March 2008)	No	N/A

Note: information presented in the table is from Lerner and Tirole (2002), Gaudeul (2007), Dahlander and Magnusson (2006), Economides and Katsamakas (2006b), Wittig and Inkinen (2004), Wheeler (2005), Netcraft (2008), Net Applications (2008).

support of its infrastructure software[44]. Generally, customers who purchase enterprise software licenses may also enter into maintenance contracts covering technical support services, regular software maintenance, and software updates and enhancements. These maintenance contracts generate persistent revenue for enterprise PRVs after the transaction of software license (e.g., Banker and Slaughter, 1997; Tan and Mookerjee, 2005).

This paper is motivated by the significant difference between the two families of software, both of which witness notably dynamic and influential OS movement in recent years. Software like database management systems (DBMS) in general requires much more maintenance service and support than software for typesetting, Web browsing or gaming purpose. The amount of support and service needed is determined in general by the complexity and scope of the software and whether the software is mission-critical[45]. For example, database types of software normally have higher complexity, are of larger scope and are more business-critical than software such as Web browser and computer games. As a result the former naturally require much more maintenance and support than the latter do. Furthermore, as support for software is complementary to software characterized by large demand for support service, the software itself and the support associated with it intervene with each other. Hence, it is important to examine them in an integrated framework; while for software with low demand for support service, examining the software alone would be sufficient.

The fundamental difference between the two families of software demands separate treatments in investigating commercial software firms' OS strategy. For example, do

[44] Source: BEA Systems annual report 2007.

[45] Poor design in software may also lead to high demand for support. In this paper, I focus on service demand driven by software characteristics and assume good quality in software design.

PRVs employ different OS strategies for software with low demand for service versus for software with high demand for service? How does the impact of OSS differ on the market structure for the two types of software market? The purpose of this Chapter is threefold. First, I investigate how activities in OS community enhance or replace the economic activities of PRVs. In particular, how does the rise of OSS have an effect on PRV's software design strategy in terms of number of features and functions bundled in the software? Second, I investigate the impact of OSS on software market structure and on pricing strategy of PRV. Third, I examine how the impacts of OSS differ between software market characterized by high demand for service and by low demand for service.

4.2 Literature Review

This paper is related with two streams of research: commercial software firms' OS strategies, and the impact of OS movement on PRS design. There is much discussion among industry experts about commercial software firms' OS strategies (e.g., Lawton, 2007; Tam, 2004; Bank, 2003). The key question among these is: "How do firms seeking to sell products compete with *free*?" (von Krogh and von Hippel, 2006) Academic researchers summarize these strategies into four categories: a) market segmentation strategy (i.e., PRVs cover a market segment different from the one covered by OSS) (Bessen, 2006; Bitzer, 2004); b) complementary product/service strategy (Lerner and Tirole, 2002; Lerner and Tirole, 2005; Dahlander and Magnusson, 2006; Haruvy, Sethi and Zhou, 2005; Schmidt and Schnitzer, 2003; Economides and Katsamakas, 2006b); c) subsidy/support strategy (i.e., PRVs participate in or subsidize OS projects) (Bessen, 2006; Rossi and Bonaccorsi, 2006; Dahlander and Magnusson, 2006; Lerner and Tirole, 2002; Lerner and

Tirole, 2005; Haruvy, Sethi and Zhou, 2005); d) releasing proprietary codes strategy (Lerner and Tirole, 2005).

Gaudeul (2005) investigates the direct competition between one OS and one proprietary project in a duopoly model where both the cost and benefit of OSS are compared with PRS. The cost of adopting OSS includes lack of coordination in development and lack of interface. The cost of adopting PRS is that developers may choose to develop a limited number of features. The author found that OSS and PRS can coexist in equilibrium, but OSS is used by low-income customers or by developers. This proves the validity and stability of the market segmentation strategy. Bitzer (2004) similarly showed that decreasing heterogeneity between OSS and PRS can drive the price of PRS below its average cost, leading to PRV's loss of incentive to develop PRS and forcing PRV to exit the market.

The complementary product/service strategy has been brought up in a number of studies. Economides and Katsamakas (2006a) developed a framework to investigate a platform firm's pricing strategy for the platform and for complementary applications to the platform (two-sided pricing strategy of a platform firm) and how the pricing strategy influence the industry structure of the proprietary system as well as the OS system. They found that the overall profit of a proprietary system is larger than an OS system when users have a strong preference for application variety; however the variety of applications in the OS system is larger than in the proprietary system. The authors also examined the competition between OS platform and proprietary platform. They found that a vertically integrated proprietary system (proprietary platform and applications to the platform are provided by one firm) has a larger market share than an OS system if there is a relatively large demand for the proprietary application, a large transition cost for the OS platform and a relatively small demand for application of the OS platform.

Although the payback of market segmentation strategy and complementary product/service strategy to commercial software firms is straightforward, researchers found that the gain of releasing codes to OS community is not guaranteed. Haruvy, Sethi and Zhou (2005) examined conditions when a software firm should open its source codes versus close its source codes. They compared the revenue of a software provider under two scenarios: the software is developed in an OS way, where the provider's revenue comes from complementary product of the OSS; the software is developed in a closed source way, where revenue comes from both software license and service. The authors demonstrated that when there are more productive in-house programmers and when the initial quality of the software is high, opening source codes may not be beneficial to a firm. On the other hand, opening source codes becomes a viable alternative to closed source only when the involvement of OS community in the project reaches a critical level. Moreover, OS is preferred when programmer wage exceeds some threshold, even with highly productive in-house programmers. Lerner and Tirole (2002) argued that PRVs release their proprietary codes to OS community in the hope that they can gain more by providing related services. However, releasing codes could be risky for PRVs if the payoff in the complementary service segment after the release does not exceed the payoff if the codes had not been converted to OS, which explains why the strategy of releasing codes to OS community is often observed in firms that are not major players in the proprietary segment.

Researchers also found that firms may have different incentives subsidizing or supporting OS projects. Some firms invest in OS development if the OSS is complement to the commercial software or hardware they produce (Schmidt and Schnitzer, 2003). Bessen (2006) investigated why firms contribute to OSS development when there is PRS from non-rival firms. The author argued that complexity of software limits the number of features included in PRS. Under-provision of features created a market for OS community, which tends to develop more complex and geekier applications. Firms participate in OSS development for their own needs of software features. The author concludes that OSS extends software market rather than replace PRS. That is why OSS can grow in the presence of a competing PRV, and PRS can coexist with OSS because they cover different market segments. Mustonen (2005) investigated when a PRV may support the development of substitute OSS. The author showed that this could happen when the support creates compatibility between the programs and that the programs exhibit network effects. So supporting a rival OSS may allow the PRS and the OSS to share the same network. The resulting strong network effect of the PRS, however, is earned at the price of the PRV losing its network of old PRS, and at the cost of providing support to the OSS. Lerner and Tirole (2002) argued that PRV may want to subsidize the OS movement for strategic reasons such as weaken a rival or lessen the dependence on other vendors.

Another line of research on OS strategy is related to OS license. Kim, Chen and Mukhopadhyay (2006) examined and compared three software pricing mechanisms: pricing of commercial software, pricing of OS product/service in OS dual license model, and pricing of open source product/service in OS support model. They investigated whether OS models are viable in monopoly and duopoly setting. They found that OS support model is viable in the presence of quality asymmetry among competitors no matter whether the quality of OSS is higher or lower than commercial software.

The second stream of research related to this paper concerns the impact of OS movement on PRS design, in particular, on the design of software features. Bitzer and Schroder (2007) stated that "the question of what impact this unusual development method (OS) has on innovation activity in the software sector has received surprisingly little attention thus far." They built a simple economic model and examined the impact of increased competition on innovation activity of both incumbents and entrants. They found that although incumbent commercial firms claim that technological progress will slow or even stop as a consequence of OSS entry, the model shows that it promotes innovation, that is, enterprises choose to increase the technological levels after the emergence of OSS. Bitzer and Schroder (2006) argued that the emergence of OSS has raised fears of a potential anti-innovation effect on two main issues. First, "the emergence of a no-cost competitor on the software market raised the question of whether commercial enterprise will be able to compete successfully…decreasing profits of commercial software producers will lower their ability to invest in R&D activity, thus resulting in slower technological progress in the software industry"; Second, "an anti-innovative effect of OSS may result if its development process is less efficient than that of commercially organized software." To examine the legitimacy of OSS's anti-innovation effect, the authors used the release history of Windows and Linux, Internet Explorer, Netscape's Web browser and Mozilla Firefox, as a potential indicator of technological progress in software industry. They found that "the data and cases presented do not offer evidence of an anti-innovative impact; on the contrary, if anything, the entry of OSS into commercial segments of the market appears to be associated with increased innovation activity." Economides and Katsamakas (2006b) extended the model developed by Economides and Katsamakas (2006a) and analyzed the innovation incentive in systems of two platforms: the OS and the proprietary platform. Comparing vis-a-vis the innovation investment of OS system and PR system, they found that the investment in application is stronger when the platform is OS. This is because the OS platform is available for free, enabling the application provider to set a larger price and capture a larger profit than the application provider for the proprietary platform. The authors also found that the level of investment in the application affects the level of investment in platform due to the complementarities between the application and the platform. In particular, the marginal benefit of investing in the platform decreases with the level of investment in the application.

Overall, most of the prior research investigate one type of software in the landscape of OS movement and do not systematically compare software that is different in nature. One exception is the research by Sen (2007). Sen (2007) compared the impact of OSS on PRV's software usability and pricing decisions for two types of software: software with strong network effect versus software with weak network effect. The author found that the two types of software markets reveal very different behavior given the entry of OSS. In software markets with low network effect, PRV is better off in the presence of competition from OSS while PRV in markets with high network effect is threatened by the presence of OSS. Furthermore, PRV has little incentive to improve the usability of their software in markets with low network effect, while the same strategy may drive PRV to exit the market in software market characterized by strong network effect. Lee and Mendelson (2008) similarly investigated the competition between a PRV and

OSS developers in software market characterized by network effect. They found that if the PRV enters the market first, it can capitalize the first-mover advantage by improving its product features.

4.3 A Model

4.3.1 A Case of MySQL

I draw on the case of MySQL to promote the model to be presented in the next Section. MySQL is a popular OS DBMS. The company was acquired by Sun Microsystems in February 2008. It was founded in 1995 and is headquartered in Sweden today. It has several offices throughout the world with thousands of community members worldwide. MySQL is in the list of OSS building blocks called LAMP, representing Linux, Apache, MySQL and Perl (Wittig and Inkinen, 2004). It is used by Google, Yahoo!, Nokia, YouTube, Amazon and Travelocity among others. MySQL has received much attention of both media (e.g., Lawton, 2007; Tam, 2004; Bank, 2003; Lacy, 2006) and academic researchers (e.g., Mustonen, 2005; Dahlander and Magnusson, 2006; Lerner and Tirole, 2005; Kim, Chen and Mukhopadhyay, 2006). Mustonen (2005), Wittig and Inkinen (2004), Wittig (2006), and Dahlander and Magnusson (2006) prepared MySQL cases and documented the history, current market position and business model of MySQL.

MySQL is known not only because of its high growth rate and large volume of installed base, but also because it challenges a high-margin database software category and could reshape the database market as Linux has reshaped the server operating systems market. In their case study of MySQL, Dahlander and Magnusson (2006) mentioned that the most prominent PRSs that are competitors of MySQL are the databases of Oracle and IBM. Industry analysts reported that "facing competition from open-source alternatives", not only Oracle, but also "IBM and Microsoft have each lowered their database prices and created low-end bundles aimed at smaller organizations and partners." (LaMonica, 2005)

MySQL used to use a so-called dual license model: the software is available under both OS license GPL and proprietary license. In 2005 MySQL introduced the MySQL Network described as "a subscription service that provides updates, alerts, notifications, knowledge base and production level support, that make it possible for companies to easily manage hundreds or thousands of MySQL servers. The core MySQL database remains open source, but these services are only available to paying customers." MySQL's EVP of sales commented "we are following the Red Hat pattern. The subscription business is growth business for us and that will overtake our OEM business" (Wittig, 2006).

MySQL has been improving its product. In October 2005, MySQL released version 5.0 which incorporated comparable features, such as stored procedures and triggers, to those of Oracle, IBM and Microsoft. These features were explicitly outlined by Oracle in 2004 to show the inferiority of MySQL (Wittig, 2006). With more advanced features released, MySQL is put in more direct competition with Oracle, IBM and Microsoft.

4.3.2 Model Setup

I model the competition between a PRV and an OSS in a software market characterized by high demand for support and service. The PRV sells software license and support service for its proprietary software; the OSS is available for users free of charge; an OS service provider adopts a support model and sells support service for the OSS. In the database market, Oracle is the PRV, MySQL database is the OSS, and MySQL is the service provider for the OS database software. I build the model on the literature of vertical differentiation. Following industry practice, I model the decision making process of the players as a three-stage noncooperative game (with no collusion).

Stage 1: Software Design.

In software market characterized by high demand for service, an OSS with *Q* features is provided free of charge[46]. A PRV chooses number of features/functions *q* for its software after observing *Q*. Let $q, Q \in [0, \bar{q}]$ where $\bar{q}$ represents the maximum number of features/functions that can be developed for a PRV to earn a nonnegative profit. The development cost is $C(q)$[47] where $C'(q) > 0$, and $C''(q) > 0$.

Number of features/functions bundled in the software is treated as the PRV's decision variable in the first stage because of the following reasons. First, the model is intended to build an integrated framework including both software and subsequent support service for the software. The software engineering literature reports that the amount of maintenance service required for one software is closely related with the complexity in

---

[46] Q is an exogenous variable in this model because the major incentive of OS developers is not to compete with PRV but for self recognition in OS community. Therefore OS developer is not modeled as a player in the game.

[47] *C(q)* includes the cost to design and develop *q* features and to ensure the quality of the software if PRV is an entrant. If PRV is an incumbent, *q* represents the overall number of features bundled in the software; *C(q)* includes the cost to design and develop the incremental features, the cost to ensure the quality of the software, including compatibility of new features to the system, maintainability, extensibility, robustness, reliability of the software, after including the incremental features.

software design; more complex software requires more maintenance and support effort[48] (e.g., Banker, Davis and Slaughter, 1998; Abreu and Melo, 1996; Briand, Morasca and Basili, 1999; Li, et al. 1995). It has been shown that function point[49] in software is related with complexity or size of software, hence is related with maintenance effort for the software (Banker, Davis and Slaughter 1998; Swanson and Beath, 1989)[50]. Second, functionality is one of the main evaluation criteria for DBMS[51] (Wittig and Inkinen, 2004), which further gives good reason for using the number of functions/features as PRV's decision variable in the software design stage.

Stage 2: Software Pricing.

PRV chooses price p for the PRS developed in stage 1. Each consumer acquires one and only one copy of the software from either the OS community or the PRV (software market is fully covered). D denotes the demand for PRS. Following Cohen and Whang (1997), D is determined by Q, q, p and the expected value of PRS service and OSS service (denoted by A and B respectively). The rest of the market adopts OSS. Demand for OSS is denoted as D^o.

[48] Poor software design may also incur high demand for support service. However, this may not be of major concern in the case of MySQL database and Oracle database. MySQL uses a systematic approach to manage the process of software design and development. Its database has an easy-to-use interface (Wittig and Inkinen, 2004). In general, OSS has to reach mature stage to be able to compete with an existing PRS. Literature has shown that OSS that reached mature stage is quite comparable with corresponding PRS in terms of quality in software design (Wheeler, 2005).
[49] A function point is a unit of measurement to express the amount of business functionality an information system provides to a user.
[50] The literature of software engineering uses a variety of metrics to measure software complexity. These are out of the scope of this paper and are not discussed here. A good discussion of these metrics can be found in Banker, Davis and Slaughter (1998).
[51] Other criteria include performance, stability/reliability, ease of use, and cost.

Stage 3: Service Pricing[52].

As discussed in previous Section, software post-sale service provides additional value to software users. In this stage, PRV chooses service price p_s for the PRS and a service provider specializing in OSS determines the service price p_s^o for OSS. I normalize service cost per customer to zero, so that service price is the profit margin. PRS users purchase service from PRV and OSS users purchase service from OSS service provider[53] if their net surplus of purchasing the service is positive. For the same type of buyers (defined in the next paragraph), I assume service is of higher value to buyers if the software purchased in stage 2 is more complex, that is, the software is bundled with more features/functions. This is reasonable because if software is simple and easy to use, users do not have much need for technical support; while if software is complicated or users lack experience in using it, technical support becomes much more important to keep the software running and to update the software to topnotch condition. Following this logic, I denote the gross value of service for PRS and OSS by $V(q)$ and $V(Q)$ correspondingly where $V'(.) > 0$.

Consumer Preference.

---

[52] Some PRVs offer a bundle of software license and support contract. Oracle offers separate prices for license and support. I follow the practice of Oracle and separate the pricing decision for license and support into two stages. The result remains the same if I combine stage 2 with stage 3. The result will change if the price for a bundle with both license and support is different from the sum of the prices if the two are sold separately. Since bundling is not the focus of this paper, I separate the pricing for license and support for simplicity.

[53] This is consistent with common industry practice.

Consumers are uniformly distributed in a unit square where the two axes represent their taste for software features and service respectively[54]. For example, consumers located at (1, 1) have the highest valuation for software features and service; consumers located at (1, 0) have the highest valuation for software features but lowest valuation for service. Consumers know their type about software design at the beginning of stage 2, but get to know their type about service only after they consume the software, i.e., at the end of stage 2. This is because consumers' type for service is closely related with consumers' experience with the software. If the software runs smoothly, they may find less value of the service. If they have difficulty with the software, they may have greater need for the service. As a result, when consumers select software in stage 2, they use the expected surplus they anticipate to derive from the service as part of the input.

4.4 The Equilibrium

I start with the last stage to solve the model.

Stage 3. Consumers' utility is $\theta' V(q) - p_s$ if they purchase PRS service and $\theta' V(Q) - p_s^o$ if they purchase OSS service. θ' represents consumers' type about service and is uniformly distributed over the interval [0,1]. Demand for service is $D_s = (1 - \frac{p_s}{V(q)})D, D_s^o = (1 - \frac{p_s^o}{V(Q)})D^o$ for PRS service and OSS service correspondingly. PRV's problem at this stage is to maxi-

[54] This approach is commonly used in modeling consumer preferences in two dimensions. Consumers' tastes for software features and service are assumed to be independent from each other because their relationship could be dependent on consumers' specific requirements for the software, their technical capabilities and prior experience with similar products. Therefore, the relationship between them could be ambiguous. Customers with high valuation for software features may have low or high valuation for service; the same applies to customers with low valuation for software features.

mize his service profit determined by $p_s D_s$. OSS service provider maximizes his profit determined by $p_s^o D_s^o$.

Proposition 1. The equilibrium service price and the consequential service market coverage satisfy

$$p_s^* = \frac{V(q)}{2}, D_s^* = \frac{D}{2}, p_s^{o*} = \frac{V(Q)}{2}, D_s^{o*} = \frac{D^o}{2}$$

for PRS and OSS correspondingly.

Clearly, service providers for PRS and OSS are monopoly in their individual service market; hence each of them covers half of the service market and prices service at half of the service value to obtain monopolistic service profit.

Stage 2. Incorporating the above results into stage 2, I derive consumers' expected net gain of using PRS service and OSS service respectively.

$$A = \int_{\frac{p_s^*}{V(q)}}^{1} (\theta' V(q) - p_s^*) d\theta' = \frac{V(q)}{8}, B = \int_{\frac{p_s^{o*}}{V(Q)}}^{1} (\theta' V(Q) - p_s^{o*}) d\theta' = \frac{V(Q)}{8}$$

Consumers' expected utility of choosing PRS is $\theta q - p + A$. θ represents consumers' type about software features and is uniformly distributed over the interval [0, 1]. $\theta q - p$ represents utility from consuming the software and A represents expected utility of adopting PRS service. Consumers' expected utility of adopting OSS is $\theta Q + B$. Let $q > Q$[55], demand is

[55] PRVs have initial market power in majority of software markets (e.g., Windows and Linux; Oracle DBMS and MySQL DBMS), hence their PRS has higher market share and more software features than OSS initially. Following this market situation, I assume $Q < q$. Besides, this assumption is consistent with the situation in the DBMS market. MySQL provides a basic, but good-enough product to users who do not need all the features of

$$D = 1 - \frac{p - \frac{1}{8}(V(q) - V(Q))}{q - Q}, D^o = \frac{p - \frac{1}{8}(V(q) - V(Q))}{q - Q}$$

for PRS and OSS correspondingly. PRV's payoff at stage 2 is π = π (software) + π (service) where

$$\pi(software) = pD, \pi(service) = p_s^* D_s^* = \frac{V(q)}{2}\frac{D}{2}.$$

PRV's problem at stage 2 is to

$$\max_p \pi = pD + \frac{V(q)}{2}\frac{D}{2}$$

Proposition 2. The equilibrium PRS price and the corresponding software market share satisfy

$$p^* = \frac{1}{2}(q - Q - \frac{1}{8}V(q) - \frac{1}{8}V(Q)), D^* = \frac{1}{2} + \frac{\frac{3}{16}V(q) - \frac{1}{16}V(Q)}{q - Q}, D^{o^*} = \frac{1}{2} - \frac{\frac{3}{16}V(q) - \frac{1}{16}V(Q)}{q - Q}.$$

Proposition 2 shows that given the setup of this model, PRV gains larger market share than OSS in software market characterized by high demand for service as long as both the PRS and the OSS stay in the competition.

Stage 1. PRV's objective at stage 1 is to

$$\max_q \pi = p^* D^* + \frac{V(q)}{2}\frac{D^*}{2} - C(q)$$

Oracle database (Wittig and Inkinen, 2004). If $q \le Q$, PRV will have zero demand. The condition will be checked later.

where p^* and D^* are determined as in proposition 2. To make the problem tractable, following Thatcher and Pingry (2004), I assume $C(q) = aq^2$ and let $V(q) = bq$ and $V(Q) = bQ$, where $a > 0$, and $b \geq 0$.

Proposition 3. If $b \in (b_1, b_2)$ where b_1 and b_2 are the roots of a quadratic function $(8+3b)^2 - 64abQ - 512aQ = 0$, PRV's payoff function is quasi-concave. The equilibrium number of features bundled in PRS satisfies

$$\frac{1}{2}\left(\frac{3b}{16(q^*-Q)} - \frac{b(3q^*-Q)}{16(q^*-Q)^2}\right)\left(q^* - Q - \frac{bQ}{8} + \frac{3bq^*}{8}\right) + \frac{1}{2}\left(\frac{1}{2} + \frac{b(3q^*-Q)}{16(q^*-Q)}\right)\left(1 + \frac{3b}{8}\right) - 2aq^* = 0$$

The sufficient condition for $q^* > Q$ is $Q < 1/(8a)$. Substituting q^* into the expressions of p^* and D^* specified in proposition 2 and into the expressions of p_s^* and p_s^{o*} specified in proposition 1, I derive the equilibrium PRS price and market share, and service prices for PRS and OSS.

4.5 Findings and Managerial Implications

Comparative analysis of the equilibrium is presented in a two-by-two framework in Table 4.2. Table 4.2 reports how the impact of OSS differs on software market characterized

Table 4.2 Comparative Analysis

	$Q = 0$ (No OSS)	$Q > 0$ (With OSS)
$b = 0$, software market with low demand for service	$q^* = \frac{1}{8a}, p^* = \frac{q^*}{2} = \frac{1}{16a}, D^* = \frac{1}{2}$	$q^* = \frac{1}{8a}, p^* = \frac{q^* - Q}{2} = \frac{1}{16a} - \frac{Q}{2}, D^* = \frac{1}{2}$
$b > 0$, software market with high demand for service	$q^* = \frac{1}{2a}(\frac{1}{2} + \frac{3b}{16})^2$ $p^* = \frac{q^*}{2} - \frac{bq^*}{16}, D^* = \frac{1}{2} + \frac{3b}{16}$	q^* as in proposition 3 $p^* = \frac{q^*}{2} - \frac{bq^*}{16} - \frac{Q}{2} - \frac{bQ}{16}, D^* = \frac{1}{2} + \frac{(3q^* - Q)b}{16(q^* - Q)}$

by low demand for service versus on software market characterized by high demand for service. The findings are summarized in the following propositions.

Proposition 4. In software market characterized by low demand for service, when $Q < 1/8a$,

a. OSS does not have an effect on the optimal number of features developed in competing PRS. The optimal level of features in PRS remains at the threshold value *1/8a*;

b. PRS is priced lower after the entry of competing OSS;

c. PRS remains the same market share after the entry of competing OSS.

The results reported in Proposition 4 are consistent with findings in prior literature that examines software market with low demand for service. For example, Bitzer and Schroder (2007) built a simple economic model to examine the impact of increased competition due to OSS entry on innovation activity of both incumbents and entrants. They found that entry of OSS promotes innovation, that is, enterprises choose to increase the technological levels after the emergence of OSS. Bitzer and Schroder (2006) empirically found support for the proposition in Bitzer and Schroder (2007). Note that although part (a) in Proposition 4 does not show an increase in PRV's investment in software feature after the entry of OSS, it does not conflict with the findings in prior research. This is because the investment decision in my model (Stage 1) refers to service-demand-improving investment in software feature. PRV may invest in other innovation activities in response to the entry of OSS, which may lead to an increase in its overall technological level.

Proposition 5[56]. In software market characterized by high demand for service, when Q is less than a threshold $\underline{q} \in (0, \frac{1}{8a})$,

a. The equilibrium number of features in PRS satisfies $Q < q^* < \frac{1}{2a}(\frac{1}{2} + \frac{3b}{16})^2$ where the upper bound is the equilibrium number of features in PRS when there is no OSS; PRV lowers the equilibrium number of features as OSS includes more features. When Q exceeds the threshold $\underline{q}$ but is less than *1/8a*, the equilibrium number of features in PRS is at the highest level $\bar{q}$. When Q exceeds *1/8a*, PRV is driven out of the market.
b. PRS is priced lower, but with larger market share when there is OSS than when there is no OSS.

Part (a) of proposition 5 implies a *magnetic phenomenon* in the design of PRS with high demand for service that is not observed in the design of software with low demand for service. That is, when the number of features in OSS is below a threshold value, the optimal number of features in PRS with high demand for service is higher than the num-

[56] Proof of proposition 5 is based on the right continuity property of PRV's payoff function at *Q=0* and by taking derivative of first and second order condition of the payoff function at optimal *q* with respect to *Q*. Complete proofs of all propositions are available upon request.

ber of features in the competing OSS; the higher the number of features in the OSS, the lower the number of features in the PRS. This situation resembles one in which the position of RV is *attracted* to that of OSS. If the position of OSS exceeds the threshold, the design of PRS is placed at the maximal distance away from the position of OSS as if the two *repelled* one another, that is, $q = q^*$.

This phenomenon can be explained by the leverage of market share effect and strategic effect discussed in the literature of product differentiation (Tirole, 1988). According to Tirole (1988), strategic effect and market share effect represent competing firms' product choice. The strategic effect dominates the market share effect when firms target their products at a maximal difference to avert intense competition; the market share effect dominates the strategic effect when firms target their products more closely to obtain a larger market share. It has been shown that in general, the strategic effect dominates the market share effect; hence firms tend to adopt the strategy of maximal difference. My findings suggest that in software market characterized by high demand for service, PRV's software design strategy depends on the features in competing OSS: there is a threshold in the number of features in OSS, below which the market share effect dominates, and above which the strategic effect dominates. However, in software markets with low demand for service, the strategic effect always dominates the market share effect.

The intuition of PRV's choice between market share effect and strategic effect is straightforward. PRVs have three options when determining number of features in response to OS entrant: decrease, increase or remain number of features in the software. For PRS with high demand for service, when the number of features in OSS is below a threshold, decreasing number of features in the PRS will decrease software price and customers' valuation for service, but the loss of margin is compensated by larger demand for software as well as for service. So the market share effect dominates the strategic effect and PRV's optimal response is to lower number of features. When OSS is positioned closer to PRS, the loss of margin in software and service exceeds the benefit of market share effect. It is optimal for PRV to play the high-end game, that is, only focus on customers that value large number of features. Hence the strategic effect dominates the market share effect. However, this strategy does not work for PRS with low demand for service. For this software category, the benefit of market share effect is not strong enough to cover the loss of margin in software license because the margin in service is zero. As a result, PRV's optimal response is to remain the same number of features.

An important implication of Proposition 5 is that in software markets characterized by high demand for service, when OSS is not very sophisticated PRV could introduce more basic version of their software in response to the growth of OS counterparts. This implication is very much inline with industry practice. In the DBMS market, IBM, Oracle and Microsoft "have each lowered their database prices and created low-end bundles aimed at smaller organizations and partners." (LaMonica, 2005a) Oracle released a low-end, free but proprietary edition of its database in October 2005, supposedly as "a reaction to the growing competitive pressure from low-end open source databases." (LaMonica, 2005b) Andrew Mendelsohn, senior vice president of Oracle's server technologies division, commented on this move: "There is definitely a market there (for low-end databases) and a demand. And we want them to be using Oracle and not MySQL or SQL Server Express." (LaMonica, 2005b)

The strategic handling of market share effect and strategic effect can also be used to explain PRVs OS strategies. Two common OS strategies employed by PRVs are: subsidize OS projects by having their employees working on these projects or making their PRS compatible with OSS (e.g., IBM, HP, SAP, Oracle, Siemens); and submit their proprietary products to the OS community (e.g., IBM submitted its database Cloudscape to Apache; BEA, Computer Associates and Sun also donated their PRS to OS community). Researchers in general believe these strategies provide PRVs with the strategic advantage to hurt rivals or the opportunity to profit through complementary products or services. My finding shows that market share may be additional benefit that drives PRV's behavior. Both releasing codes and making PRS compatible with OSS help PRVs gain larger market share by attracting users of OS products. For example, by making its DBMS compatible with Linux, Oracle effectively attracts Linux users to their products run on Linux.

The finding of PRV's pricing strategy (part b in Proposition 5) is also consistent with industry practice. For example, BEA Systems reported in its annual report of 2007 that "competitive pressures and open source availability of functionally competitive software could require us to reduce the price of our products and related services." Some analysts believe that MySQL gave Oracle's corporate customers more price bargaining power than in the past despite the lack of functionality of MySQL compared with Oracle database (Wittig and Inkinen, 2004).

The results presented in Propositions 4 and 5 allow me to compare the impact of OSS on market structure of two categories of software markets: software with high service demand and software with low service demand. The finding implies that OSS will be stronger in terms of market share in markets with low demand for service than in market with high demand for service; PRV gains larger market share in software market characterized by high service demand than in market with low service demand. These hold true so long as OSS is not very advanced. The results are consistent with Economides and Katsamakas (2006a). They found that a vertically integrated proprietary system (proprietary platform and applications to the platform are provided by one firm) has a larger market share than OS system if there is a relatively large demand for the proprietary application, a large transition cost for the open source platform and a relatively small demand for application of the open source platform.

The intuition is that in software markets with high demand for service, PRV has the advantage of manipulating software design and optimizing pricing for both software and service. OS service provider however does not have as much influence on the design of OSS as the influence PRV has on its own software. This disadvantage of OS support provider leads to smaller market share of OSS in markets with high demand for service than in markets with low demand for service. To certain extent this resembles the case of an integrated system (proprietary system including software and service) versus a disintegrated system (OS system). Interestingly although MySQL CEO Marten Mickos believed that eventually Oracle, Microsoft, and IBM would realized that MySQL could be a threat to their database business, my model shows that if MySQL does not have full control on the development of MySQL database, and is not able to align its development objective with the objective of MySQL database OS developers, it will not dominate the market of DBMS.

Proposition 6[57]. PRS with high demand for service has more features and larger market coverage than PRS with low demand for service regardless of whether there is OSS.

Proposition 6 is intuitive. In software market with high demand for service, PRV can derive higher margin selling service by increasing the number of features. The superiority of this type of PRS in terms of features helps increase demand for corresponding service, which in turn improves the expected value consumers can derive from purchasing the software. Naturally, market coverage for PRS with high demand for service is higher.

Proposition 7[58].

a. OSS hurts PRV's payoff in both types of software markets. When OSS is sufficiently advanced, PRV is driven out of the market.

b. PRV of software with high demand for service has higher payoff than PRV of software with low demand for service regardless of whether there is OSS.

Proposition 7 is very intuitive. Part (a) implies that competing with OSS in the same market is ultimately a bad strategy for PRV. In the long run, PRV has to find a way to cooperate with OS community.

4.6 Conclusions

---

[57] Proof of proposition 6 is straightforward, hence is omitted.
[58] Proposition 7 can be proved by studying the direct effect of Q and b on PRV's optimal payoff.

OSS has become prominent in many categories of software. I investigate how the rise of OSS influence software market structure and PRV's software design strategy. In particular, I examine how the impacts of OSS differ between software market characterized by high demand for service and by low demand for service. I found that when there is competition between OSS and PRS, PRS is priced lower because of the entry of OSS. OSS have larger market share in markets with low demand for service than in markets with high demand for service. PRV's software design strategy in response to OSS differs between software with high demand for service and software with low demand for service: PRV's software design demonstrates a *magnetic* phenomenon for software with high demand for service, the purpose of which is to increase software market share, which further promotes more revenue in service; this magnetic phenomenon is not found in software with low demand for service.

APPENDIX A

Interaction between R&D Intensity of MNEs and Host Country Firms

Combining R&D expenditure and output certainly makes the analysis more complicated. It can be established that R&D activity of MNEs in host country lowers domestic R&D intensity if and only if:

$\frac{2(B(bR - fIK) + fGIJ) + ((bR - fIK)((1-\alpha)/m + E) + fGIL)X_D}{L} < 0$. Whether this condition satisfies crucially depends on domestic spillover coefficient α_F. α_F represents the ease for innovating domestic firms to absorb the R&D spillover of MNEs. Figure A.1 plots the change in domestic R&D intensity in response to R&D activity of MNEs in host country as a function of α_F for various levels of MNE spillover coefficient γ_D. γ_D represents the ease for innovating MNEs to absorb the R&D spillover of domestic firms. If it is easy for innovating domestic firms to learn from MNEs, the R&D intensity of domestic firms increases with the R&D intensity of MNEs. If it is difficult for innovating domestic firms to learn, the R&D intensity of domestic firms decreases with the R&D intensity of MNEs.

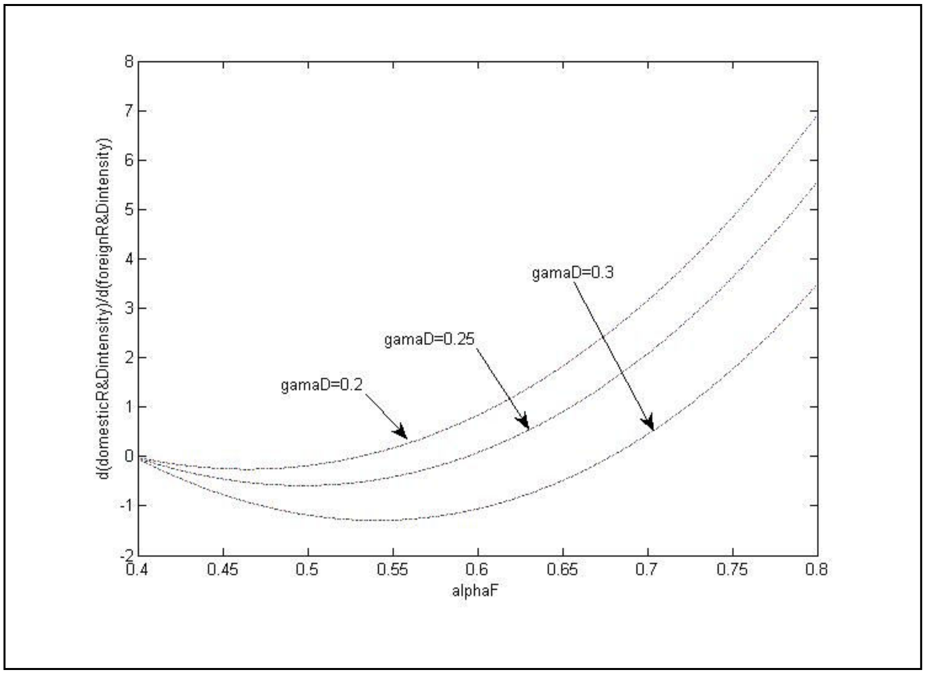

Figure A.1: Impact of Ease of Learning on the Relation between Foreign R&D and Domestic R&D

APPENDIX B

List of Neighboring Provinces

Table B.1: List of Neighboring Provinces

Province	Neighboring Provinces
Beijing	Tianjin, Hebei
Tianjin	Beijing, Hebei
Hebei	Beijing, Tianjin, Shanxi, Inner Mongolia, Liaoning, Shandong, Henan
Shanxi	Hebei, Inner Mongolia, Henan, Shaanxi
Inner Mongolia	Hebei, Shanxi, Liaoning, Jilin, Heilongjiang, Shaaxi, Gansu, Ningxia
Liaoning	Hebei, Inner Mongolia, Jilin
Jilin	Inner Mongolia, Liaoning, Heilongjiang
Heilongjiang	Inner Mongolia, Jilin
Shanghai	Jiangsu, Zhejiang
Jiangsu	Shanghai, Zhejiang, Anhui, Shandong
Zhejiang	Shanghai, Jiangsu, Anhui, Fujian, Jiangxi
Anhui	Jiangsu, Zhejiang, Jiangxi, Henan, Hubei
Fujian	Zhejiang, Jiangxi, Guangdong
Jiangxi	Zhejiang, Anhui, Fujian, Hubei, Hunan, Guangdong
Shandong	Hebei, Jiangsu, Henan
Henan	Hebei, Shanxi, Anhui, Shandong, Hubei, Shaanxi
Hubei	Anhui, Jiangxi, Henan, Hunan, Chongqing, Shaanxi
Hunan	Jiangxi, Hubei, Guangdong, Guangxi, Chongqing, Guizhou
Guangdong	Fujian, Jiangxi, Hunan, Guangxi, Hainan
Guangxi	Hunan, Guangdong, Hainan, Guizhou, Yunnan
Hainan	Guangdong, Guangxi
Chongqing	Hubei, Hunan, Sichuan, Guizhou, Shaanxi
Sichuan	Chongqing, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai
Guizhou	Hunan, Guangxi, Chongqing, Sichuan, Yunnan
Yunnan	Guangxi, Sichuan, Guizhou, Tibet
Gansu	Inner Mongolia, Sichuan, Shaanxi, Qinghai, Ningxia, Xinjiang
Qinghai	Sichuan, Tibet, Gansu, Xinjiang
Ningxia	Inner Mongolia, Shaanxi, Gansu
Xinjiang	Tibet, Gansu, Qinghai

REFERENCES

Aardal, K., Pochet, Y., and Wolsey, L. 1995. Capacitated facility location: valid inequalities and facets. Mathematics of Operations Research (20:3), 562–582.

Abreu, B.F., Melo, W.L. 1996. Evaluating the Impact of Object-Oriented Design on Software Quality. Proc. Third Int'l Software Metrics Symp.

Adams, J.D., 1990. Fundamental stocks of knowledge and productivity growth. Journal of Political Economy (98:4), 673-702.

Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. Econometrica (60:2), 323-351.

Aitken, B.J., Harrison, A.E., 1999. Do domestic firms benefit from direct foreign investment? Evidence from Venezuela. The American Economic Review (89:3), 605-618.

Alvarez, I., Molero, J., 2005. Technology and the generation of international knowledge spillovers: An application to Spanish manufacturing firms. Research Policy (34:9), 1440-1452.

Ambos, B., 2005. Foreign direct investment in industrial research and development: A study of German MNCs. Research Policy (34:4), 395-410.

Anonymous 2006. American hotel & lodging association. http://www.ahla.com/index.asp.

Arnold, D.J., Quelch, J.A., 1998. New strategies in emerging economies. Sloan Management Review (40:1), 7-20.

Arntzen, B., Brown, G., Harrison, T., and Trafton, L. 1995. Global supply chain management at digital equipment corporation. Interfaces (25:1), 69–93.

Baker, M., Boehmer, J., and Davis, C. 2007. Domestic daily costs near top of sellers cycle. Business Travel News (24:3), 4–28.

Bank, D. Open Source Database Poses Oracle Threat. Wall Street Journal. (Eastern Edition). New York, N.Y.Jul 9, 2003, B1.

Banker, R., Davis, G., Slaughter, S. 1998. Software development practices, software complexity, and software maintenance performance: a field study. Management Science (44:4), 433-450.

Banker, R., Slaughter, S. 1997. A field study of scale economies in software maintenance. Management Science (43:12), 1709-1725.

Bardhan, A., Jaffee, D., 2005. Innovation, R&D and offshoring. Working Paper, University of California Berkeley.

Bardhan, A., 2006. Managing globalization of R&D: Organizing for offshoring innovation. Working Paper, F-602, Haas School of Business, University of California, Berkeley.

Basant, R., Fikkert, B., 1996. The effects of R&D, foreign technology purchase, and domestic and international spillovers on productivity in Indian firms. The Review of Economics and Statistics (78:2), 187-199.

Belderbos, R., Lykogianni, E, Veugelers, R. 2005. Strategic R&D location by multinational firms: spillovers, technology sourcing, and competition. Working paper, EARIE Conference.

Bessen, J. 2006. Open source software: free provision of complex public goods. The Economics of Open Source Software Development, 57-81.

Bitzer, J. 2004. Commercial versus open source software: the role of product heterogeneity in competition. Economic Systems (28), 369-381.

Bitzer, J., Schroder, P. 2006. The impact of entry and competition by open source software on innovation activity. The economics of open source software development, 219-246.

Blomstrom, M., Kokko, A., 1998. Multinational corporations and spillovers. Journal of Economic Survey (12:2), 1-31.

Blomstrom, M., Kokko, A., 2001. Foreign direct investment and spillovers of technology. International Journal of Technology Management (22:5/6), 435-454.

Boyd, E. A. and Bilegan, I. C. 2003. Revenue management and e-commerce. Management Science (49:10), 1363–1386.

Briand, L.C., Morasca, S., Basili, V.C. 1999. Defining and Validating Measures for Object-Based High-Level Design. IEEE Trans. Software Eng. (25:5), 722-743.

Cachon, G. and Zhang, F. 2006. Procuring fast delivery: Sole sourcing with information asymmetry. Management Science (52:6), 881–896.

Chaudhry, S., Forst, F., and Zydiak, J. 1993. Vendor selection with price breaks. European Journal of Operational Research (70), 52–66.

Chen, R., Roundy, R., Zhang, R., and Janakiraman, G. 2005. Efficient auction mechanisms for supply chain procurement. Management Science (51:3), 467–482.

Chung, W., Alcacer, J., 2002. Knowledge seeking and location choice of foreign direct investment in the United State. Management Science (48) 1534-1554.

Cohen, A. 2000. Finding balance can sometimes be tricky. Financial Times. 10 February 2000 p. 3.

Cohen, A. 2004. BTN data benchmarking CT100: show flat 2003, rising 2004. Business Travel News (21:12), 10–12.

Cohen, M. and Lee, H. 1988. Strategic Analysis of Integrated Production-Distribution Systems: Models and Methods. Operations Research (36:2), 216–228.

Cohen, M., Whang, S. 1997. Competing in product and service: a product life-cycle model. Management Science (43:4), 535-545.

Cohen, W., Klepper, S., 1992. The anatomy of industry R&D intensity distributions. The American Economic Review (82:4), 773-799.

Cohen, W.M., Levin, R.C., Mowery, D.C., 1987. Firm Size and R & D Intensity: A Re-Examination. The Journal of Industrial Economics (35:4), 543-565.

Corbett, C. and de Groote, X. 2000. A Supplier's Optimal Quantity Discount Policy Under Asymmetric Information. Management Science (46:3), 444–450.

Crespo, N., Fontoura, M.P., 2007. Determinant factors of FDI spillovers - What do we really know? World Development (35:3), 410-425.

D'Aspremont, C., Jacquemin, A., 1988. Cooperative and non cooperative R&D in duopoly with spillovers. The American Economic Review (78:5), 1133-1137.

Dahlander, L., Magnusson, M. 2006. Business models and community relationships of open source software firms. The Economics of Open Source Software Development, 111-130.

Daniel, F., Lohrke, F.T., Fornaciari, C.J., Turner, R.A., 2004. Slack resources and firm performance: a meta-analysis. Journal of Business Research (57:6), 565-574.

Davis, C. 2001. CT100 summit highlights. Business Travel News (18:18), 8–16.

Davis, C. 2002. The corporate travel 100: benchmarks severe change. Business Travel News (19:16), 8–16.

Degraeve, Z., Labro, E., and Roodhooft, F. 2000. An evaluation of vendor selection models from a total cost of ownership perspective. European Journal of Operational Research (125:1), 34–58.

Degraeve, Z., Labro, E., and Roodhooft, F. 2004. Total cost of ownership purchasing of a service: the case of airline selection at alcatel bell. European Journal of Operational Research (156:1), 23–40.

Del Canto, J.G., Gonzalez, I.S., 1999. A resource-based analysis of the factors determining a firm's R&D activities. Research Policy (28:8), 891-905.

Deng, S. and Elmaghraby, W. 2005. Supplier Selection via Tournaments. Production and Operations Management (14:2), 252–267.

Driffield, N., 2001. The impact on domestic productivity of inward investment in the UK. The Manchester School (69:1), 103-17.

Economides, N., Katsamakas, E. 2006a. Two-sided competition of proprietary vs. open source technology platforms and the implications for the software industry. Management Science (52:7), 1057-1071.

Economides, N., Katsamakas, E. 2006b. Linux vs. Windows: a comparison of application and platform innovation incentives for open source and proprietary software platforms. The Economics of Open Source Software Development, 207-218.

Elmaghraby, W. 2000. Supply Contract Competition and Sourcing Policies. Manufacturing & Service Operations Management (2:4), 350–371.

Elmaghraby, W. and Keskinocak, P. 2003. Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions. Management Science (49:10), 1287–1309.

Engardio, P. R&D offshoring: is it working? Businessweek Online, May 10, 2006.

Fung, K., 2005. Trade and Investment among China, the United States, and the Asia-Pacific Economies: An Invited Testimony to the U.S. Congressional Commission, China as an Emerging Regional and Technological Power: Implications for U.S. Economic and Security Interests, Hearing before the U.S.-China Economic and Security Review Commission, One Hundred Eighth Congress, Second Session, U.S. Government Printing Office, Washington, D.C., 2004, p.30-36. Available from :< http://www.uscc.gov/hearings/2004heari.

Gaudeul, A. Competition between open-source and proprietary software: the LaTex case study. Working paper, 2005. University of East Anglia – Norwich and ESRC Center for Competition Policy.

Gaudeul, A. 2007. Do open source developers respond to competition? The LaTex case study. Review of Network Economics (6:2), 239-263.

Gersbach, H., Schmutzler, A., 2006. Foreign direct investment and R&D offshoring. Working Paper, University of Zurich.

Greve, H.R., 2003. A behavioral theory of R&D expenditures and innovations: Evidence from shipbuilding. Academy of Management Journal (46:6), 685-702.

Haddad, M., Harrison, A., 1993. Are there positive spillovers from direct foreign investment? : Evidence from panel data for Morocco. Journal of Development Economics (42:1), 51-74.

Haller, S., 2004. The impact of multinational entry on domestic market structure and R&D. Working Paper, European University Institute.

Haruvy, E., Sethi, S., Zhou, J. Open source development with a commercial complementary product or service. 2005. Working paper. University of Texas Dallas.

Hemphill, T., 2005. US offshore outsourcing of R&D: accommodating firm and national competitiveness perspectives. Innovation: management, policy & practice (7:4), 351-356.

Hoskisson, R.E., Eden, L., Lau, C.M., Wright, M., 2000. Strategy in Emerging Economies. The Academy of Management Journal (43:3), 249-267.

Inkmann, J., Pohlmeier, W., 1995. R&D Spillovers, technological distance and innovation success. Discussion Paper, University of Konstanz.

Jaffe, A.B., 1986. Technological opportunity and spillovers of R & D: Evidence from firms' patents, profits, and market value. The American Economic Review (76:5), 984-1001.

Javorcik, B.S., 2004. Does foreign direct investment increase the productivity of domestic firms? In search of spillovers through backward linkages. The American Economic Review (94:3), 605-627.

Jones, C., 1995. R&D based model of economic growth. Journal of Political Economy (103: 4), 759-784.

Kaiser, U., 2002. Measuring knowledge spillovers in manufacturing and services: An empirical assessment of alternative approaches. Research Policy (31:1), 125-144.

Karabuk, S., Wu, D. 2003. Coordinating strategic capacity planning in the semiconductor industry. Operations Research (51:6), 839-849.

Katsuno, M., 2005. Status and overview of official ICT indicators for China. OECD Science, Technology and Industry Working Paper, 2005/4.

Kerstetter, J. Will IBM's Gift Keep on Giving? Business Week Online, 8/4/2004.

Kokko, A., 1994. Technology, markets characteristics, and spillovers. Journal of Development Economics (43:2), 279-293.

Kouvelis, P., Chambers, C., and Wang, H. 2006. Supply chain management research and production and operations management: Review, trends, and opportunities. Production and Operations Management (15:3), 449–469.

Kuemmerle, W., 1999. The drivers of foreign direct investment into research and development: an empirical investigation. Journal of International Business Studies (30), 1-24.

Kumar, M., Vrat, P., and Shankar, R. 2006. A fuzzy programming approach for vendor selection problem in a supply chain. International Journal of Production Economics (101), 273–285.

Kyriazidou, E., 1997. Estimation of a panel data sample selection model. Econometrica (65:6), 1335-1364.

Labbe, M. and Yaman, H. 2006. Polyhedral Analysis for Concentrator Location Problems. Computational Optimization and Applications (34:3), 377–407.

Labbe, M., Yaman, H., and Gourdin, E. 2005. A branch and cut algorithm for hub location problems with single assignment. Mathematical Programming (102:2), 371–405.

Lacy, S. Open Source: Now It's an Ecosystem. Business Week Online; 10/3/2005.

Laderman, D., Maloney, W., Serven, L., 2003. Lessons from NAFTA from Latin American and Caribbean countries. Washington, D.C.: World Bank.

LaMonica, M. Oracle buys open-source database firm. News.com, October 10, 2005a.

LaMonica, M. Oracle to offer free database. News.com, October 28, 2005b.

Lawton, C. Business Technology: Linux Shoots for Big League of Servers; Low-Cost Operating System Pushes Into Microsoft, Sun Strongholds. Wall Street Journal. (Eastern Edition). New York, N.Y. Jun 19, 2007, B.5

Le Bas, C., Sierra, C., 2002. Location versus home country advantages in R&D activities: some further results on multinationals' locational strategies. Research Policy (31), 589-609.

Lee, C. Y., 2005. A new perspective on industry R&D and market structure. Journal of Industrial Economics (53:1), 101-122.

Lee, C .Y., Sung, T., 2005. Schumpeter's legacy: A new perspective on the relationship between firm size and R&D. Research Policy (34:6), 914-931.

Lee, D., Mendelson, H. Divide and Conquer: Competing with Free Technology Under Network Effects. Production and Operations Management 2008 (17:1), 12-28.

Lerner, J., Tirole, J. Some simple economics of open source. Journal of Industrial Economics 2002 (L:2), 197-234.

Lerner, J., Tirole, J. The economics of technology sharing: open source and beyond. Journal of Economic Perspectives 2005 (19:2), 99-120.

Li, W., Henry, S., Kafura, D., Schulman, R. 1995. Measuring Object-Oriented Design. J. Object-Oriented Programming (8:4), 48-55.

Lieberman, J., 2004. Offshore outsourcing and American's competitive edge: losing out in the high technology R&D and services sectors. White paper by Office of Senator, Washington D. C., 20510.

Marin, A., Bell, M., 2006. Technology spillovers from foreign direct investment (FDI): an exploration of the active role of MNC subsidiaries in the case of Argentina in the 1990s. The Journal of Development Studies (42:4), 678-697.

McCartney, S. 2006. The middle seat: Exploiting travelers' 'visual guilt'; firms find that employees pinch pennies if they book business trips themselves. Wall Street Journal. 25 July 2006 p. D4.

Meng, Q., Li, M., 2002. New Economy and ICT Development in China. Information Economics and Policy (14:2), 275-295.

Meyer, D. 2005. Corporate travel 100: spending back above $8b. Business Travel News (22:12), 3–54.

Meyer, D. 2006. Corporate travel 100: air spending soars. Business Travel News (23), 3–57.

Mustonen, M. 2005. When does a firm support substitute open source programming? Journal of Economics and Management Strategy (14:1), 121-139.

Narasimhan, S. and Pirkul, H. 1992. Hierarchial concentrator location problem. Computer Communications (15:3), 185–191.

National Bureau of Statistics, 2005. The analysis of high technology product import and export of 2004 (2004 Nian Gao Ke Ji Chan Pin Jin Chu Kou Fen Xi)(in Chinese), Available from :< http://www.sts.org.cn/tjbg/jsmy/documents/2004/0320.htm>

Net Applications. Top Browser Market Share for February, 2008. Retrieved on 2008-03-31.

Netcraft's Survey on web server market share at http://news.netcraft.com/

OECD, 2005. OECD finds that China is biggest exporter of information technology goods in 2004, surpassing US and EU, available from:< http://www.oecd.org/document/8/0,2340,en_2649_201185_35833096_1_1_1_1,00.html>.

Ornaghi, C. 2006. Spillovers in product and process innovation: evidence from manufacturing firms. International Journal of Industrial Organization (24), 349-380.

Pack, H., Saggi, K., 2006. Is there a case for industrial policy? A critical survey. World Bank Research Observation (21:2), 267-297.

Parkes, D. and Kalagnanam, J. 2005. Models for Iterative Multiattribute Procurement Auctions. Management Science (51:3), 435–451.

Petit, M., Sanna-Randaccio, F., 2000. Endogenous R&D and foreign direct investment in international oligopolies. International Journal of Industrial Organization (18), 339-367.

Pirkul, H. 1987. Efficient algorithms for the capacitated concentrator location problem. Computers and Operations Research (14:3), 197–208.

Potterie, B.P., Lichtenberg, F., 2001. Does foreign direct investment transfer technology across borders? The Review of Economics and Statistics (83:3), 490-497.

Poundarikapuram, S. and Veeramani, D. 2004. Distributed Decision-Making in Supply Chains and Private E-Marketplaces. Production and Operations Management (13:1), 111–121.

Riordan, M. and Sappington, D. 1987. Awarding Monopoly Franchises. The American Economic Review (77:3), 375–387.

Romer, P., 1990. Endogenous technological change. Journal of Political Economy (98:5), S71-S102.

Rossi, C., Bonaccorsi, A. 2006. Intrinsic motivations and profit-oriented firms in open source software: do firms practice what they preach? The Economics of Open Source Software Development. 83-107.

Roth, A. and Menor, L. 2003. Insights into service operations management: A research agenda. Production and Operations Management (12:2), 145–164.

Russo, B., Braghin, B., Gasped, P., Sillitti, A., Succi, G. 2005. Defining TCO for the Transition to Open Source Systems, in Proceedings of First International Conference on Open Source (OSS2005), M. Scotto and G. Succi (eds.), Genoa, Italy, July 11- 15, 2005, 108-112.

Saggi, K. 2002. Trade, foreign direct investment, and international technology transfer: a survey. The World Bank Research Observer (17:2), 191-235.

Sanna-Randaccio, F., 2002. The impact of foreign direct investment on home and host countries with endogenous R&D. Review of International Economics (10:2), 278-298.

Schmidt, K., Schnitzer, M. 2003. Public subsidies for open source? Some economic policy issues of the software market. Harvard J of Law and Technology (16:2), 474-502.

Schumpter, J.A., 1950. Capitalism, socialism, and democracy. 3rd ed. Harper & Row, New York.

Semykina, A., Wooldridge, J., 2005. Estimating panel data models in the presence of endogeneity and selection: theory and application. Working paper.

Sen, R. 2007. A strategic analysis of competition between open source and proprietary software. Journal of Management Information Systems (24:1), 233-257.

Serlen, B. 2002. Hotel chain adds discounting tier. Business Travel News (19:20), 1–3.

Serlen, B. 2004. E-agencies offer dynamic room rates. Business Travel News (21:7), 1–2.

Serlen, B. 2005. Setting up a hotel program. Business Travel News (22:7), 59–68.

Seshadri, S., Chatterjee, K., and Lilien, G. 1991. Multiple Source Procurement Competitions. Marketing Science (10:3), 246–263.

Snir, E. and Hitt, L. 2004. Vendor Screening in Information Technology Contracting With a Pilot Project. Journal of Organizational Computing and Electronic Commerce (14:1), 61–88.

Spencer, B., Brander, J., 1983. International R&D rivalry and industrial strategy. Review of Economic Studies (50:163), 707-722.

Stembridge, B., 2007. Eastward Ho!- the Geographic Drift of Global R&D. Knowledge-Link newsletter from Thomson Scientific.

Stone, A. For Novell, Linux Means New Life. Business Week Online, 3/31/2004.

Swanson, E. B., Beath, C. M. Maintaining information systems in organizations. John Wiley and Sons, New York, 1989.

Tam, K. Y. and Hui, K. L. 2001. A Choice Model for the Selection of Computer Vendors and Its Empirical Estimation. Journal of Management Information Systems (17:4), 97–124.

Tam, P. W. H-P Increases Its Support For 'Open Source' Software. Wall Street Journal. (Eastern Edition). New York, N.Y. Jun 1, 2004, B.4.

Tan, J., Peng, M.W., 2004. Organizational slack and firm performance during economic transitions: Two studies from an emerging economy. Strategic Management Journal (25:3), 307-307.

Tan, Y. and Mookerjee, V. 2005. Comparing uniform and flexible policies for software maintenance and replacement. IEEE Transactions on Software Engineering (31:3), 238-255.

Tempelmeier, H. 2002. A simple heuristic for dynamic order sizing and supplier selection with time-varying data. Production and Operations Management (11:4), 499–515.

Thatcher, M., Pingry, D. 2004. An Economic Model of Product Quality and IT Value. Information Systems Research (15:3), 268-287.

The Boston Consulting Group. 2005. Globalizing R&D: building a pathway to profits.

Thursby, J., Thursby, M., 2006. Here or there? A survey on the factors in multinational R&D location, Report to the government-university-industry research roundtable. Washington, DC: National Academies Press.

Tirole, J. The theory of industrial organization, MIT Press, Cambridge, MA, 1988.

Tullous, R. and Utecht, R. 1992. Multiple or single sourcing? Journal Business and Industrial Marketing (7:3), 5–18.

Vara, V. Oracle Takes Aim At Linux Support Sold by Red Hat. Wall Street Journal. (Eastern Edition). New York, N.Y.Oct 26, 2006. B.4.

Veugelers, R., Cassiman B., 1999. Make and buy in innovation strategies: evidence from Belgian manufacturing firms. Research Policy (28), 63-80.

Veugelers, R., Houte, P.V., 1990. Domestic R&D in the presence of multinational enterprises. International Journal of Industrial Organization (8:1), 1-15.

Weber, C. and Current, J. 1993. A multiobjective approach to vendor selection. European Journal of Operational Research (68), 173–184.

Wheeler, D. Why Open Source Software/Free Software (OSS/FS, FLOSS, or FOSS)? Look at the Numbers! November 2005 (available online at http://www.dwheeler.com/oss_fs_why .html).

Wittig, C. 2006. MySQL open source database in 2006 (B). Stanford case SM-124 (B), Graduate School of Business.

Wittig, C., Inkinen, S. 2004. MySQL open source database in 2004. Stanford case SM-124, Graduate School of Business.

Wooldridge, J., 1995. Selection corrections for panel data models under conditional mean independence assumptions. Journal of Econometrics (68), 115-132.

Wu, D. and Kleindorfer, P. 2005. Competitive options, supply contracting, and electronic markets. Management Science (51:3), 452–466.