In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of ACCOUNTING

The University of British Columbia
Vancouver, Canada

Date AUGUST 18, 1995
ABSTRACT

This dissertation develops variants of the well-known Hotelling's location model to examine the nature of competition in the audit market where audit firms make strategic specialization and pricing decisions.

In a multi-period spatial oligopoly model of auditing competition, audit firms obtain market power through their service specialization with respect to client characteristics relevant to audit production. This market power allows audit firms to price discriminate among clients. Competition among audit firms is localized: an audit firm optimally charges a client, to whom it has the lowest auditing cost to serve, the marginal auditing cost of the second lowest-cost audit firm. These equilibrium audit firms' pricing strategies result in an allocation of clients' surplus and audit firms' profits that lies in the core of the economy. The existence of a specialization-pricing equilibrium is also established. In equilibrium, given its rivals' specializations, each audit firm's profit is maximized by choosing a specialization that maximizes the social welfare (the sum of clients' surplus and audit firms' profits). Moreover, audit firms never choose the same specialization in equilibrium. Instead, in order to earn rents as 'local monopolists', audit firms differentiate themselves from each other. This result is consistent with a widely held notion that audit firms search for 'niche' markets, such as industry specialization, to increase their profits.

The dissertation then focuses on a two-period spatial duopoly model in which the market power created by audit firm specialization is now further fortified by the presence of auditors' learning and clients' switching costs. In this case, audit firms optimally price discriminate among clients by offering them 'specialization-and-relationship-specific' audit fee schedules. The practice of 'low-balling' is found to be a natural consequence of the competition among audit firms. However, low-balling occurs only in a certain market segment where audit firms compete quite fiercely. The analysis also demonstrates how equilibrium audit fee
schedules, audit firms' specializations and profits, clients' surplus, and social welfare depend on the auditing costs, the learning rate, and the switching costs. Some interesting policy implications are illustrated. Finally, the model is used to analyze the impact of banning audit firms from the practice of low-balling. It is demonstrated that even though a policy of banning low-balling always reduces competition, it improves social efficiency in some cases.
## TABLE OF CONTENTS

Abstract  
Table of Contents  
List of Figures  
Acknowledgement  

Chapter 1 Introduction  

Chapter 2 A Multi-Period Spatial Model of Auditing Competition  
  2.1 A Brief Literature Review of Spatial Competition  
  2.2 The Model  
  2.3 The Pricing Equilibrium  
  2.4 Specialization Equilibria  
  2.5 Concluding Remarks  

Chapter 3 A Two-Period Model of Auditing Competition  
  with Learning and Switching Costs  
  3.1 The Model  
  3.2 Analysis of Pricing Subgames  
    3.2.1 The Second-Period Price Equilibrium  
    3.2.2 The First-Period Price Equilibrium  
  3.3 Equilibrium for the Full Game  
  3.4 Implications of Changes in Auditing Costs, Learning Rate, and Switching Costs  
  3.5 Concluding Remarks  

Chapter 4 'Low-balling' and Efficiency  
  4.1 Low-balling  
  4.2 Welfare Implications of Low-Balling  
    4.2.1 Equilibrium Outcomes without Low-Balling  
    4.2.2 Welfare Comparison  

iv
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3 Concluding Remarks</td>
<td>74</td>
</tr>
<tr>
<td>Chapter 5 Conclusion</td>
<td>76</td>
</tr>
<tr>
<td>References</td>
<td>82</td>
</tr>
<tr>
<td>Appendix A The Potential Benefits of External Auditing:</td>
<td>89</td>
</tr>
<tr>
<td>An Information Economic Analysis</td>
<td></td>
</tr>
<tr>
<td>A.1 The Basic Borrower-Lender Model without Auditing</td>
<td>92</td>
</tr>
<tr>
<td>A.1.1 The Model</td>
<td>92</td>
</tr>
<tr>
<td>A.1.2 Characterizing the Equilibrium</td>
<td>96</td>
</tr>
<tr>
<td>A.2 The Setting with Auditing</td>
<td>100</td>
</tr>
<tr>
<td>A.3 Concluding Remarks</td>
<td>112</td>
</tr>
<tr>
<td>Appendix B Proofs of Propositions</td>
<td>113</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Chapter 3

Figure 1 The Second-Period Equilibrium Audit Fee Schedule 63
Figure 2 The First-Period Equilibrium Audit Fee Schedule 64
ACKNOWLEDGEMENT

I wish to express my most sincere appreciation to the members of my dissertation committee, Professor Dan Simunic and Professor Mukesh Eswaran, for their excellent assistance and, especially, to my dissertation supervisor, Professor Gerald Feltham, for his encouragement and professional guidance. I would also like to thank seminar participants at Duke, Hong Kong University of Science and Technology, Maryland, Pennsylvania State, UBC, Waterloo and Yale for helpful comments. Professor Jack Hughes at Duke University also deserves special thanks for providing early encouragement. Last, but not least, I wish to thank all the faculty members and Ph.D. students at UBC for providing intellectual stimulation, constant encouragement and friendship. Financial support from the MacPhee Graduate Fellowship and the University Graduate Fellowship is also gratefully acknowledged.
Chapter 1

Introduction

Prior empirical studies have suggested that significant variation exists in accounting and auditing practices across industries.\textsuperscript{1} This is consistent with the conventional wisdom that audit firms invest in specialized resources, such as SEC reporting expertise, taxation advice, computer audit and management consultation, to yield economies of scale and scope for services rendered to particular market segments.\textsuperscript{2} This dissertation proposes that service specialization of audit firms is not only the result of an adaptation by audit firms to particular technological or institutional conditions but also reflects strategic positioning of audit firms in the market. Through specialization, audit firms are able to create their own market niches in which they possess some monopoly power and generate economic rents. The competitive forces in the market then induce audit firms to achieve (constrained) efficient utilization of specialized resources. Hence, this strategic scenario suggests the importance of service specialization considerations in modelling auditing competition. However, even though the nature of competition within the public accounting profession has received increasing attention from researchers and practitioners in recent years, there has been no formal model of

\begin{itemize}
\item \textsuperscript{1}As stated in Danos and Eichenseher (1982): "The factors of production used in producing audit services are diverse in nature. Some can be used across all audit engagements, while others are unique to specific client industries. Moreover, the importance of industry-specific factors tends to vary with the complexity of accounting and auditing rules unique to the client's industry (p. 606)."
\item \textsuperscript{2}For example, Dopuch and Simunic (1980), Danos and Eichenseher (1982), Eichenseher (1985), Rhode et al. (1974) and Schiff and Fried (1976) find evidence of industry specialization by the then Big Eight audit firms.
\end{itemize}
audit firm specialization.³

Extant research on auditing competition considers the audit market as an *ex-ante* perfectly competitive market and has primarily focused on the pricing behaviour of audit firms. Especially, the issue of ‘low-balling’, i.e., setting audit fees below total current costs on an initial audit engagement, has been singled out as a very important research topic and received considerable attention. This attention not only comes from academics (e.g., Coate and Loeb (1994), DeAngelo (1981a, 1981b), Dye (1991), Gigler and Penno (1995), Kanodia and Mukherji (1994), Magee and Tseng (1990), and Simunic (1980)), but also from the profession itself (e.g., The Cohen Commission Report (1978)). The interest in low-balling stems from the hypothesized link to impaired auditor independence.

Based on a multi-period contestable market model of auditing, DeAngelo (1981a) suggests that low-balling occurs if there are rents to be earned by audit firms. In her model, switching audit firms is costly to a client, which in turn, provides a quasi-rent to an incumbent audit firm on future audit engagements with the client. But since the market for auditing services is *ex-ante* perfectly competitive, competition among audit firms for the right to become the incumbent in the initial engagement drives the total quasi-rent for the audit firm to zero, implying below-cost initial audit fees. Thus, DeAngelo concludes that the existence of client-specific quasi-rents to incumbent audit firms both lowers the amount of auditor independence and leads to low-balling in the initial period. Magee and Tseng (1990) basically concur with DeAngelo’s conclusion after modelling the audit pricing game by a dynamic programming approach.⁴ However, the effects of quasi-rents on the auditor’s independence are less significant in the Magee and Tseng’s framework than in DeAngelo’s. DeAngelo asserts that the existence of quasi-rents is a necessary condition for reducing au-

³The study of auditing competition was stimulated first by the Metcalf Staff Report (U.S. Senate 1976) and later by the Cohen Commission Report (AICPA 1978). The former argues that there is insufficient competition, whereas the latter believes it to be excessive.

⁴To be more correct, Magee and Tseng (1990) look at ‘price-cutting’ rather than low-balling, where the former is defined by them as the difference between the second- and first-period audit fees. While low-balling implies price-cutting, the converse is not true.
ditor independence. However, she does not describe how and to what extent the quasi-rents would affect the auditor’s independence. Magee and Tseng extend DeAngelo’s model in a game theoretic setting and obtain five necessary conditions under which a client-specific rent may lead to a compromise of auditor independence. They further argue that those conditions are usually not satisfied. Among other things, Magee and Tseng point out that when audit firms possess all the bargaining power and there is no disagreement among audit firms on the proper interpretation of generally accepted accounting principles (GAAP), clients have nothing to gain by threatening termination of incumbent audit firms and there is no impairment of auditor independence. Thus, Magee and Tseng conclude that there is little pressure on the auditor’s independence despite the existence of client-specific quasi-rents.

Dye (1991) argues that DeAngelo (1981a) implies a division of bargaining power (in terms of audit fees determination) favouring the audit firm; and low-balling would not exist in the absence of the audit firms’ bargaining power over their clients.\(^5\) To provide an explanation for the existence of low-balling, Dye turns to a model with asymmetric information between the audit firm, client and financial statement users. He shows that low-balling is induced by the auditor selection mechanism designed by the client, who has all the bargaining power. Similar to Dye (1991), Kanodia and Mukherji (1994) also assume that the client has almost all the bargaining power. In addition, they assume that the incumbent audit firm has an informational advantage on the auditing costs in subsequent periods following the initial engagement. Kanodia and Mukherji then use contract theory to derive an equilibrium in which low-balling occurs. Using a setting similar to Kanodia and Mukherji, but basing the analysis on auction theory rather than contract theory, Coate and Loeb (1994) also find low-balling occurs in equilibrium. Dye, Kanodia and Mukherji, and Coate and Loeb all find that the motivation for low-balling is to reduce informational rents accrued to the incumbent audit firms, and it does not induce the audit firm to compromise its audit decisions.

\(^{5}\)As argued by Goldman and Barley (1974), since the attestation service the audit firm provides is valuable to the client, it confers power to the audit firm. However, it is now a commonly held notion that the power held by audit firms is diminishing due to fierce competition in the audit market.
The conclusions drawn from the abovementioned theoretical analyses lead to a consensus that there is no causal relation between low-balling and impaired auditor independence. This conclusion does not depend on an assumption that the audit firm has all the bargaining power, as in DeAngelo (1981a) and Magee and Tseng (1990); or that the client has all of it, as in Coate and Loeb (1994), Dye (1991) and Kanodia and Mukherji (1994); or that they somewhat share it. Hence, it seems that the hypothesized link between low-balling and impaired auditor independence may be unwarranted. This dissertation shifts the focus from issues of auditor independence to the economic relation among audit pricing policies (includes low-balling), audit firms' specialization decisions, and social efficiency in the audit market.

It is clear now that the extant analytical literature on audit pricing is built on the assumption that the audit market is ex-ante perfectly competitive. Given a perfectly competitive audit market, and if there is no causal relationship between audit pricing policy and auditor independence as suggested by the extant theoretical literature, then an audit firm's pricing behaviour will only affect how the benefits from an audit are divided between the client and the audit firm, with the social efficiency being held fixed. Thus, in order to address a meaningful social efficiency issue, one has to depart from the perfectly competitive paradigm. For this reason, this dissertation shifts the focus from ex-ante perfect competition to imperfect competition and emphasizes the strategic interactions among audit firms.

The only published work in the auditing competition literature which also emphasizes market imperfections is a recent article by Gigler and Penno (1995). However, their treatment of market imperfection is primitive. Gigler and Penno assume that audit firms have substantial market power because they are randomly endowed with different auditing costs. In other words, the cost differences which are modelled by them as the primary source of market power are not the result of audit firms' equilibrium behaviour. Rather, audit firms in their model are assumed to be ex-ante heterogeneous. On the contrary, this dissertation examines a setting where audit firms are ex-ante identical and strategically choose to
become *ex-post* heterogeneous in terms of their audit production costs by means of their service specialization. Thus, audit firm specialization is not a mute issue as in the existing models on auditing competition. By explicitly recognizing the strategic purpose of audit firm specialization, the models in this dissertation capture the widely held notion that audit firm specialization is the primary source of the market power and, hence, the economic rents accrue to the audit firms. As such, the models in this dissertation enrich the traditional audit pricing models by expanding the strategy spaces of the audit firm; audit firms strategically make both pricing and specialization decisions. Our emphasis on an imperfect audit market and the importance of audit firms' specialization decisions should provide insights that augment the studies of DeAngelo (1981), Coate and Loeb (1994), Dye (1991), Gigler and Penno (1995), Magee and Tseng (1990) and Kanodia and Mukherji (1994).

More specifically, this dissertation postulates a spatial perspective to examine the nature of competition in the audit services market where audit firms make strategic specialization and pricing decisions. The spatial perspective is borrowed from the spatial economics literature which represents a recent breakthrough in the development of a new industrial organization theory. Recently, many economists have increasingly recognized that the perfect competition paradigm is inappropriate to the explanation of pricing behaviour in many real life markets characterized by a significant separation between buyers and sellers (see Greenhut, Norman and Hung (1987) and Beath and Katsoulacos (1991)). Unlike the perfect competition paradigm, the spatial perspective recognizes the dispersed nature of many real life markets, and more importantly, the market power conferred to the suppliers because of natural market separation created by space. Moreover, because market activities are performed at dispersed points in space, each supplier finds only a few competitors in its immediate neighborhood. Accordingly, competition in space occurs “among the few” which are deemed as close substitutes by the buyers, thus leading to the analysis of the problem.

---

6The use of spatial analysis in audit market research is first discussed by Simunic and Stein (1987). However, they do not provide a formal spatial model to examine audit firms’ choices of both audit fee schedules and service specialization.
as a game of strategy.

The essence of the spatial perspective is that it not only provides natural market separation, but also provides a powerful analogy for some apparently nonspatial issues. The major concern of this dissertation can then be viewed as making this analogy explicit by applying the spatial analysis to audit firms’ specialization and pricing problems. Applying the spatial approach, the models in this dissertation explicitly recognize the dispersed nature of the audit market, namely that it embodies a large number of audit purchasers with different ‘characteristics’ relevant to audit production and relatively few audit suppliers who differ in their area of specialization with respect to client characteristics. In this framework, all audit clients are unique. They operate in different businesses, have different management organizations, employ different philosophies, are subject to different risks, and have different information and control environments. It follows that audit firms bidding on audit engagements have to customize their production of services to meet the unique characteristics of each client. It also creates an incentive for audit firms to specialize their services. Through specialization, an audit firm achieves a comparative cost advantage over its rivals for all clients whose characteristics are closer to its area of specialization. The interest of this dissertation is on the kinds of audit firms’ specializations that the market mechanism will provide and the implications for welfare distribution of alternative audit market environments. Moreover, explicit account is taken of the ability that audit firms have to acquire and exercise market power and of the strategic interactions among audit firms.

It is noteworthy that the spatial models developed in this dissertation are different from the ones that are generally used in economic literature to explain the phenomenon of production differentiation. While the standard product differentiation models assume that product differentiation involves making a particular firm’s product either really or apparently different from its rivals, the starting point in the models of this dissertation is that the differences

---

7One thing in common is that all audit clients, regardless of their own characteristics, require external audit services to carry out a systematic program of financial audit. This creates the demand for external auditing services.
in audit firms' specializations are grounded on the underlying differences of clients' characteristics. More specifically, it is assumed that corporate financial statement audits are homogeneous across audit firms from the viewpoint of the users.\(^8\) The rationale underlying this assumption is that professional standards impart homogeneity across audit reports, in effect causing audit services to be identical across audit firms.\(^9\) Therefore, as long as professional standards and qualifications are maintained, the users of financial statements have no reason to distinguish among audit firms. Consequently, a client's perception of audit services is assumed to be independent of the identity of the audit firm.\(^10\) That is, even though auditing activities are performed at dispersed points in client-characteristics space, each audit firm provides the same service at a given point.

The assumption that audit services are identical across audit firms does not remove all the frictions in the audit market. The reason is that audit firms differ in general in their area of service specialization with respect to client characteristics. Hence, there is heterogeneity on the cost side of audit services simply because audit firms are responsible for customizing their production of audit services to meet clients' characteristics and professional standards.

The analysis in this dissertation will take as given the nature of clients' characteristics, in its final report, the Cohen Commission states its belief that there is little or no product differentiation in the audit profession: "Public accounting firms go to considerable lengths to develop superior services for their clients, but there is little effective product differentiation from the viewpoint of the present buyer of the service, that is, the management of the corporation.... A 'clean opinion' obtained from one reputable firm is about as valuable to the competent, honest financial manager as one from another reputable firm (AICPA 1978, p. 111)."

\(^8\)Audit firms are constrained to provide a minimum level of audit quality to comply with generally accepted auditing standards (GAAS).

\(^9\)In contrast to this view, the importance of product differentiation in explaining observed market shares in the market for audit services is first asserted by Dopuch and Simunic (1980). Subsequent studies tie audit (quality) differentiation to the pricing of audit services and the relationship between audit firm size and auditor independence (see DeAngelo (1981a, 1981b) and Dopuch and Simunic (1980, 1982)). There is considerable amount of subsequent evidence for this issue which is consistent with the 'quality differentiation' (see Francis (1984) and Palmrose (1986)). In fact, one might argue that both differences on client characteristics and quality differentiation exist in the real-world audit market. However, for simplicity, this dissertation only considers the former, which has a greater effect on audit service production, and simply assumes that audit quality is not at issue. In this way, one may argue that the models in this dissertation are more relevant to a regulation requiring specific audit procedures (e.g., confirmation of accounts receivable or observation of inventory) than to regulations aimed at quality control (e.g., rules requiring proper supervision or peer review). Allowing audit quality to vary in the spatial framework considerably complicates the analysis but is, of course, a promising topic for future research.
the specification of audit technology, and a suitable notion of what would constitute an equilibrium in the particular problem under consideration. In all cases, the latter will be defined in terms of audit prices and audit firms’ specializations. The analysis provides predictions of the nature of audit firm specialization that would emerge in market equilibrium. This of course is a question of the ‘positive’ economics of auditing. However, this dissertation also considers the ‘normative’ economics of auditing. In particular, audit firms are subject to increasing degrees of regulatory controls. For example, controls or guidelines have been set with respect to audit pricing policies and switch of audit suppliers.\textsuperscript{11} In all cases, the question that would be central in the mind of a regulator is the implications of these market equilibria for market power and the welfare of the participants in the audit services market. An answer to this question involves comparing the market equilibrium with a relevant social optimum. If the market equilibrium outcome is incompatible with the social optimum and there are ways to reduce the efficiency loss, then there might exist a role for government intervention. However, the extant regulatory control on the audit market is primarily rooted in economic principles derived from classical competitive economics, while regulation is applied almost by definition to imperfectly competitive markets.\textsuperscript{12} Therefore, it is necessary that the methodological foundation of the regulation be re-examined, and perhaps, much of the regulation should be re-evaluated. Adopting the spatial approach, the analyses and results in this dissertation are shown to carry interesting policy implications with respect to policies concerning audit industry regulations such as policies concerning switching of audit suppliers studied in the chapter 3, and audit firms’ practices of ‘low-balling’ studied in chapter 4. As such, the spatial framework developed in this dissertation may shed light on a number of important audit industry issues and might provide regulators with a more

\textsuperscript{11}For example, Accounting Series Release (ASR) No. 250 requires disclosure of “fee arrangements where the audit firm has agreed to a fee significantly less than a fee that would cover expected direct costs in order to obtain the client”, whereas ASR No. 165 et al. require disclosure of both the resignation of the prior audit firm and the engagement of the new audit firm. See also footnote 13 for details.

\textsuperscript{12}For example, changes in CPAs’ codes of ethic during the 1980’s were designed to stimulate the competitiveness of audit firms, suggesting that the audit market was less than perfectly competitive when the process started.
adequate foundation on which to base regulatory judgements.

This dissertation is organized as follows. Formal models of auditing competition in a spatial context are presented in chapters 2, 3, and 4. Conclusions and suggestions for future research are the subject of the final chapter. A simple model demonstrating the existence of a demand for voluntary external auditing is provided in appendix A. All proofs of results are given in appendix B.

This dissertation is built on the belief that understanding the specialization and pricing decisions of audit firms is the cornerstone of modern audit market research. Thus, to expand that understanding, this dissertation introduces the spatial framework into the auditing competition literature. In chapter 2 a multi-period spatial oligopoly model is introduced to study how audit firms make strategic specialization and pricing decisions. Audit firms are modelled as Bertrand oligopolists who simultaneously choose specializations with respect to client characteristics and then compete in setting audit fees. It is shown that through specialization, each audit firm obtains some market power and is able to price discriminate across clients by offering 'specialization-specific' audit fee schedules. We find that, given a specialization configuration, each audit firm optimally charges the minimum of the marginal auditing costs of its rivals on services to clients whose characteristics are closer to its own specialization. Given these pricing strategies of the audit firms, the assignment of audit firms to clients is simply a function of audit cost conditions; clients purchase audit services from the least-cost supplier. The resulting allocation of clients' surplus and audit firms' profits is shown to be in the core of the economy. This means that, at the induced allocation, no group of clients can move to another audit firm for a mutually advantageous auditor-client re-alignment. Turning to the specialization decision, each audit firm optimally chooses a specialization so as to maximize its own expected profits, given the equilibrium audit pricing strategies and the specializations of its rivals. The existence of a specialization-price equilibrium is established. Surprisingly, we find that in a subgame perfect Nash equilibrium choice of audit firm specializations, given its rivals' specializations, each audit firm specializes
so as to maximize the expected social welfare (the sum of the total profits to audit firms and the aggregate surplus to clients), rather than maximize its own expected profit. It is also demonstrated that, in order to avoid intense price competition, and thus, earn rents as ‘local monopolists’, audit firms would like to differentiate themselves from each other in equilibrium. This result is consistent with a widely held notion that audit firms search for ‘niche’ markets, such as industry specialization, to increases their profits. As such, the model provides a theoretical link between audit firm specializations and the observed market segmentation in which clients with similar characteristics buy from the same audit firm, which has a cost efficiency advantage in serving them.

Chapter 3 simplifies the model developed in chapter 2 by focusing on a two-period spatial duopoly auditing competition model. This simplification allows us to add more institutional details into the model. More specifically, auditors’ learning and clients’ costs of switching audit firms are introduced to capture salient economic features of an audit market with ‘relationship-specific economic interests’. As in the multi-period spatial oligopoly model, audit firms make strategic specialization and pricing decisions. Through specialization, an audit firm achieves a comparative cost advantage over its rival for all clients whose characteristics are closer to its area of specialization. This comparative cost advantage is further fortified by the presence of learning and switching costs. Thus audit firms are able to price discriminate across clients by offering ‘specialization-and-relationship-specific’ audit fee schedules. The analysis demonstrates that the practice of ‘low-balling’ is a natural consequence of the competition among audit firms. However, low-balling occurs in a certain market segment where audit firms compete quite fiercely. Furthermore, the analysis shows how equilibrium audit fee schedules, audit firms’ specialization decisions and profits, clients’ surplus, and social welfare depend on the auditing costs, the learning rate, and the switching costs. Some of the results of the analysis are shown to carry interesting policy implications. For example, the analysis enables us to understand why there may be a conflict between the regulations (e.g., Securities Act Release No. 34-9344, ASR No. 165, ASR No. 194 and ASR No. 247)
the audit firms in the audit industry would like to adopt and those the regulators and/or the clients might want to impose.\textsuperscript{13} In this respect, our results suggest that if the objective of the regulators (particularly the SEC) is to maximize the expected social welfare, then the regulators should impose regulations that induce lower switching costs. This policy may raise audit firms' profits at the expense of clients' aggregate surplus, but it improves overall efficiency.

The issue of low-balling is further scrutinized in chapter 4. We first compare the primary similarities and differences between the predictions on low-balling as well as 'price cutting' of our model developed in chapter 3 and those of the existing literature. Then, our focus shifts to examining the welfare implications of low-balling which have not been fully considered by academics and regulators concerned with low-balling by audit firms. To this end, we compare the equilibrium outcomes derived in chapter 3 with those derived in an otherwise-equivalent economy where low-balling is not allowed, i.e., everything is the same as in the model in chapter 3 except that audit firms are required to price at or above their auditing costs. The result of this analysis provides theoretical support for banning the practice of low-balling.

Finally, chapter 5 offers some conclusions on the dissertation and points out some directions for future research.

\textsuperscript{13}The principal reporting requirements under ASR No. 165 et al. are disclosure of both the resignation of the prior audit firm and the engagement of the new audit firm, and the existence of any significant disagreement with the prior audit firm within the two most recent fiscal years. The client must request the prior audit firm to respond to the filing, and its response is appended as an exhibit. In addition, financial statement disclosure of the effect of the disagreement, if material, is required.
The purpose of this chapter is to expand our understanding about the specialization and pricing decisions of audit firms. To this end, the nature of competition in the audit services market is re-examined from a spatial perspective, which is discussed in chapter 1. In the spatial framework, the dispersed nature of the audit market is recognized: audit clients are unique and have different 'characteristics' relevant to audit production.\textsuperscript{1} The fact that clients have different characteristics leads to the natural consequence that audit firms bidding on audit engagements have to customize their production of services to meet the unique characteristics of each client.\textsuperscript{2} It also creates an incentive for audit firms to specialize their services with respect to client characteristics. Through specialization, an audit firm achieves a comparative cost advantage over its rivals for all clients whose characteristics are

\textsuperscript{1}For example, a client firm in a regulated industry requires the use of specialized financial rules for filings with a government agency; some accounting rules and applications are unique to a given industry; and client-specific, as well as industry-specific, knowledge is necessary to the audit supplier in identifying potential problem areas and communicating with client personnel.

\textsuperscript{2}As stated in Arens and Loebbecke (1984): "An extensive understanding of the client's business and industry and knowledge about the company's operations are essential for doing an adequate audit (p. 200)." Understanding the client's business at least includes an appreciation for its business and related inherent risks, and its information system and control environments. In addition, O'Keefe, Simunic and Stein (1994) find evidence of significant influences of client size, complexity, and inherent risk on the production of auditing services.
closer to its area of specialization. Together with the assumption that audit outputs, i.e., audited financial statements, are homogeneous across audit firms from the viewpoint of the users, the assignment of audit firms to clients is thus simply a function of audit firm cost conditions. That is, clients purchase audit services from the least-cost supplier.

The assumption that audit services are identical across audit firms does not remove all the frictions in the audit market. The reason is that audit firms differ in general in their area of service specialization, and therefore, are heterogeneous in terms of their auditing costs to a particular client. These cost differences then create market power for the audit firms. Thus, as a result of specialization, audit firms possess some monopoly power even though clients perceive alternative audit firms as perfect substitutes. In particular, since clients cannot resell audit contracts, their perception of homogeneous audit services does not preclude the audit firms from having the ability to engage in price discrimination.

To illustrate the above argument, this chapter presents a dynamic oligopoly model of spatial competition with price discrimination to analyze the nature of auditing competition. The model falls in the domain of a hybrid framework of Hotelling (1929) and Hoover (1937). In this framework, audit firms behave as Bertrand oligopolists and price discriminate across clients with respect to the client characteristics when providing audit services to spatially dispersed clients. That is, audit firms quote different audit fee schedules for services to different clients according to their characteristics. Owing to the competition among audit firms, the natural inference is that the cost effectiveness of audit firms determines their ultimate market shares. It is demonstrated that given a configuration of audit firm specializations,
each audit firms serves a client, to whom it has a comparative cost advantage, at an audit fee equal to the auditing cost of the second lowest-cost audit firm to that client. This equilibrium audit firm's pricing strategy is shown to be efficient as the induced allocation of clients' surplus and audit firms' profits is contained in the core of the economy. That is, at the induced allocation no group of clients can move to another audit firm for a mutually advantageous auditor-client re-alignment. When making their specialization decisions, audit firms respond to the pricing and specialization decisions of other rivals. The competitive forces in the market then induce audit firms to achieve constrained efficient utilization of specialized resources. Under some innocuous assumptions commonly used in spatial models, it is demonstrated that a specialization-price equilibrium is obtained when each audit firm maximizes expected social welfare given the specializations of its rivals.\footnote{The definition of social welfare is given in section 2.4.} Moreover, it can never be beneficial for an audit firm to choose a specialization arbitrarily close to any of its rivals'. This result implies that audit firms tend to differentiate themselves from each other by means of service specialization. As such, the model provides a theoretical link between audit firm specializations and the observed market segmentation in which clients with similar characteristics buy from the same audit firm, which has a cost efficiency advantage in serving them.

The rest of the chapter is organized as follows. As is obvious from the discussion thus far, the dissertation heavily draws on the field of spatial economics for input. A brief review of this literature appears in section 2.1. Section 2.2 presents a very general multi-period oligopoly spatial auditing competition model, provides a precise specification of the demand and supply sides of the audit market, and defines a specialization-price game for the audit firms and an appropriate solution concept. Section 2.3 derives the unique audit pricing equilibrium for the model. Section 2.4 establishes the existence of a specialization-price equilibrium. In addition to existence, some interesting properties of the specialization equilibria are also demonstrated. Section 2.5 concludes the chapter.
2.1 A Brief Literature Review of Spatial Competition

Beginning with the work of Hotelling (1929), the study of spatial competition has provided important insights into markets for differentiated product. The distinctive feature of the Hotelling's spatial model, compared to other models of product differentiation, is that it allows an explicit representation of product choice by oligopolistic firms. Specifically, the form of differentiation introduced by Hotelling (1929) can be described as 'horizontal' in the sense that no product (location) is unanimously preferred by all consumers. In this context, products differ only because they are offered at different locations. In his model, Hotelling considers two identical firms that produce a single homogeneous product with a constant production cost in a bounded linear market over which consumers with inelastic demand are uniformly distributed. The firms compete in location and price and the consumers purchase the product from the cheapest source and pay a transport cost which is assumed linear with respect to the distance between the locations of the consumer and the firm. For each pair of locations chosen by the firms, Hotelling calculates the equilibrium prices they would set. To study the location tendencies, Hotelling introduces these equilibrium prices back into the firms' profit functions. In this respect, Hotelling can be said to have studied a subgame perfect Nash equilibrium in a two-stage location-price game. Hotelling claims that a Nash equilibrium in locations for the two firm market exists and yields 'back-to-back' locations at the center of the market.

Some problems with Hotelling's analysis later become apparent. D'Aspremont, Gabszewicz and Thisse (1979) find that no equilibrium (in pure strategies) in prices exists when firms are located too close to each other. But, if no price equilibrium exists for certain locational choices, then there is no way for firms to estimate the profitability of those locations. That is, the disturbing result of D'Aspremont, Gabszewicz and Thisse is that location tendencies cannot be derived because the outcome of the price game is not well-defined.

In contrast, two products are said to be 'vertically' differentiated if all consumers unanimously rank unit quantities of them. Thus, if they are sold at the same price, all consumers purchase the same product.
Many resolutions of the existence problem have been discussed in the literature. Basically, they can be divided into four areas: (1) changing the transport cost function, (2) allowing for mixed strategies over prices, (3) focusing on vertical as opposed to horizontal location problems, and (4) allowing for discriminatory pricing.

D’Aspremont, Gabszewicz and Thisse point out that the problem of nonexistence of a noncooperative equilibrium arises from the fact that, with linear transport cost, the firms’ demand functions are discontinuous and their profit functions are discontinuous and non-concave. Consequently, the price-competition problem is not well-behaved. To obviate this nonexistence problem they assume quadratic as opposed to linear transportation costs. The equilibrium locations are at the two ends of the linear market. Similar results are derived in Economides (1984). He shows that a pricing equilibrium exists and firms locate far apart when consumers have a maximal or reservation distance. Unfortunately, since the equilibrium depends heavily on the form of transportation cost function, few applications have been developed.

The fact that there is no equilibrium in pure strategies (over prices) does not preclude the existence of a mixed strategy equilibrium (see Dasgupta and Maskin (1986)). Nevertheless, it is of interest that Hotelling’s model with linear transportation costs and bounded reservation prices possesses no equilibrium even in mixed strategies. In games where mixed strategy equilibria do exist (see Gal-Or (1982) and Osborne and Pitchik (1987)), their complexity effectively rules out comparative static analysis.

A completely different approach was taken by Shaked and Sutton (1982, 1983). They attain important positive results in price-location theory by examining vertical rather than horizontal product differentiation. In their model, firms compete over quality and price. Quality choice is a ‘vertical’ location problem because all consumers prefer higher quality to lower quality. By contrast, in ‘horizontal’ location problems, changing the product specification is a move towards some consumers and away from others. Their results are
encouraging, but it has proved difficult to develop generalizations that include horizontal product differentiation.

In the traditional Hotelling setting, Hotelling assumes there is no price discrimination. Consumers pay the costs of transporting the product from firm to home plus the mill price set by the firm. An alternative approach to spatial competition, pioneered by Hoover (1937), relaxes this constraint and allows firms to price discriminatorily. This situation is plausible if firms can identify consumer locations. It is of interest that even if consumer locations are not directly observable, price discrimination may still be possible if firms choose delivered price schedules over a space which precludes consumer arbitrage. In such situations, unless there exist regulations dictating otherwise, the firm has the potential for price discrimination. In his original work, Hoover analyzes spatial price discrimination for firms with fixed locations. He concludes that a firm serving a market point would have a local price constrained by the marginal cost of service of other firms. In situations where demand elasticity is not too high, this will result in delivered prices at market points equal to the marginal cost of the firm in the market with the second lowest marginal cost. This research agenda has then been further developed by Hurter and Lederer (1985), Lederer and Hurter (1986) and Hamilton, Thiss and Weskamp (1989). They show that a two-stage perfect Nash equilibrium of prices and locations exists when firms are allowed to set discriminatory prices. Above all, the resulting game typically involves, as strategic variables, price schedules specifying the delivery prices at which each firm is willing to supply the consumer at each point of space. In other words, if firms are allowed to price discriminate in a spatial market, their decision variables are price functions instead of price scalars. The models in this dissertation basically follow this particular line of research. However, the key difference in the models in this dissertation is that firms will compete for consumers over a time horizon rather than a single period. Such multi-period extensions give rise to the consideration of the firm's opportunity to learn and the consumer's switching costs, which are the subjects of examination in chapter 3. As one would expect, the resulting multi-period models are much richer than their single-
period counterpart, since the firm's strategy set is expanded. Moreover, the incorporation of learning-by-doing and switching costs into the spatial competition model is novel in the economic literature.

2.2 The Model

Consider an economy that lasts for $T$ periods. There is a continuum of client firms (henceforth called clients) distributed over a convex compact subset $Z \subset \mathbb{R}^N$, where $Z$ is the domain of client's 'characteristics' which are relevant to audit production. That is, the characteristics of each client are fully described by a vector $z \in Z$. Moreover, the density of clients at $z \in Z$, $h(z)$, is positive and continuous on $Z$. For expositional convenience, the client(s) located at $z$ will be referred to as 'client $z$'.

Each client would like to acquire one unit of audit service from an external audit firm in every period, provided that the benefit is higher than the cost of the audit. It is assumed that auditing standards are maintained by each audit supplier, such that audit services are qualitatively homogeneous across audit firms from the viewpoint of the clients. Furthermore, a unit of audit service, regardless of the identity of the audit supplier, is assumed to give a client $z$ a gross benefit of $b_z^p$ per period, where $b_z^p$ can be interpreted as the highest audit fee that client $z$ would be willing to pay for a unit of audit service. It is assumed that $b_z^p$ is sufficiently large, so that each client would like to purchase the audit service in every period. Other than the audit fee paid to his audit firm, there are no additional transaction costs incurred by the client regarding the hiring or switching of audit firms.

---

3. $T$ can be finite or infinite.
10 Characteristic space' is a natural criterion for the separation of the market for audit services in the context of imperfect competition.
11 For a useful background on the various sources of demand for audit services, see Jensen and Meckling (1976), Ng (1978, 1979), Benston (1985), Watts and Zimmerman (1986), and Berry and Wallace (1986). A simple model to explain the existence of a voluntary demand for external auditing is also provided in appendix A.
12 There will be no qualitatively change of the analysis if $b$ is also time-dependent. Here, the assumption is made for notational convenience.
13 This assumption is relaxed in chapter 3.
The audit market consists of \( n \) independent (non-colluding) audit firms bidding on audit engagements. They are indexed by \( i \in \{1, 2, \ldots, n\} \) and may only differ in specialization of services with respect to client characteristics. At the beginning of the first period, audit firms choose their area of specialization simultaneously; once chosen, the specializations are fixed forever. For simplicity, each audit firm is only allowed to choose a single type of service specialization.\(^\text{14}\) Then, for each period, they simultaneously quote audit fees to each client. It is assumed that no multi-period offers are permitted.\(^\text{15}\)

Let \( l_i \in Z \) denote the specialization of audit firm \( i \) in the economy. Audit firms are responsible for customizing their production of audit services to meet clients’ characteristics. The audit technology available to each audit firm is the same. When the difference between audit firm \( i \)'s specialization and client \( z \)'s characteristics is equal to \( \|l_i - z\| \), the auditing cost per period is given by a function \( m(\|l_i - z\|) \), where \( \| \cdot \| \) is a norm defined on \( Z \).\(^\text{16}\) \( m(\|l_i - z\|) \) is increasing and continuous in \( \|l_i - z\| \) with \( m(0) = 0 \). Again, for simplicity, it is assumed that there is no learning-by-doing advantages by the incumbent audit firm, so that \( m(\cdot) \) is independent of time.\(^\text{17}\)

The assumptions of the audit production function imply that an audit firm’s auditing cost to a particular client can be reduced by simply choosing a specialization that is closer to the characteristics of that client. That is, the benefit from cost reduction gives an incentive

---

\(^\text{14}\)Presumably, one can expect there is a fixed cost of specialization which is an irreversible investment (e.g., costs associated with the technology adopted and the human capital/expertise of professional staff), otherwise audit firms will simply specialize at all points at which there are clients and demand will be perfectly satisfied. More specifically, it is assumed that the fixed cost is low enough for the audit firm to make nonnegative profit but high enough to prevent it from having more than one specialization. Similarly, the fixed cost is assumed to be high enough to prevent more than \( n \) audit firms to co-exist in the audit market. Put differently, the audit market defined by \( Z \) is assumed to be just large enough to allow exactly \( n \) audit firms to make positive profits (net offixed cost). However, it will be clear later in the analysis that the fixed cost does not play any important role in the analysis. Therefore, it is intentionally omitted to reduce the notational burden.

\(^\text{15}\)In other words, audit firms are not able to make binding long term commitments for future audit fees. In fact, such binding long term commitments are rare in practice, probably because of prohibitions on audit contracts that are contingent on the content of audited reports.

\(^\text{16}\)Given arbitrary points \( z_1, z_2 \) and \( z_3 \in Z \), a norm is a real-valued function which satisfies (i) \( \|z_1 - z_2\| = 0 \) if, and only if, \( z_1 = z_2 \); (ii) \( \|z_1 - z_2\| + \|z_2 - z_3\| \geq \|z_1 - z_3\| \); and (iii) \( \|z_1 - z_2\| = \|z_2 - z_1\| \geq 0 \).

\(^\text{17}\)Again, this assumption is relaxed in chapter 3.
for audit firms to specialize their services. By means of specialization, an audit firm becomes more cost efficient, compared to its rivals, to serve clients whose characteristics are closer to its area of specialization. Consistent with this line of thinking, empirical researchers find that audit firms specialize by industry (Dopuch and Simunic (1980), Danos and Eichenseher (1982), and Eichenseher (1985)). As one would expect, specialists in the client’s industry are likely to enjoy cost advantages over nonspecialists.

Formally, the setting is a $T$-period, $T + 1$-stage complete information game. In the first stage (the specialization stage) which occurs at the beginning of the first period, the audit firms simultaneously choose their specializations in $Z$. Then, each audit firm becomes aware of its rivals’ specializations. It implies that, after the audit firms choose their specializations, everyone knows who the most cost efficient audit firm is for a particular client. In the second stage (the first period pricing stage), the audit firms simultaneously quote audit fees to each client. The client at $z$ acquires auditing services from the audit firm quoting the lowest audit fee. When audit fees are equal, it is assumed the audit firm with the lowest auditing cost provides the auditing services to the client. This may be rationalized by noting that the most cost efficient audit firm can always offer a strictly better audit fee schedule to the client. Furthermore, if two or more audit firms have equal lowest costs of auditing client $z$ and quote equal lowest audit fees to him, the client chooses the one with the lowest index. Generally, the set of clients and audit firms for which $m(||i - z||) = m(||l - z||)$, $i \neq j$, is negligible. Then, it follows that the ‘tie-breaking’ rule used in this latter case is of no consequence in the equilibrium analysis. The first-period pricing stage will then repeat from periods 2 to $T$. Common to both audit firms and clients is the assumption that there is a one-period time-independent discount factor $\delta \in (0, 1)$ for future revenues (benefits) and

---

18 The assumption on the sequence of audit firms’ decisions is motivated by the fact that the choice about specialization occurs prior to the decisions on audit fees and specialization is an irreversible investment.

19 It is clear that the audit fee setting and competition stage is closely related to the model of price competition studied by Bertrand (1883).

20 This assumption explicitly avoids defining an equilibrium in terms of an $\varepsilon$-equilibrium where the cost efficient audit firm slightly undercut the other’s auditing cost.

21 See proposition 2.5.
costs.

The equilibrium concept employed is Selten’s (1975) subgame perfect Nash equilibrium (SPNE) and attention is restricted to pure strategy equilibria only. A set of pure strategies for a game is an SPNE if it is a Nash equilibrium for the entire game and its relevant action rules are a Nash equilibrium for every proper subgame. This is the appropriate equilibrium concept for a complete information game. In this model, a strategy for audit firm $i$ is an ordered pair, $(l_i, \{(F_{it})_{t=1}^T\})$, that consists of the audit firm’s specialization, $l_i$, and a sequence of time-dependent functions, $\{F_{it}\}_{t=1}^T$, mapping every possible observed combination of $(l_1, l_2, ..., l_n)$ and audit fee history $\{(F_{1t}, F_{2t}, ..., F_{nt})\}_{t=1}^{t-1}$ into a sequence of period-$t$ audit fee schedules quoted to client $z$, $\{f_{zt}\}_{z \in Z}$. A strategy for client $z$ is a sequence of time-dependent functions $\{Q_t^z\}_{t=1}^T$ that maps every possible combination of $(f_{1t}, f_{2t}, ..., f_{nt})$ in period $t$ into $\{1, 2, ..., n\}$, where $\{1, 2, ..., n\}$ is the set of audit firms in the market. Hence, an SPNE strategy choice is an ordered pair, $(\{(l_i^*, (F_{it})_{t=1}^T)\}_{i=1}^n, \{Q_t^z\}_{t=1}^T)_{z \in Z}$, such that (i) no player can improve his payoff by unilaterally deviating, (ii) $\{F_{it}\}_{t=1}^T$ constitutes SPNE choices of $\{f_{zt}\}_{z \in Z}$, $i = 1, 2, ..., n$ and $t = 1, 2, ..., T$, for every possible prior choice of $(l_1, l_2, ..., l_n)$ and $\{(F_{1t}, F_{2t}, ..., F_{nt})\}_{t=1}^t$, and (iii) $\{Q_t^z\}_{z \in Z}$ constitutes SPNE choices of audit suppliers for every possible combination of $(f_{1t}, f_{2t}, ..., f_{nt})$, for each $z \in Z$, and for all $t = 1, 2, ..., T$.

The characterization of the SPNE proceeds in two steps. The first step is to characterize the SPNE for subgames starting from stages 2 to $T + 1$ (henceforth called the pricing subgames) defined by every possible audit firm specialization choices of $(l_1, l_2, ..., l_n)$. Once this has been done, the SPNE for the specialization stage is readily solved.

---

22The concept of subgame perfect Nash equilibrium captures the idea that, when audit firms choose their specialization, they all anticipate the consequences of their choice on future audit fee competition. In particular, they are aware that this competition will be more severe if their specializations are close to one another, rather than far apart.

23Obviously, the audit fee history at the beginning of the first period is a null set.
2.3 The Pricing Equilibrium

Consider the pricing stage under a subgame defined by \((l_1, l_2, \ldots, l_n)\). Given that (i) auditing cost to a particular client \(z\) in period \(t\) is time-independent and unaffected by the costs to other clients, and (ii) clients are not able to resell audit contracts, audit fee schedules quoted by an audit firm to a particular client across different periods or to different clients in a particular period are strategically independent. It follows that the equilibrium audit fee for each client should be the same for each period. Since audit fees are stable over time, no client will ever switch auditors in equilibrium.\(^{24}\) Thus, the pricing subgames are equivalent to a \(T\)-period repeated Bertrand pricing game at each point in client-characteristics space under asymmetric auditing cost conditions. The equilibrium of the pricing subgames can then be characterised by a set of client-specific and time-independent Bertrand pricing equilibria, one for each client at \(z \in Z\) repeated for \(T\) periods. Since all decision variables are time-independent, the subscript for time is suppressed from now on. In the sequel, the SPNE of the pricing subgames will be derived as in a one-shot pricing game.

Notice that, as a result of specialization and the fact that clients cannot resell audit contracts, audit firms possess some monopoly power and are able to price discriminate across clients as they take into consideration the heterogeneity (by characteristics) among clients.\(^ {25}\) In general, since the audit fee schedule \(f_i^z\) specifies the audit fee at which audit firm \(i\) is willing to supply audit services to client \(z \in Z\), it must cover its auditing cost to that client. Moreover, in order to induce a client to accept an offer, the audit fee must not exceed the maximum the client would be willing to pay. Thus, formally, \(f_i^z\) is in the set:

\[ F_i \equiv \{ f_i^z : \text{a nonnegative function defined on } Z, \text{ measurable and such that,} \]
for all \( z \in Z, b^z \geq f^z_i \geq m(||l_i - z||). \)

It is also worthwhile to mention that, in a complete information game, each audit firm knows the characteristics of the client, and knows the specialization of its rivals. Thus, each audit firm can calculate the audit fee offered by the other audit firms, and respond with an audit fee that is attractive to the client but still leaves it with a monopolistic rent. As a result, each client receives a set of audit fee schedules which depends on his characteristics relative to the audit firms' specializations.

Suppose client \( z \) receives and accepts an audit fee \( f^z_i \) from audit firm \( i \), the one-period surplus for him is defined as

\[
S^z(f^z_i) \equiv b^z - f^z_i.
\]

The client accepts the offer which gives him a nonnegative and highest surplus. As stated before, it is assumed that when a client receives equal surplus from two or more audit firms, the client chooses to patronize the audit firm that has the lowest auditing cost to serve him. If client \( z \) rejects all offers, his surplus is zero. Recall that an audit is not mandated in our model. Therefore, it is clear that \( S^z(f^z_i) \geq 0 \) since client \( z \) has the right to reject any offer that gives him a negative surplus. Thus, given a configuration of specializations \( (l_i, l_{-i}) \) and a set of audit fee schedules, \((f^z_i, f^z_{-i})\), where \(-i \equiv \{1, 2, ..., i - 1, i + 1, ..., n\}\), quoted to client \( z \), the one-period profit that audit firm \( i \) earns from client \( z \) is\(^{26}\)

\[
\Pi^z_i(l_i, l_{-i}, f^z_i, f^z_{-i}) = \begin{cases} 
 f^z_i - m(||l_i - z||) & \text{if } S^z(f^z_i) > S^z(f^z_j) \text{ for all } j \neq i, \\
 0 & \text{if } S^z(f^z_i) \geq S^z(f^z_j) \text{ for at least one } j \neq i.
\end{cases}
\]

Hence, given any specialization configuration \((l_1, l_2, ..., l_n)\), an SPNE audit fee schedule in pure strategies is an n-tuple \((F^*_1, F^*_2, ..., F^*_n)\) of audit fee schedules such that

\[
\Pi^z_i(l_i, l_{-i}, f^z_i, f^z_{-i}) \geq \Pi^z_i(l_i, l_{-i}, f^z_i, f^z_{-i}),
\]

for all \( f^z_i \in \mathcal{F}_i, i = 1, 2, ..., n \) and \( z \in Z \).

\(^{26}\)Note that audit firm \( i \)'s profit is not directly influenced by its rivals' specialization choice. Instead, audit firm \( i \)'s direct concern is only the current fees offered by its rivals. However, those fees are in general dependent on the rivals' specializations.
Following a standard Bertrand argument in spatial models, in equilibrium the audit firm with the lowest auditing cost to serve client $z$ will exclusively audit him since it can profitably undercut any audit fee set by a rival. Thus we have:

**Proposition 2.1.** There exists a unique SPNE audit fee schedule which is given by

$$f_{i}^{*} = \begin{cases} \min_{j \neq i} m(||l_j - z||) & \text{if } m(||l_i - z||) < m(||l_j - z||) \text{ for all } j \neq i, \\ m(||l_i - z||) & \text{otherwise}, \end{cases}$$

for $i = 1, 2, \ldots, n$.

Proposition 2.1 is a very strong result and depends on only two innocuous assumptions: (i) audit firms are able to set discriminatory audit fees according to client characteristics, and (ii) clients cannot resell audit contracts. Both of them are believed to be prevalent in the audit markets.\(^2\) Other than these two assumptions, the existence of an SPNE audit fee schedule equilibrium (in pure strategies) is guaranteed for any configuration of specializations $(l_1, l_2, \ldots, l_n)$, audit technology $m(\cdot)$, and distribution of clients. Moreover, the general structure of the equilibrium audit fee schedule is robust to arbitrary client distributions, auditing cost functions, multidimensional client-characteristics space, and many audit firms. The SPNE audit fee schedule is such that the lowest-cost audit firm serves a client at an audit fee equal to the auditing cost of the second lowest-cost audit firm to that client. This implies that competition among audit firms becomes 'localized'. An audit firm’s pricing strategies will have a powerful impact on those rivals whose specializations are very similar to it, but will only have a weak impact on those rivals whose specializations are very different from it.

The SPNE pricing strategies characterized in proposition 2.1 provide some interesting empirical implications. Prior empirical studies on audit pricing have suggested that there is a direct relationship between client characteristics and the audit fee charged to that client.

\(^2\)The client-specific nature of audit services guarantees the satisfaction of condition (ii), and makes it virtually impossible for regulators to impose restrictions that would violate condition (i).
(e.g., Simunic (1980) and Palmrose (1986)). Our result suggests that this relationship is indirect. Instead, the equilibrium audit fee to a client is directly related to the cost of the closest substitute in the audit market, i.e., the auditing cost of the second lowest-cost audit firm to that client. Since the cost of the closest substitute depends on the difference between the client's characteristics and the specialization of the second lowest-cost audit firm, the client's characteristics only indirectly affect the audit fee charged to him by the lowest-cost audit firm through their influence on the auditing cost of the second lowest-cost audit firm. In other words, while client characteristics directly affect the supplier's cost, the supplier's fee charged to the client is only indirectly affected by the effect of client characteristics on the closest competitor's cost. Nevertheless, our result is not at odds with the empirical findings which document a positive relationship between client characteristics and audit fee. It is because the client's characteristics may be a good proxy for the difference between the client's characteristics and the specialization of the second lowest-cost audit firm. However, our result implies that one has to be cautious when interpreting the empirical results regarding the relationship between client characteristics and audit fee.

The following caveat is in order before proceeding. The above analysis implicitly assumes that audit firms make a take-it-or-leave-it offer to the client, while the latter is not allowed to respond with a counter offer (which would start a process of bargaining). This implies that the lowest-cost audit firm is assigned superior bargaining power relative to that of the client in the auditor-client matching game. In spite of the lowest-cost audit firm's superior bargaining position, the availability of the other audit firms in the market provides the client with an option that defines his bargaining position (i.e., his reservation surplus) when the lowest-cost audit firm makes an offer to him. That is, the client uses an offer from the second lowest-cost audit firm to obtain a lower offer from the lowest-cost audit firm. In this respect, one might argue that a more direct approach would be to use a bargaining game to study the interaction among audit firms and clients. However, the difficulty of this approach is that the outcome is sensitive to the way the extensive form of the bargaining game is
defined (see Bester (1989)). Above all, even though our approach is somewhat arbitrary (the same comment can equally apply to almost all the existing models on audit pricing), the equilibrium pricing strategies in our model can be shown to result in an allocation of clients’ surplus and audit firms’ profits that lies in the core of the economy. It means that, at the final allocation, no group of clients can move to another audit firm for a mutually advantageous auditor-client re-alignment. To see this, let \( Z = Z \cup \{1, 2, \ldots, n\} \) be the set of participants in the audit market (clients and \( n \) audit firms), \( S \subseteq Z \) be an arbitrary coalition, and \( A(S) \) be the set of feasible allocations for a coalition \( S \). An allocation \( \Theta \in A(Z) \) is said to be in the core if, and only, if there does not exist a coalition \( S \subseteq Z \) and an allocation \( \hat{\Theta} \in A(S) \) in which all members of \( S \) are better off. As such, if profits to audit firms and surplus to clients in the coalition can be increased, the current allocation is not in the core. Clearly, since clients cannot resell audit contracts and we do not allow clients to collude, there is no interaction among clients. Similarly, since audit firms are not allowed to collude, audit firms can only gain by making deals with clients. Hence, a blocking coalition must contain a nonempty subset of clients and at least one audit firm. To see that, given a configuration of specializations, the allocation induced by the equilibrium pricing strategies described in proposition 2.1 is contained in the core, let us recall how equilibrium audit pricing strategies are determined. For each client \( z \in Z \), the equilibrium audit fee is constructed to maximizing profit for the supplying audit firm (the lowest-cost audit firm) by extracting all the surplus in excess of the surplus the client could obtain from the second lowest-cost audit firm. Thus no other offer could make both the client \( z \) and the supplying audit firm better off, and no other offer for the client \( z \) could be profitably provided by the other audit firms. Therefore, given a configuration of specializations, the allocation of clients’ surplus and audit firms’ profits induced by the SPNE audit fee schedule is in the core, and hence, is efficient.
2.4 Specialization Equilibria

In the specialization stage, audit firms choose specializations looking ahead to the pricing stage outcome derived in the previous section. Under the equilibrium audit fee schedule and the assumed tie-breaking rule, the audit markets served by each of the audit firms can be defined as

\[ Z_i(l_i, l_{-i}) = \hat{Z}_i(l_i, l_{-i}) \cup \partial \hat{Z}_i(l_i, l_{-i}), \]

where

\[ \hat{Z}_i(l_i, l_{-i}) = \{ z \in Z \mid m(||l_i - z||) < m(||l_j - z||) \text{ for all } j \neq i \}, \]

\[ \partial \hat{Z}_i(l_i, l_{-i}) = \{ z \in Z \mid m(||l_i - z||) \leq m(||l_j - z||) \text{ for all } j \neq i, \text{ and } \]

\[ m(||l_i - z||) = m(||l_j - z||) \text{ for at least one } j \in \{ i+1, i+2, \ldots, n \}. \]

\( Z_i \) is the audit market exclusively served by audit firm \( i \). It is assumed that \( \bigcup_{i=1}^{n} Z_i = Z \), such that the whole audit market will be covered. It is also easy to see that under the SPNE audit fee schedule, the audit firms in \( \partial \hat{Z}_i \) earn zero profit. Thus, audit firm \( i \)'s total expected profit (in present value) at the beginning of the first period can be written as

\[ \Pi_i(l_i, l_{-i}, f_i^{2*}, f_{-i}^{2*}) = \frac{\delta(1 - \delta^T)}{1 - \delta} \int_{Z_i} \left[ \min_{j \neq i} m(||l_j - z||) - m(||l_i - z||) \right] h(z) \, dz. \]

A noncooperative SPNE choice of audit firm specializations in pure strategies is an \( n \)-tuple \( (l_1^*, l_2^*, \ldots, l_n^*) \) of audit firm specializations such that

\[ \Pi_i(l_i^*, l_{-i}^*, f_i^{2*}, f_{-i}^{2*}) \geq \Pi_i(l_i, l_{-i}, f_i^{2*}, f_{-i}^{2*}), \]

for all \( l_i \in Z, i = 1, 2, \ldots, n \). That is, in the specialization stage, an SPNE of specializations obtains when each audit firm chooses its specialization so as to maximize its total expected profit given its rivals' specialization choices.

Before proceeding to analyze an SPNE choice of audit firm specializations, a few preliminary results will prove helpful and add insight to the properties that equilibrium specializations of audit firms must obey. Let define \( C(l_1, l_2, \ldots, l_n) \) and \( S(l_1, l_2, \ldots, l_n) \) be the total
expected costs for auditing services (in present value) and the aggregate expected surplus to 
clients (in present value) given a specialization configuration \((l_1, l_2, ..., l_n)\), respectively.

\[
C(l_1, l_2, ..., l_n) = \frac{\delta(1-\delta^T)}{1-\delta} \int_{\mathbb{Z}} \left[ \min_i m(||l_i - z||) \right] h(z) \, dz,
\]

\[
S(l_1, l_2, ..., l_n) = \frac{\delta(1-\delta^T)}{1-\delta} \sum_{i=1}^{n} \int_{Z_i} \left[ b^z - \min_{j \neq i} m(||l_j - z||) \right] h(z) \, dz.
\]

Furthermore, denote by \(W(l_1, l_2, ..., l_n)\) the expected social welfare (in present value) which 
is taken to be the sum of the total expected profit to audit firms and the aggregate expected 
surplus to clients, i.e.,

\[
W(l_1, l_2, ..., l_n) = \sum_{i=1}^{n} \Pi_i(l_1, l_2, ..., l_n) + S(l_1, l_2, ..., l_n).
\]

It follows immediately that \(W(l_1, l_2, ..., l_n)\) can be rewritten as

\[
W(l_1, l_2, ..., l_n) = \frac{\delta(1-\delta^T)}{1-\delta} \int_{\mathbb{Z}} b^z h(z) \, dz - C(l_1, l_2, ..., l_n).
\]

The following proposition states the economic relation between the cost-minimization 
and welfare-maximization audit firm specializations.

**Proposition 2.2.** A specialization configuration \((l_1, l_2, ..., l_n)\) that maximizes the 
expected social welfare also minimizes the total expected costs for auditing services.

Because demand for audit service is perfectly inelastic, if social welfare is defined as 
\(W(l_1, l_2, ..., l_n)\), then the maximization of social welfare with respect to specialization re-
duces to the minimization of total expected auditing costs.

The next proposition states that a specialization configuration that maximizes the ex-
pected social welfare (or equivalently, minimizes the total expected costs for auditing services) 
always exists.

---

28Since the model assumes that corporate financial statement audits are homogeneous across audit firms 
from the viewpoint of the users, the welfare of the end users of audited financial statements should not be 
affected by the result of the auditor-client matching.
Proposition 2.3. There always exists a specialization configuration that maximizes the expected social welfare.

Next, it is demonstrated that the existence of an SPNE choice of audit firm specializations depends on the existence of specializations that maximize the expected social welfare. Since, proposition 2.3 provides the existence of such specializations, then the existence of an SPNE choice of audit firm specializations is always assured.

Proposition 2.4. An SPNE choice of audit firm specializations, \((l^*_i, l^*_{-i})\), exists and satisfies

\[ W(l^*_i, l^*_{-i}) \geq W(l_i, l^*_{-i}), \]

for all \( l_i \in Z, i = 1, 2, ..., n. \)

Proposition 2.4 states that the existence of an SPNE choice of specializations hinges on the existence of specializations such that each specialization chosen by an audit firm maximizes the expected social welfare given its rivals' specializations. Such specializations exist by the result of proposition 2.3. Thus, under the equilibrium audit fee schedule derived in the previous section and given the specializations of its rivals, each audit firm chooses a specialization that maximizes the expected social welfare. However, it does not imply that the welfare-maximization specializations are the only equilibrium specializations. In fact, the set of welfare-maximization specializations is likely to be a proper subset of the equilibrium specializations. This is because, even though \( l^*_i \) must locally maximize \( W(l_i, l^*_{-i}) \) given \( l^*_{-i} \), it does not imply that \( (l^*_i, l^*_{-i}) \) globally maximizes \( W(l_i, L_{-i}) \).

The next proposition states another important property of an SPNE choice of special-

\[ ^{29}\text{In other words, there may exist multiple equilibria. As such, the predictive ability of the model will be reduced. In order to avoid multiple equilibria, more structure has to be given to the model such that the equilibrium is unique. The simplifications we made in chapter 3 ensure that the equilibrium is unique and, hence, there is no ambiguity regarding the predicted consequences on welfare of changes in the model's parameters.} \]
Proposition 2.5. Audit firms will choose different specializations in equilibrium.

The intuition behind proposition 2.5 is that if an audit firm chooses the same specialization with at least one rival, profits are driven to zero by intense price competition. Anticipating this outcome, audit firms will never choose the same specialization. Therefore, Bertrand competition drives audit firms to choose different specializations in order to earn positive profits. In other words, in an SPNE of audit firm specializations, audit firms have a tendency to differentiate themselves in order to relax price competition.\(^3\)

2.5 Concluding Remarks

This chapter re-examines the nature of competition in the audit market from a spatial perspective. Audit firms in the model make strategic specialization and pricing decisions. Through specialization, an audit firm achieves a cost advantage over its rival for all clients whose characteristics are closer to its area of specialization. Thus, each audit firm obtains some market power and is able to price discriminate across clients by offering ‘specialization-specific’ audit fee schedules.

The analysis demonstrates that the unique SPNE choice of audit fee schedules requires

\(^3\)In the terminology of spatial economics, our result finds that the principle of minimum differentiation does not hold. On the contrary, using a modified model as in D'Aspremont, Gabszewicz and Thisse (1979) but assuming firms can collude on prices, Friedman and Thisse (1993) find that the unique equilibrium outcome involves all firms choosing the same specialization right at the middle of the market. That is, Friedman and Thisse restore the principle of minimum differentiation. The same result is obtained in Chan (1993) where he assumes inelastic demand and firms can set discriminatory prices. All in all, it is not surprising that the presence of price collusion induces more supplier concentration because, contrary to Bertrand competition, price collusion does not lead to zero profits when firms choose the same specializations. What is perhaps more surprising is that price collusion induces no differentiation at all. The reason for this seemingly surprising result is as follows. In the spatial framework, choosing the same specialization means that firms' ability to punish each other for defection is maximized once the equilibrium specializations are selected. In this case the non-cooperative equilibrium profits are zero. Thus once the same specialization has been chosen by all firms, the punishment for defecting is naturally the most severe punishment possible within the model. Hence, the collusion is sustainable.
each audit firm to charge the minimum of the marginal auditing costs of its rivals on services to clients whose characteristics are in the vicinity to its own specialization. Given the specializations of audit firms, these pricing strategies induce an allocation of clients’ surplus and audit firms’ profits that is in the core of the economy. The existence of an SPNE choice of specializations is also established. We find that, given the specializations of its rivals, an SPNE choice of specializations requires each audit firm to specialize such that the expected social welfare is maximized. Moreover, audit firms will not choose the same specialization in equilibrium. Instead, in order to earn rents as ‘local monopolists’, audit firms will search for ‘niche’ markets such as industry specialization. Thus, the model provides a theoretical link between audit firm specializations and the observed market segmentation.
Chapter 3

A Two-Period Spatial Model of Auditing Competition with Learning and Switching Costs

In the first year of an audit, audit firms incur substantial ‘start up’ costs when learning about new clients’ operations and checking their financial statements.\(^1\) If clients terminate the relationship with their incumbent audit firms and establish another with new audit firms, these start up costs must be incurred again. In addition, once an audit firm has performed an initial audit for a given client, it has acquired specialized knowledge of that client and can therefore reduce its auditing costs when serving this particular client in future periods. Hence, the existence of ‘learning’, which includes both the start-up costs and learning-by-doing advantages, in the provision of audit services provides comparative cost advantages to an incumbent audit firm when recontracting occurs.\(^2\)

\(^1\) Arens and Loebbecke (1984) provide three reasons for the existence of significant start-up costs entailed in initial audit engagements: (1) it is necessary to verify the details making up those balance sheet accounts that are of a permanent nature, such as fixed assets, patents, and retained earnings; (2) it is necessary to verify the beginning balances sheet accounts on an initial engagement; and (3) the audit firm is less familiar with the client’s operations in an initial audit (p. 150-1).

\(^2\) In practice there is a cycle to recurring audits. The completion of one year’s audit naturally leads to and provides inputs for the planning phase of the following year’s audit. The knowledge gained from previous audits accumulates and contributes to cost advantages. For example, assessment of a client’s inherent risk is based on the audit firm’s cumulative audit knowledge and its updated understanding of the client’s business, information, accounting and control systems. The nature and level of inherent risk directly influences the
On the other hand, termination of an audit firm can also impose costs on the client. In general, a client has to incur ‘switching costs’ if he employs an audit firm that he did not hire in the previous period. The client’s switching costs arise from the need to solicit presentations from a potential audit firm and, therefore, include the cost of adapting from one audit firm to another.\(^3\) Needless to say, it does not pay for the client to build up a new relationship with another audit firm if the benefit from switching to a new audit firm does not fully cover the switching costs. Thus, clients may display loyalty and continue to use their incumbent audit firms simply because of the existence of switching costs.

The above discussion suggests that the presence of audit firms’ learning and clients’ costs of switching audit firms creates ‘relationship-specific economic interests’ which provide the joint incentive to continue an auditor-client relationship once established.\(^4\) That is, both audit firm and client tend to lose in economic terms if an established relationship is terminated. This creates a ‘lock in’ effect and provides the incentive for audit firms to enlarge their market shares in competing for initial audit engagements. As a result, owing to the competition among audit firms, the natural inference is that the existence of economic interests of an established relationship induces ‘low-balling’, i.e., audit firms bid below total auditing costs in their initial audit engagements. In this way, the phenomenon of vigorous price-cutting on initial audit engagements can be viewed as a competitive weapon utilized by audit firms seeking to achieve market dominance. Consequently, the competition among audit firms, both for the initial audit and at the time of recontracting, will govern the extent to which an incumbent audit firm can benefit from learning and switching costs. In fact, the costs of changing audit firms in the future period partially induce clients to continue using

\(^3\)The client’s switching costs may include search costs of finding a new audit firm, the costs of the additional time spent by management explaining the system to the new audit firm, and the costs of complying with regulation which mandates disclosure of the circumstances surrounding a change of audit firms, etc.. In addition, non-economic costs may also be incurred. For example, the client has to interact with an auditor whose style and personality is quite different from his incumbent auditor.

\(^4\)A relationship-specific economic interest is an asset that is non-marketable or non-transferable in transactions involving different trading partners other than the old ones. In this way, the existence of a relationship-specific economic interest creates a ‘lock in’ effect by making it costly to switch trading partners.
the audit firms they initially selected. As a result, clients display loyalty and audit firms have an incentive to raise the fees for their audit services in the future period. Thus, in a subgame perfect Nash equilibrium, the audit market is ‘less’ competitive with higher profits after the initial period. However, the dependence of future profits on the number of ‘locked-in’ clients also leads to ‘more’ competitive behaviour in the initial period (before clients have attached themselves to audit firms) than if there were no relationship-specific economic interests. This means that relationship-specific economic interests will change the structure of demand in the first period as well as the future period. In the first period, audit firms are more willing to cut their fees. In other words, in the audit market with relationship-specific economic interests, audit firms are willing to compete quite fiercely to build up a larger client base, i.e., they behave more competitively, in the first period. On the other hand, audit firms also have an incentive to exploit their previous clients, i.e., they behave less competitively, in the future period. Therefore, the effect of the presence of relationship-specific economic interests on overall competition is potentially ambiguous. Thus, one of the goals of the analysis in this chapter is to identify the conditions under which the overall competition is ‘excessive’ or ‘insufficient’ (from a social welfare perspective) in an audit market with learning and switching costs. Among other things, it is shown that while the relationship between learning and the strength of overall competition is monotonic, the relationship between switching costs and the strength of overall competition is not. Nevertheless, the analysis establishes the conditions under which overall competition will be further fortified when switching costs increase.

This chapter is related to the growing literature on audit pricing which is reviewed in chapter 1. Our work is distinguished from this area of research on the basis of its focus on an imperfect competitive audit market. The only published work in the auditing competition literature which also emphasizes market imperfections is a recent article by Gigler and Penno (1995). While Gigler and Penno look at audit firms who have substantial market power

---

5The Metcalf Staff Report (U.S. Senate 1976) argues that there is insufficient competition in the audit market, whereas the Cohen Commission Report (AICPA 1978) believes it to be excessive.
because of their stochastic endowment of different auditing costs, we examine a setting where audit firms are *ex-ante* identical and strategically choose to become differentiated (in terms of their audit production costs) by means of service specialization. The focus of the analyses is also very different. Gigler and Penno, following Magee and Tseng (1990), focus on the pricing contest between two audit firms for serving a single client. In contrast, our model is a market setting where audit firms compete for an infinite number of clients by offering client-specific audit fees. The spatial approach that we adopt explicitly recognizes the dispersed nature of the audit market, namely that it embodies a large number of clients with different characteristics relevant to audit production and relatively few audit suppliers who may differ in their area of service specialization with respect to client characteristics. It is a widely held notion that audit firm specialization is the primary source of the cost differences among audit firms. The cost differences, in turn, are believed to be the source of market power and, hence, the economic rents which may accrue to the audit firms. In this respect, we provide the first formal spatial model to examine how audit firms acquire market power by means of service specialization and the effect of audit firms' specializations on their audit pricing decisions.

In this model, the audit firms' learning is exogenous, but the audit firms' specializations as well as the clients' switching costs are controlled (at least partially) by the audit firms. The endogeneity of the audit firms' specializations and the clients' switching costs is not found in either the existing auditing competition models in the accounting literature or the existing switching cost models in the economics literature. In this framework, it is demonstrated that social welfare (the sum of audit firms' profits and clients' surplus) is influenced by the audit firms' specialization decisions, which in turn are influenced by the learning and switching costs. As such, it is possible to examine the welfare implications of changes in learning and switching costs. These welfare implications are absent in the extant accounting literature. It is because the social welfare is fixed in a pure pricing game with

---

6For discussion and models of switching costs see Klemperer (1987) and the references cited there.
inelastic demand. Any changes in the learning and switching costs will simply lead to a transfer of economic interest from one party to the other, and will not change the sum.

One of the conclusions in this chapter is that the audit market is less efficient in the presence of switching costs (i.e., social welfare is lower). This inefficiency is driven by the fact that, in the presence of relationship-specific economic interests created by switching costs, audit firms are able to relax price competition and achieve partial collusion by differentiating themselves through specialization of services. However, whether the switching costs, which are often cited as the source of audit firms’ economic rents, may actually increase the economic rents to the audit firms is not obvious. In fact, it may happen that increasing switching costs decreases the audit firms’ profits, and increases the benefit to clients. Such a case is possible if an increase in switching costs induces more aggressive pricing and specialization decisions of audit firms in order to compete for clients in the initial period. Then, audit firms may be worse off if the increased competition drives their first-period fees so low that their final profits are reduced even if their future-period rents increased by the increase in the switching costs. Clients are better off in this case since they pay lower audit fees owing to a more competitive audit market. This result is consistent with the finding of Gigler and Penno, although our conclusion is based on an intuition that is different from theirs.

The effect of the presence of learning on social welfare is equally subtle. On the one hand, the presence of learning has a direct effect of reducing total auditing costs, given audit firm specialization choice. On the other hand, it also has an indirect effect of distorting the audit firms’ specialization decisions. Therefore, the impact of learning on social welfare is potentially ambiguous. A similar argument applies to the effect of learning on the total profits to audit firms. An increase in learning allows each audit firm to reduce its total auditing costs to its clients in the future period. This in turn provides each audit firm an incentive to enlarge its own market share by pricing more aggressively in the initial period. While it is individually rational for each audit firm to do so, all audit firms taken together are made worse off by the increased competition. Unlike the cases of social welfare and audit
firms' profits, the effect of learning on clients' surplus is clear. Clients are better off because they pay lower audit fees owing to increased competition as learning increases.

The rest of the chapter is organized as follows. Section 3.1 presents the model, and defines a specialization-price game for the audit firms and an appropriate solution concept. Section 3.2 derives a duopoly equilibrium by first analyzing how the second-period price equilibrium depends on the first-period market shares. Knowledge of this dependence allows one to solve for the first-period price equilibrium, and hence the first-period specialization equilibrium and the outcome of the full game in section 3.3. Section 3.4 studies the implications of changes in auditing costs, learning rate, and switching costs. Section 3.5 concludes the chapter.

3.1 The Model

The model presented in the previous chapter is very general. In order to get some qualitatively stronger results, it is necessary to give the model more structure. To this end, the analysis in this chapter focuses on a simple two-period spatial duopoly model where client characteristics are distributed on a one-dimensional compact space. Formally, it is assumed that there is a common index function, \( G(z) : \mathbb{R}^N \rightarrow \mathbb{R} \), mapping a client's vector of characteristics \( z \) into a one-dimensional point \( z \), where for sake of simplicity, \( z \) is further assumed to be uniformly distributed along the line segment \([0, 1]\) with unit density. On the other hand, in order to capture salient economic features of an audit market with auditor-client relationship-specific economic interests, audit firms' learning and clients' cost

---

7The analysis becomes increasingly complex, if not unmanageable, as additional periods, audit firms or dimensions of client characteristics are added. Nevertheless, the basic economic forces that drive the results in a two-period duopoly model are believed to be present in a more general model.

8While such a distribution is empirically rare, it provides a setting in which interesting parametric variations can be investigated. Specifically, the assumption of a uniform distribution has the advantage of eliminating the effect of nonuniformity of distribution as a possible explanation of equilibrium specialization. A nonuniform distribution (e.g., unimodal or bimodal) may lead to agglomeration (i.e., audit firms choose the same specialization) or differentiation (i.e., audit firms choose different specializations), and confounds the effect of competition, which is what this chapter attempts to analyze.
of switching audit firms are introduced.

The audit market consists of two independent audit firms which bid on audit engagements and may only differ in specialization of services with respect to client characteristics. Let \( l_1 \) and \( l_2 \) denote the respective specializations of audit firm 1 and audit firm 2, where \( 0 \leq l_1 \leq l_2 \leq 1 \).\(^9\) The auditing technology is the same for each audit firm. It costs \( m_1(z) = c|l_i - z| \), \( c > 0 \) to audit a client \( z \in [0, 1] \) in the first period, where \( |l_i - z| \) measures the absolute difference between audit firm \( i \)'s specialization and client \( z \)'s characteristics.\(^10\) In the second period, because of learning by the incumbent audit firm, the auditing cost of audit firm \( i \) to client \( z \) is

\[
m_i^2(z) = \begin{cases} 
\beta c|l_i - z| & \text{if audit firm } i \text{ audited client } z \text{ in the first period}, \\
|l_i - z| & \text{otherwise},
\end{cases}
\]

where \( 0 < \beta < 1 \) characterizes the degree of learning of the incumbent audit firm (i.e., the lower the beta, the higher the learning rate, which is equal to \( 1 - \beta \)). The effect of \( \beta \) is to give the incumbent audit firms a comparative cost advantage over their rivals in the second period.

On the demand side, each client voluntarily acquires one unit of audit service from one of the audit firms in every period. If an audit firm is hired, the audit services it provides will give a client a gross benefit of \( b \) per period, where \( b > c(2 + \beta)/2 \).\(^11\) The cost of the audit to the client is an audit fee, \( f \). In the second period, in addition to an audit fee paid to an audit firm, a client has a switching cost \( k(l_2 - l_1) \), \( 0 < k < \beta c/2 \), of hiring an audit firm that he has not previously hired. The switching costs of a client are assumed to be proportional

\(^9\)This assumption imposes a coordination device concerning the ranking of the audit firms' specialization along the line segment \([0, 1]\). This device can be interpreted as a collusive rule which restricts the audit firms' strategy spaces. In the absence of this restriction the two audit firms find themselves in a coordination game. This results in an infinity of mixed strategy equilibria. See Bester et al (1991) for details.

\(^10\)A higher \( c \) implies that it is more difficult for an audit firm to 'adjust' its production process so as to service clients whose characteristics are different from its specialization. For example, suppose the line segment \([0, 1]\) represents the domain of the client industry. Then when \( c \) is high, 'nonspecialists' will find it more difficult to efficiently service clients whose industries do not fall into their area of specialization.

\(^11\)Actually, \( b \) just needs to be sufficiently large, so that every client would like to purchase the audit services. Notice also that \( b \) can be client-specific and time-dependent without affecting the qualitative results derived in this chapter.
to the difference between the specializations of the two audit firms. This assumption reflects
the fact that the switching costs of a client arise from the need to adapt from his incumbent
audit firm to a replacement audit firm.\textsuperscript{12}

The following assumption summarizes the specifications of the parameter values in this
model.

**Assumption (A1).** Let $c > 0$, $b > c(2 + \beta)/2$, $0 < k < \beta c/2$ and $0 < \beta < 1$.

Notice that the ranges chosen for $b$ and $k$ are sufficient to ensure that in equilibrium, even
in the presence of relationship-specific economic interests in clients, the whole audit market is
covered and either audit firm can compete for the whole audit market in both periods. It also
implies that audit firms cannot charge the monopoly audit fee to their previous purchasers
in the second period.

Clients are assumed to have rational expectations in the sense that they foresee at any
time the equilibrium of the rest of the game and behave accordingly. The one-period surplus
to client $z$ at an audit fee $f_{it}^z$, if he hires audit firm $i$ in period $t$ is given by

$$S_1^*(f_{it}^z) = b - f_{it}^z,$$

$$S_2^*(f_{it}^z) = \begin{cases} b - f_{it}^z & \text{if client } z \text{ hired audit firm } i \text{ in the first period,} \\ b - f_{it}^z - k(l_2 - l_1) & \text{otherwise.} \end{cases}$$

Each period, the client chooses which audit firm to hire given the audit fees offered to him
and the transaction costs of switching audit firms.

On the supply side, in order to induce a client to accept an offer, the audit fee must
not exceed the maximum amount the client would be willing to pay. However, audit firms

\textsuperscript{12}It is assumed that clients face no adaption cost with the initial audit engagement. In fact, there is no
change in the analysis if such an adaption cost is independent of the audit firm’s specialization (it can be
part of $b$). The analysis would change if the adaption cost was specialization-specific (which is the case
with the switching cost). However, it can be shown that there would be no qualitative change in the results
obtained in this chapter. Above all, incorporating a client’s adaption cost in the initial period does not give
any additional insights to the analysis. Hence, it is omitted for simplicity.
are not restricted to price at or above their marginal auditing costs in this model.\textsuperscript{13} In fact, an audit firm might price below its marginal auditing cost if, say, audit firm $i$ were confident that audit firm $j$ would undercut its audit fee in the initial period and thus serve the client in question. Such behaviour is of no direct benefit to audit firm $i$, but serves to force audit firm $j$ to charge a lower audit fee. Therefore, this behaviour might be strategically important to audit firm $i$ in attempting to discourage audit firm $j$ from choosing a particular specialization. Nevertheless, the range chosen for $k$ ensures that, in equilibrium, audit firms always charge positive audit fees. Hence, in each period audit firms offer audit fees from the feasible range $[0, b]$ to each client. Suppose that the audit fees offered to a client $z$ in period $t$ by audit firms $i$ and $j$ are $f_{it}^z$ and $f_{jt}^z$, respectively. If the client purchases from audit firm $i$, then the one-period profit for audit firm $i$ earned from this client is equal to

$$
\Pi_{it}^z(l_i, l_j, f_{it}^z, f_{jt}^z) = f_{it}^z - m_{it}^z(l_i).
$$

Without loss of generality, it is assumed that there is no discounting to the second-period revenues (benefits) and costs.

Formally, the setting is a two-period, three-stage complete information game with the first two stages occurring in the first period and the third stage occurring in the second period. In the first stage of the game (the specialization stage), the audit firms simultaneously choose their specializations in the line segment $[0, 1]$. In the second stage (the first-period pricing stage), the audit firms simultaneously quote audit fees to each client. Given the first-period audit fee schedules set by the two audit firms, clients make their audit firm choices. At the beginning of the second period, auditor-client relationships have been established. Learning is realised by the incumbent audit firms and transactions costs must be incurred if the clients choose to switch audit firms. As a result, in the third stage (the second-period pricing stage), both audit firms simultaneously quote audit fees which reflect the effects of learning and switching costs. Clients have rational expectations. At each point of time,

\textsuperscript{13}This restriction and its welfare implications are examined in chapter 4.
client \( z \) acquires auditing services from the audit firm giving him the highest, nonnegative (expected) total surplus.

Once again, the equilibrium concept employed is Selten’s (1975) subgame perfect Nash equilibrium (SPNE) and attention is restricted to pure strategy equilibria only. For the following analysis, a variable with an asterisk denotes an SPNE strategy for the subgame in question. Moreover, the following tie-breaking rules are adopted for simplifying the analysis: (i) a client will patronize the audit firm with the lower auditing cost if he is indifferent between the audit firms’ offers in the first period; and (ii) a client will stay with his incumbent audit firm if he is indifferent between his incumbent audit firm’s offer and that of the replacement audit firm in the second period. A strategy for audit firm \( i \) is a triplet \( (l_i, F_{1i}, F_{2i}) \), that consists of the audit firm’s specialization, \( l_i \), a function \( F_{1i} \), mapping every possible observed combination of \( l_1 \) and \( l_2 \) into a sequence of first-period audit fee schedules, \( \{f_{1i}^x\}_{x \in [0, 1]} \), and a function, \( F_{2i} \), mapping every possible observed combination of \( (l_1, l_2, F_{11}, F_{21}) \) into a sequence of second-period audit fee schedules, \( \{f_{2i}^x\}_{x \in [0, 1]} \), where \( f_{it}^x \) is the audit fee schedule quoted to client \( z \) by audit firm \( i \) in period \( t \). A strategy for client \( z \) is a pair, \( (Q_1^z, Q_2^z) \), that consists of functions \( Q_i^z \) that map every strategy of \( f_{1i} \) and \( f_{2i} \) into \( \{1, 2\} \) in period \( t \), where \( \{1, 2\} \) is the set of audit firms in the market.

3.2 Analysis of Pricing Subgames

This section finds and characterizes the SPNE for subgames starting from stages 2 to 3. It is demonstrated that there exists a symmetric equilibrium in which all clients to the left (right) of \( (l_1 + l_2)/2 \) purchase from audit firm 1 (2) and clients do not switch audit firms in equilibrium.\(^{14}\) Then, the symmetric equilibrium will be proven to be the unique equilibrium for the game given the restriction on the audit firms’ strategies on specialization, i.e., \( 0 \leq l_1 \leq l_2 \leq 1 \). This conjecture can be rationalized by noting that audit firms have the

\(^{14}\)Looking for a symmetric equilibrium seems the natural procedure given the symmetric structure of the game.
same audit cost functions before clients have established relationships with them in the first period, and that audit firm 1 (2) has a cost advantage over the other audit firm with respect to all clients whose characteristics are on the left (right) of \((l_1 + l_2)/2\). This allows audit firm 1 (2) to offer a strictly better audit fee schedule to all clients whose characteristics are on the left (right) of \((l_1 + l_2)/2\). Unless audit firms choose the same specialization, the set of clients whose characteristics are ‘equidistant’ from both audit firms is of measure zero.

As stated in the previous chapter, since audit firms possess some monopoly power by means of specialization and clients cannot resell audit contracts, the two audit firms are able to price discriminate across clients as they take into consideration the heterogeneity (by specialization and purchase history) among clients.\(^{15}\) Thus, in each period, each client receives a pair of audit fee schedules, which depends on his characteristics and his relationship to the two audit firms, and chooses which audit firm to hire. Given complete information, rational players in the game accurately anticipate all actions taken by all the other players.

Under the configuration described above, the first-period audit fee schedules \(\{f^*_1\}_{z \in [0, 1]}\) and \(\{f^*_2\}_{z \in [0, 1]}\) result in first-period profits \(\Pi_{11}\) and \(\Pi_{21}\) and market segments \([0, (l_1 + l_2)/2]\) and \(((l_1 + l_2)/2, 1]\).\(^{16}\) It then follows that audit firms’ second-period choices of fee schedules \(\{f^*_1\}_{z \in [0, 1]}\) and \(\{f^*_2\}_{z \in [0, 1]}\) and their second-period profits \(\Pi_{12}\) and \(\Pi_{22}\) depend on these market segments. As usual, this dependence must be examined first since to compute the first-period equilibrium one must know how future profits depend on first-period market segments.

3.2.1 The Second-Period Price Equilibrium

This subsection analyzes the second period of an audit market with the presence of learning and switching costs, given auditor-client relationships established in the first period. The audit firms’ optimal fee schedules are computed as functions of their first-period market

\(^{15}\)The purchase history of a client tells which audit firm audited the client in the previous period.

\(^{16}\)Without loss of generality, the client at \((l_1 + l_2)/2\) is assigned to audit firm 1 for technical convenience.
segments, which are characterized by their specializations in the first period. More formally, consider the second-period pricing stage under a subgame defined by \((l_1, l_2, F_{11}, F_{21})\). The two audit firms are able to set discriminatory audit fees according to a client’s purchase history and characteristics. Since the economy only lasts for two periods, the second-period audit fee schedule \(f^*_2\), which specifies the audit fee at which audit firm \(i\) is willing to supply audit services to a client at \(z \in [0, 1]\), must cover the unit marginal auditing cost to the client. Again, in order to induce a client to accept an offer, the audit fee must not exceed the maximum amount the client would be willing to pay. Therefore, in equilibrium, \(f^*_2\) must be in the range \([c|l_i - z|, 6]\) for all \(z \in [0, 1]\).

As stated before, the cost advantage of audit firm 1 (2) over the other audit firm with respect to all clients whose characteristics are on the left (right) of \((l_1 + l_2)/2\) in the first period ensures that, in equilibrium, all clients in the interval \([0, (l_1 + l_2)/2]\) \(((l_1 + l_2)/2, 1]\) hire audit firm 1 (2) in the first period. As such, clients will switch from the incumbent audit firm to the replacement audit firm in the second period if, and only if, the difference in client surplus from the two audit fees is strictly greater than the client’s cost of switching. By assumption, an indifferent client stays with the incumbent audit firm. Hence, clients who patronized audit firm 1 will stay with it in the second period if \(f^*_1 \leq f^*_2 + k(l_2 - l_1)\). Similarly, none of the clients who bought from audit firm 2 in the first period will purchase from audit firm 1 in the second period if \(f^*_1 \leq f^*_2 + k(l_2 - l_1)\).

Now, consider any client \(z\) in the interval \([0, (l_1 + l_2)/2]\) who bought from audit firm 1 in the first period. In the second period, to any client \(z\) in the interval \([0, (l_1 + l_2)/2]\), the most favourable audit fee schedule that audit firm 2 can offer, subject to it at least breaking even, is \(f^*_2 = c(l_2 - z)\). The surplus of client \(z\) if he accepts the offer is then given by

\[
S^*_2(f^*_2) = b - c(l_2 - z) - k(l_2 - l_1) > 0,
\]

for all \(c > 0, b > c(2 + \beta)/2, 0 < k < \beta c/2, 0 < \beta < 1\) and \(0 < l_1 \leq l_2 \leq 1\).

Being an incumbent audit firm for any client \(z \in [0, (l_1 + l_2)/2]\) in the second period,
audit firm 1 must give client \( z \) a surplus no less than \((3.1)\) to induce him to stay with it. Thus, the optimal audit fee schedule quoted by audit firm 1 to client \( z \) solves: \((P3.1)\)

\[
\begin{align*}
\max_{f_{12}^*} & \quad \Pi_{12}^*(l_1, l_2, f_{12}^*, f_{22}^*) \\
\text{s.t.} & \quad S^*_2(f_{12}^*) \geq S^*_2(f_{22}^*). 
\end{align*}
\] (3.2) (3.3)

It is not difficult to show that \((3.3)\) is binding and the unique maximum solution for \((P3.1)\) is

\[ f_{12}^* = c(l_2 - z) + k(l_2 - l_1). \]

That is, in the second-period price equilibrium, the incumbent audit firm (i.e., audit firm 1) matches the rival’s audit fee such that the difference in client surplus from the two audit fees exactly equals the client’s switching costs. In other words, audit firm 1 offers to a client \( z \in [0, (l_1 + l_2)/2] \) an audit fee that makes him indifferent between staying and switching; and by assumption, the client stays. Hence, in the second period, audit firm 1 would have its audit fee for client \( z \in [0, (l_1 + l_2)/2] \) constrained by the marginal auditing cost of audit firm 2 and the switching costs of the client. Accordingly, the profit that audit firm 1 earns by offering this audit fee schedule to client \( z \) in the second period is

\[
\Pi_{12}^*(l_1, l_2, f_{12}^*, f_{22}^*) = c(l_2 - z - \beta |l_1 - z|) + k(l_2 - l_1) > 0,
\]

for \( l_1 \leq l_2, 0 < k < \beta c/2 \) and \( 0 < \beta < 1 \).

Notice that the profit that audit firm 1 earns from client \( z \) in the second period is composed of two basic elements: (i) the difference between the auditing costs of servicing client \( z \) by the incumbent audit firm and that of the replacement audit firm; and (ii) the switching costs that the incumbent audit firm extracts from client \( z \). Other things being equal, an increase in the auditing costs (as \( c \) increases), the learning (as \( \beta \) decreases), or the switching costs (as \( k \) increases) increases the second-period profit that audit firm 1 can earn from client \( z \in [0, (l_1 + l_2)/2] \). The effects of the latter two factors on the second-period profit are very easy to understand. Only the effect of the first factor needs some
Audit firms’ profits (both in the first-period and the second-period) rise with \( c \) because an increase in the cost parameter increases the barrier to competition and results in higher profit mark-ups earned by the audit firms.\(^{17}\) The key here is that the audit production cost is proportional to \( c \) and the learning benefit for any given client is also proportional to \( c \). The former would hold even if the learning benefit was additive instead of multiplicative. Similarly, one can calculate the equilibrium pairs of audit fee schedules for all clients in the interval \( ((l_1 + l_2)/2, 1] \). A complete solution for the second-period pricing stage, \( (F_{12}, F_{22}, \{Q^z_2\}_{z\in[0,1]}) \), is then obtained. In the second-period price equilibrium, the incumbent audit firm must price at the marginal auditing cost of its competitor plus the client’s switching costs on services to clients whose characteristics are in the proximity of its own specialization. Hence, the unique second-period SPNE audit fee schedules of audit firm 1 and 2, respectively, are given by, for \( z \in [0, 1] \),

\[
\begin{align*}
&f_{12}^{*z}(l_1, l_2) = \begin{cases} 
  c(l_2 - z) + k(l_2 - l_1) & \text{if } 0 \leq z \leq (l_1 + l_2)/2, \\
  c(z - l_1) & \text{if } (l_1 + l_2)/2 < z \leq 1,
\end{cases} \\
&f_{22}^{*z}(l_1, l_2) = \begin{cases} 
  c(l_2 - z) & \text{if } 0 \leq z \leq (l_1 + l_2)/2, \\
  c(z - l_1) + k(l_2 - l_1) & \text{if } (l_1 + l_2)/2 < z \leq 1.
\end{cases}
\end{align*}
\]

(Figure 1 about here)

The second-period equilibrium audit fee schedule is illustrated in figure 1 where it is represented by the heavy line. The audit market is segmented at \( (l_1 + l_2)/2 \); audit firm 1 (2) serves segment \( [0, (l_1 + l_2)/2] \) \( ((((l_1 + l_2)/2, 1]) \) at audit firm 2’s (1’s) marginal auditing cost plus the client’s switching costs. Over the interval \( [l_1, l_2] \) the second-period equilibrium audit fee to client \( z \) falls as the difference between client \( z \)’s characteristics and his incumbent audit firm’s specialization rises since the incumbent audit firm has to meet the competition.\(^{18}\)

Notice that clients are loyal to their incumbent audit firms in the second period because of

\(^{17}\)It is interesting to point out that if the cost parameter \( c \) is firm-specific and audit firms can do something to affect their own cost, then while it may be individually rational for an audit firm to decrease its own cost, all audit firms taken together may be made worse off by the decreased profit mark-ups.

\(^{18}\)Notice that the interval \( [l_1, l_2] \) is the market segment where services supplied by the two audit firms are deemed as ‘close substitutes’ to the clients in that segment.
the presence of clients’ switching costs. This in turn weakens price competition between the audit firms. Thus, given the audit firms’ specializations, switching costs make the outcome more collusive in the second period.

3.2.2 The First-Period Price Equilibrium

Now, go back one stage and consider the subgame defined by \((l_1, l_2)\) in the first period. After the two audit firms have chosen their specialization simultaneously, each audit firm sets its fee schedules while taking into account not only the effect on its first-period profitability, but also the effect on its first-period market segment and hence its second-period profitability.

Similar to the second-period pricing stage, in the first period each client receives a pair of audit fee schedules which depend on his characteristics relative to the two audit firms’ specializations (although there is no client’s purchase history in the first period on which to condition the audit fee). Every client has rational expectations and accepts the most favourable offer which gives him the highest, nonnegative total two-period (expected) surplus. Thus, given a pair of audit fee schedules in the first period, each client wants to maximize his total (expected) surplus over the two periods. Again, since audit firm 1 (2) has a cost advantage over the other audit firm with respect to all clients whose characteristics are on the left (right) of \((l_1 + l_2)/2\) in the first period, in equilibrium, all clients whose characteristics are on the left (right) of \((l_1 + l_2)/2\) purchase from audit firm 1 (2).

It is worthwhile mentioning that, because of the presence of learning and transactions costs of switching audit firms in the second period, there exists a nonempty interval of clients inside \([l_1, (l_1 + l_2)/2]\) \([(l_1 + l_2)/2, l_2]\) who will stay with audit firm 2 (1) if they purchased from it in the first period. To see this, suppose there exists a client \(z\) who is in the interval \((l_1, (l_1 + l_2)/2)\) and bought from audit firm 2 in the first period. The client will stay with audit firm 2 (his incumbent audit firm) in the second period if \(f_{22}^z \leq f_{12}^z + k(l_2 - l_1)\). Since by assumption client \(z\) will stay with his incumbent audit firm if he is indifferent
between his incumbent audit firm's offer and that of the rival audit firm in the second period, then following the standard Bertrand argument, audit firm 2 will optimally offer 
\[ f_{22}^* = c(z - l_1) + k(l_2 - l_1) \] to client z. It is because the best audit fee schedule for client z that audit firm 1 could offer without suffering a loss is \( c(z - l_1) \). Audit firm 2 offers client z the same amount of surplus using the above audit fee. Any higher audit fee would mean audit firm 2 loses client z whereas any lower audit fee would mean it gives up potential profits.

Needless to say, it is only rational for audit firm 2 to make such an offer to client z if its second-period audit fee can cover its second-period auditing cost, i.e., \( \beta c(l_2 - z) \leq c(z - l_1) + k(l_2 - l_1) \). Let \( z^1 \equiv l_1 + \frac{\beta c - k}{1 + \beta} c(l_2 - l_1) \). Then \( f_{22}^* = c(z - l_1) + k(l_2 - l_1) \) is the optimal second-period audit fee that audit firm 2 could offer to client z only if \( z \in [z^1, (l_1 + l_2)/2] \).

On the other hand, if \( z \in [0, z^1) \), \( \beta c(l_2 - z) > c(z - l_1) + k(l_2 - l_1) \). Then, the best second-period audit fee that audit firm 2 could offer to client z is its marginal auditing cost, i.e., \( f_{22}^* = \beta c(l_2 - z) \). In this case, audit firm 2 earns zero profits from client z in the second period.

The next step is to determine the audit fee \( f_{21}^* \) that audit firm 2 would like to offer to client \( z \in [0, (l_1 + l_2)/2] \) in the first period. At first, notice that the second-period profit that audit firm 2 can earn from any client \( z \in [0, (l_1 + l_2)/2] \) is

\[
\Pi_{22}^z = \begin{cases} 
0 & \text{if } 0 \leq z < z^1, \\
c(z - l_1) + k(l_2 - l_1) - \beta c(l_2 - z) & \text{if } z^1 \leq z \leq (l_1 + l_2)/2.
\end{cases}
\]

Consequently, in the first period, to any client z in the interval \([0, (l_1 + l_2)/2]\), the most favourable audit fee schedule that audit firm 2 can offer, subject to it at least breaking even, is

\[
f_{21}^z = \begin{cases} 
c(l_2 - z) & \text{if } 0 \leq z < z^1, \\
c(l_2 - z) - [c(z - l_1) + k(l_2 - l_1) - \beta c(l_2 - z)] & \text{if } z^1 \leq z \leq (l_1 + l_2)/2.
\end{cases}
\]

\(^{19}z^1\) is obtained by setting \( \beta c(l_2 - z^1) = c(z_1 - l_1) + k(l_2 - l_1) \). It is easy to verify that \( z^1 \in (l_1, (l_1 + l_2)/2) \) for all \( c > 0, 0 < k < \beta c/2 \) and \( 0 < \beta < 1 \).
i.e., in the first period audit firm 2 is willing to offer an audit fee to client \( z \in [0, (l_1 + l_2)/2] \) which is so low that its second-period potential profit from client \( z \) is exactly cancelled. Clearly, \( f_{21}^* < c(l_2 - z) \) for \( z \in [z^1, (l_1 + l_2)/2] \). It means that in order to attract client \( z \in [z^1, (l_1 + l_2)/2] \) to patronize it, audit firm 2 is willing to turn over its second-period potential profit to the client in the form of an initial discount.

Given the above audit fee schedules, a rational client \( z \in [0, (l_1 + l_2)/2] \) who purchases from audit firm 2 in the first period will have a total two-period (expected) surplus of

\[
S_1^z + S_2^z = 2b - c(1 + \beta)(l_2 - z) > 0. \tag{3.4}
\]

Hence, in the first period, looking ahead to the second-period equilibrium audit fees, audit firm 1 must give him a total two-period (expected) surplus no less than (3.4) to induce client \( z \in [0, (l_1 + l_2)/2] \) to patronize it. Thus, the optimal first-period audit fee schedule quoted by audit firm 1 to client \( z \in [0, (l_1 + l_2)/2] \) solves: (P3.2)

\[
\max_{f_{11}^z} \quad \Pi_{11}^z(l_1, l_2, f_{11}^z, f_{21}^{**}) + \Pi_{12}^z(l_1, l_2, f_{11}^z, f_{21}^{**}) \tag{3.5}
\]

\[
\text{s.t.} \quad S_1^z(f_{11}^z) + S_2^z(f_{12}^{**}) \geq S_1^z + S_2^z. \tag{3.6}
\]

It is not difficult to show that (3.6) is binding and the unique maximum solution for (P3.2) is

\[
f_{11}^{**} = \begin{cases} 
  c(l_2 - z) & \text{if } 0 \leq z < z^1, \\
  \beta c(l_2 - z) - k(l_2 - l_1) & \text{if } z^1 \leq z \leq (l_1 + l_2)/2.
\end{cases}
\]

Notice that for all \( c > 0, 0 < k < \beta c/2 \) and \( 0 < \beta < 1, f_{11}^{**} > 0 \) for all \( z \in [0, (l_1 + l_2)/2] \).

Accordingly, the profit that audit firm 1 earns by offering this audit fee schedule to client \( z \) in the first period is

\[
\Pi_{11}^z(l_1, l_2, f_{11}^{**}, f_{21}^{**}) = \begin{cases} 
  c(l_2 - z - |l_1 - z|) & \text{if } 0 \leq z < z^1, \\
  c[\beta(l_2 - z) - (z - l_1)] - k(l_2 - l_1) & \text{if } z^1 \leq z \leq (l_1 + l_2)/2.
\end{cases}
\]

Similarly, one can calculate the equilibrium pairs of audit fee schedules for all clients in the interval \(((l_1 + l_2)/2, 1]\). A complete solution for the first-period pricing stage,
The unique first-period SPNE audit fee schedules of audit firm 1 and 2, respectively, are given by

\[
\begin{align*}
\text{\textit{f}}_{11}^*(l_1, l_2) &= \begin{cases} 
  c(l_2 - z) & \text{if } 0 \leq z < z^1, \\
  \beta c(l_2 - z) - k(l_2 - l_1) & \text{if } z^1 \leq z \leq (l_1 + l_2)/2, \\
  c[(1 + \beta)(z - l_1) - (l_2 - z)] - k(l_2 - l_1) & \text{if } (l_1 + l_2)/2 < z \leq z^2, \\
  c(z - l_1) & \text{if } z^2 < z \leq 1,
\end{cases} \\
\text{\textit{f}}_{21}^*(l_1, l_2) &= \begin{cases} 
  c(l_2 - z) & \text{if } 0 \leq z < z^1, \\
  c[(1 + \beta)(l_2 - z) - (z - l_1)] - k(l_2 - l_1) & \text{if } z^1 \leq z \leq (l_1 + l_2)/2, \\
  \beta c(z - l_1) - k(l_2 - l_1) & \text{if } (l_1 + l_2)/2 < z \leq z^2, \\
  c(z - l_1) & \text{if } z^2 < z \leq 1,
\end{cases}
\end{align*}
\]

where \( z^2 = l_2 - \frac{\beta c - k}{1 + \beta} l_1 \) is obtained by setting \( \beta c(z^2 - l_1) = c(l_2 - z^2) + k(l_2 - l_1) \).\textsuperscript{20}

The first-period equilibrium audit fee schedule is illustrated in figure 2 where it is represented by the heavy line. Again, the audit market is segmented at \((l_1 + l_2)/2\) and over the interval \([l_1, l_2]\) the first-period equilibrium audit fee to client \( z \) falls as the difference between client \( z \)'s characteristics and the supplying audit firm's specialization rises. Audit firm 1 (2) serves segment \([0, z^1]\) \(([z^2, 1]\)) at audit firm 2's (1's) auditing cost, which is greater than its own auditing cost. However, over the interval \([z^1, z^2]\) the first-period equilibrium audit fee to client \( z \) is below the auditing cost of the supplying audit firm since the audit firm has to meet the competition. That is, low-balling occurs only over the interval \([z^1, z^2]\) where competition between audit firms is quite keen.\textsuperscript{21} It is also easy to show that the interval \([z^1, z^2]\) increases as \( \beta \) decreases and/or \( k \) increases. The reason is that market share is more valuable in the second period in the presence of relationship-specific economic interests created by learning and switching costs. Thus, given specializations of audit firms, each audit firm competes more aggressively than it otherwise would to capture that market share. As a result, both audit firm 1 and 2 choose lower first-period audit fees. Moreover, an increase in learning or switching costs has the effect that the competitor will anticipate higher future profits if

\textsuperscript{20}It is easy to verify that for all \( c > 0, 0 < k < \beta c/2 \) and \( 0 < \beta < 1, z^2 \in ((l_1 + l_2)/2, l_2) \) and \( f_{11}^{**}, f_{21}^{**} > 0 \) for all \( z \in [0, 1] \).

\textsuperscript{21}A detailed discussion of low-balling is deferred to chapter 4.
it manages to attract the client. Higher future profits imply that the competitor can stand a lower audit fee today. Therefore, this drives the supplying audit firm’s fee downwards. As a result, audit firms’ first-period profits are lower if learning and/or switching costs increase. Recall that audit firms’ second-period profits increase with an increase of the learning and/or switching costs. Thus, the effects of learning and switching costs on audit firms’ total two-period profits are potentially ambiguous. A clear conclusion cannot be reached without considering the corresponding effects on audit firms’ specialization decisions.

3.3 Equilibrium for the Full Game

In the specialization stage, audit firms choose their specializations looking ahead to the outcomes of the pricing stages derived in the previous section. Using the results of the pricing stages, audit firm 1’s total two-period profit is given by

$$
\Pi_1(l_1, l_2) = \Pi_{11}(l_1, l_2, F^*_1, F^*_2) + \Pi_{12}(l_1, l_2, F^*_1, F^*_2) \\
= \int_0^{z_1} c(l_2 - z - |l_1 - z|) \, dz + \int_{z_1}^{1+\frac{\alpha}{2}} \{c[\beta(l_2 - z) - (z - l_1)] - k(l_2 - l_1)\} \, dz \\
+ \int_0^{1+\frac{\alpha}{2}} [c(l_2 - z - \beta|l_1 - z|) + k(l_2 - l_1)] \, dz,
$$

since $q^*_{1t} = 0$ for all $z \in ((l_1 + l_2)/2, 1]$ and $t = 1, 2$. By the same token, one can obtain audit firm 2’s total two-period profit, which is symmetric to that of audit firm 1. In the specialization stage, each audit firm will choose its specialization to maximize its total two-period profit given its rival’s specialization choice.

Before proceeding to solve the SPNE choice of audit firm specializations, it is helpful to define some useful terms. Similar to chapter 2, let define $\Pi(l_1, l_2)$, $C(l_1, l_2)$, $S(l_1, l_2)$ and $W(l_1, l_2)$, respectively, as the total profit to both audit firms, the total costs for auditing services, the aggregate surplus to clients and the social welfare given an audit firm specialization pair $(l_1, l_2)$. Then, under the equilibrium audit fee schedules, they are given by

$$
\Pi(l_1, l_2) = \Pi_1(l_1, l_2) + \Pi_2(l_1, l_2)
$$
\[ C(l_1, l_2) = \int_0^{\frac{1 + l_2}{2}} (1 + \beta) c|l_1 - z| \, dz + \int_{\frac{1 + l_2}{2}}^1 (1 + \beta) c|l_2 - z| \, dz, \]
\[ S(l_1, l_2) = \int_0^{\frac{1 + l_2}{2}} (b - f_{11}^* \ast) \, dz + \int_{\frac{1 + l_2}{2}}^1 (b - f_{21}^* \ast) \, dz \]
\[ + \int_0^{\frac{1 + l_2}{2}} (b - f_{12}^* \ast) \, dz + \int_{\frac{1 + l_2}{2}}^1 (b - f_{22}^* \ast) \, dz, \]
\[ W(l_1, l_2) = \Pi(l_1, l_2) + S(l_1, l_2) \]
\[ = 2b - C(l_1, l_2). \]

It is easy to show that \( W(l_1, l_2) \) is continuous and strictly concave in \((l_1, l_2)\) (or, equivalently, \( C(l_1, l_2) \) is continuous and strictly convex in \((l_1, l_2)\)). Then, it follows immediately from the results of propositions 2.1 and 2.2 that a specialization pair that maximizes the social welfare also minimizes the total costs for auditing services, and such a specialization pair always exists. The next proposition states the unique welfare-maximization (or equivalently, cost-minimization) audit firm specialization pair.

**Proposition 3.1.** Given \((A1)\), the audit firm specialization pair which maximizes the social welfare is \((l_{1W}, l_{2W}) = (1/4, 3/4)\).

Because of the assumption of perfectly inelastic demands for audit services and uniform distribution of client characteristics, social welfare is maximized when the pair of audit firm specializations is \((l_{1W}, l_{2W})\). In this case, the two audit firms cooperatively choose symmetric specializations, each at a difference of one-fourth from the middle of the whole market segment, and split the market in half. The maximum difference between the characteristics of any client and the specialization of the supplying audit firm is only one-fourth of the total market space. As such, by choosing the specializations \((l_{1W}, l_{2W})\), the audit firms maximize the social welfare. However, these welfare-maximization specializations are not likely to be sustainable as the outcome of a noncooperative equilibrium.

The next proposition shows the SPNE choice of audit firm specializations. The SPNE
choice of audit firm specializations is a pair \((l_1^*, l_2^*)\) such that each audit firm's profit is maximized given its rival’s specialization. It is demonstrated that, in the presence of learning and switching costs, the noncooperative specialization choices that maximize audit firms' profits are in general different from the ones that maximize social welfare.

**Proposition 3.2.** Given \((A1)\), the unique symmetric SPNE choice of audit firm specializations, \((l_1^*, l_2^*)\), is given by

\[
l_1^* = 1 - l_2^* = \frac{(3 + \beta)(c^2 + \beta^2 c^2 + 2ck - 2\beta ck + 2k^2)}{2(4c^2 + 4\beta c^2 + 6\beta^2 c^2 + 7\beta^3 c^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2)}.
\]

It is clear that given the equilibrium audit fee schedules, audit firms never choose an identical specialization in equilibrium. The reason is that when audit firms choose the same specialization, profits are driven down by intense price competition.\(^{22}\) Anticipating this outcome, audit firms will never choose an identical specialization. Therefore, Bertrand competition drives audit firms to disperse in order to earn positive profits. In other words, in an SPNE specialization-price equilibrium, audit firms have a tendency to differentiate themselves in order to relax price competition. To see this, notice that the equilibrium audit firm specialization choices are a consequence of conflicts among three effects: the market-share effect, the strategic effect, and the cost effect. The market-share effect induces each audit firm to choose a specialization that is closer to the middle of the market so as to enlarge its market share. The strategic effect, on the other hand, gives an incentive to each audit firm to differentiate itself from the other so as to soften the price competition with its rival.\(^{23}\) The cost effect induces each audit firm to specialize so as to minimize its total costs. As a result, audit firms choose distinct specializations in the interior of the market segment \([0, 1]\).

\(^{22}\)It is easily shown that if \(\beta = 1\), Bertrand competition will drive profits to both audit firms to zero when audit firms choose the same specialization.

\(^{23}\)The movement towards audit market agglomeration (differentiation) decreases (increases) an audit firm’s profit mark-up but increases (decreases) its market share, given the specialization of its competitor.
Now, let us define the equilibrium specialization difference, $l_2^* - l_1^*$, as the degree of competitiveness of the audit market. As a benchmark, the welfare-maximization equilibrium specialization difference is equal to $l_2^W - l_1^W$. Given this, we have the following definition.

**Definition:** From a social welfare perspective, competition in the audit market is 'excessive' (‘insufficient’) if, and only if, $l_2^* - l_1^* < (>) l_2^W - l_1^W$, or equivalently, $l_1^* > (<) l_1^W$. \(^{24}\)

In the spatial framework, competition can be excessive if an increase in total auditing costs (as a result of an undesirable pair of audit firm specializations) exceeds an increase in aggregate surplus to clients. It happens when the SPNE audit firm specializations, $(l_1^*, l_2^*)$, are closer together compared to the welfare-maximization ones, $(l_1^W, l_2^W)$. On the other hand, competition can be insufficient if an increase in total auditing costs comes along with a decrease in aggregate surplus to clients. It is the case when the SPNE audit firm specializations are more distinct from each other compared to the welfare-maximization ones.

The next corollary provides some preliminary insights about the roles of learning and switching costs in determining the equilibrium audit firm specializations. A more detailed analysis and discussion is deferred to the next section.

**Corollary 3.1.** Given (A1), then (i) as $\beta$ approaches one in the limit, competition in the audit market is insufficient, i.e., $\lim_{\beta \to 1} l_1^* < l_1^W$; (ii) as $k$ approaches zero in the limit, competition in the audit market is excessive, i.e., $\lim_{k \to 0} l_1^* > l_1^W$; (iii) as $\beta$ approaches one and $k$ approaches zero in the limit, the SPNE audit firm specializations are the same as the welfare-maximization specializations.

Corollary 3.1 demonstrates that the presence of learning brings in the market-share effect and the presence of switching costs brings in the strategic effect. Hence, in the absence of switching costs (learning), i.e., as $k$ approaches zero ($\beta$ approaches one) in the limit, \(^{24}\)Since $l_2^* = 1 - l_1^*$ and $l_2^W = 1 - l_1^W$ in equilibrium.
the equilibrium audit firm specialization choices are a consequence of conflicts between the market-share (strategic) effect and the cost effect. As a result, audit firms choose distinct specializations inside (outside) the welfare-maximization specializations \((l_1^W, l_2^W)\). When \(\beta\) is very close to one and \(k\) is very close to zero, the market-share effect and the strategic effect are negligible. Then the cost effect alone dictates the welfare-maximization specializations.

Finally, given the equilibrium audit firms' profit-maximization specializations \((l_1^*, l_2^*)\), where \(l_2^* = 1 - l_1^*\), the total two-period profit to both audit firms in the duopoly auditing market is given by

\[
\Pi^*(l_1^*) \equiv \Pi(l_1^*, 1 - l_1^*) = \frac{1}{2c(1 + \beta)^2} \left[ c^2 + 7\beta c^2 + \beta^2 c^2 - \beta^3 c^2 - 4ck + 8\beta ck + 4\beta^2 k - 6k^2 - 2\beta k^2 \right. \\
\left. + 4l_1^*(c^2 - 3\beta c^2 + \beta^2 c^2 + \beta^3 c^2 + 5ck - 6\beta ck - 3\beta^2 ck + 6k^2 + 2\beta k^2) \right. \\
- 2l_1^2(5c^2 + \beta c^2 + 7\beta^2 c^2 + 3\beta^3 c^2 + 12ck - 8\beta ck - 4\beta^2 ck + 12k^2) \\
+ 4\beta k^2 \right],
\]

(3.7)

and the corresponding aggregate surplus to clients and social welfare, respectively, are given by

\[
S^*(l_1^*) \equiv S(l_1^*, 1 - l_1^*) = 2b - \frac{1}{4c(1 + \beta)^2} \left[ 3c^2 + 17\beta c^2 + 5\beta^2 c^2 - \beta^3 c^2 - 8ck + 16\beta ck \\
+ 8\beta^2 ck - 12k^2 - 4\beta k^2 + 4l_1^*(c^2 - 9\beta c^2 - \beta^2 c^2 + \beta^3 c^2 + 10ck - 12\beta ck \right. \\
- 6\beta^2 ck + 12k^2 + 4\beta k^2) - 4l_1^2(3c^2 - 5\beta c^2 + \beta^2 c^2 + \beta^3 c^2 + 12ck - 8\beta ck \\
- 4\beta^2 ck + 12k^2 + 4\beta k^2 \right],
\]

(3.8)

\[
W^*(l_1^*) \equiv W(l_1^*, 1 - l_1^*) = 2b - \frac{c(1 + \beta)(1 - 4l_1^* + 8l_1^2)}{4}.
\]

(3.9)

The presence of switching costs has no direct impact on social welfare. To see this,
observe that in equilibrium no switching costs are incurred. It is because incumbent audit firms always set their second-period audit fees to prevent entry so that clients will not change audit firms in equilibrium. The presence of switching costs only influences the transfer of economic interests from the clients to their incumbent audit firms. However, there is an indirect effect of switching costs on the social welfare through their influence on equilibrium specializations of audit firms. The next section will discuss this further.

3.4 Implications of Changes in Auditing Costs, Learning Rate, and Switching Costs

This section analyzes the implications of changes in auditing costs, learning rate, and switching costs on the audit firms' specializations and profits, the aggregate surplus to clients, and social welfare. Before going on, lemmas 3.1 and 3.2 below provide the indirect effects (i.e., the effects through the changes of the equilibrium audit firms' specializations) of the auditing costs, the learning rate and the switching costs, respectively, on the total profit to both audit firms, the aggregate surplus to clients and the social welfare. The corresponding direct effects will be given in lemmas 3.3-3.5.

**Lemma 3.1.** Given (A1), (i) \( \partial l_1^* / \partial \beta < 0 \). Furthermore, (ii) \( \partial l_1^* / \partial c < (>) 0 \) and (iii) \( \partial l_1^* / \partial k > (<) 0 \) if,

\[
c^2 - 3\beta^2 c^2 + 4ck + 4\beta ck + 2k^2 > (<) 0.
\]  

(3.10)

Part (i) of lemma 3.1 shows that a decrease in \( \beta \) fortifies the market-share effect so that audit firms choose specializations that are closer to the middle of the market. The reason for this is that the monopolistic rents increase with higher learning (i.e., lower \( \beta \)) if audit firms can enlarge their market shares in the initial period. This can be done by choosing a specialization that is closer to the middle of the market. As such, it makes an
audit firm become more cost efficient to serve most of the clients in the market, and hence, increases its market share. Part (ii) and (iii) of lemma 3.1 show that, under some conditions (say, inequality (3.10) is positive), while an increase in c induces audit firms to be more differentiated, an increase in k induces them to choose specializations that are closer to the middle of the market. These imply the audit market may becomes more or less competitive as the cost parameter or switching costs increase.

Lemma 3.2. Given (A1) and holding c, β and k constant. If \( l_1^* \) and \( l_2^* = 1 - l_1^* \) shift exogenously, then we have (i) \( \frac{\partial \Pi^*}{\partial l_1^*} < 0 \) and (ii) \( \frac{\partial S^*}{\partial l_1^*} > 0 \). Furthermore, (iii) \( \frac{\partial W^*}{\partial l_1^*} < (>) 0 \) if,

\[
2c^2 - 2\beta c^2 + 5ck - 6\beta ck - 3\beta^2 ck + 6k^2 + 2\beta k^2 > (<) 0.
\]

Lemma 3.2 states the effects of an exogenous shift of \( l_1^* \) and \( l_2^* = 1 - l_1^* \) on the total profit to both audit firms, the aggregate surplus to clients and the social welfare. Here, we hold c, β and k constant, and do not ask why \( l_1^* \) and \( l_2^* = 1 - l_1^* \) shift. Part (i) and (ii) of lemma 3.2 are very intuitive. They say that if the equilibrium audit firms’ specializations are closer to each other, intense price competition drives down the total profit to both audit firms but raises the aggregate surplus to clients. However, an increase in price competition may hurt the society as a whole as stated in part (iii) of lemma 3.2. In such a case, i.e., if inequality (3.11) is positive, social welfare decreases as the equilibrium audit firms’ specializations are closer to each other and, therefore deviate farther from the welfare-maximization specializations. As a consequence, the total auditing cost in the industry increases. In this sense, excessive competition is detrimental to the society as a whole.

Depending on the sign of inequalities (3.10) and (3.11) there are two possible sets of comparative statics to be considered: (i) when both inequalities (3.10) and (3.11) are positive; and when (ii) both inequalities (3.10) and (3.11) are negative.\(^{25}\) Since the derivation of

\(^{25}\)It is easily verified that the other two cases (i.e., (iii) when inequality (3.10) is positive but inequality (3.11) is negative; and (iv) when inequality (3.10) is negative but inequality (3.11) is positive) do not exist given the set of restrictions imposed on the parameters in the model.
comparative statics in both cases is similar, only the case in which both inequalities (3.10) and (3.11) are positive is examined. The other case can be similarly derived and is left to the interested reader.

It is easy to ascertain that solving the positive inequalities (3.10) and (3.11) simultaneously implies $0 < \beta < 0.885618$ and $\min \{0.171573c, \beta c/2\} < k < \beta c/2$. For expositional convenience, the audit market in which the above conditions are fulfilled is referred to as a high learning (i.e., $\beta$ is small), high switching cost audit market.

**Assumption (A2).** Let $c > 0$, $b > c(2 + \beta)/2$, $\min \{0.171573c, \beta c/2\} < k < \beta c/2$ and $0 < \beta < 0.885618$. 

The following two propositions summarize the results of the above analyses.

**Proposition 3.3.** Given (A2), competition in the audit market is excessive.

In a high learning, high switching cost audit market, market share is very valuable in the second period. This provides an incentive for each audit firm to compete more aggressively to capture that market share. As a result, the SPNE choice of audit firm specializations are closer together than the welfare-maximization ones. This implies that competition is excessive in a high learning, high switching cost audit market.

**Proposition 3.4.** Given (A2), in the SPNE, the audit market is more competitive (i.e., audit firm specializations are closer together) as (i) the auditing cost parameter, $c$, decreases; (ii) the learning parameter, $\beta$, decreases; or (iii) the switching cost parameter, $k$, increases.

Recall that proposition 3.3 states that in a high learning, high switching cost audit market, audit firms compete more aggressively and choose to specialize inside the welfare-
maximization specializations \((l_1^W, l_2^W)\). This implies that the total auditing costs given the SPNE audit firm specializations are higher as compared to that of the welfare-maximization ones. Hence, an increase in \(c\) fortifies the cost effect and induces audit firms to choose specializations that are more distinct from each other but closer to the welfare-maximization specializations as stated in part (i) of proposition 3.4. The intuition underlying part (ii) of proposition 3.4 is that a decrease in \(\beta\) fortifies the market-share effect since the monopolistic rents increase with higher learning if audit firms can enlarge their market shares in the initial period. Hence, a decrease in \(\beta\) enhances audit firms' aggressive behaviour towards market share. The intuition behind part (iii) of proposition 3.4 is similar to that of part (ii). In a high learning, high switching cost audit market, an increase in \(k\) increases the incentive of the audit firms to enlarge their market shares. As a consequence, audit firms choose specializations that are closer to the middle of the market and compete more aggressively.

The following lemmas provide the direct effects (i.e., when specializations of audit firms are fixed) of the auditing costs, the learning rate and the switching costs, respectively, on the total profit to both audit firms, the aggregate surplus to clients and the social welfare. The results of these lemmas will be useful in the proofs of the remaining propositions.

**Lemma 3.3.** Given (A2) and holding \(l_1^*\) and \(l_2^* = 1 - l_1^*\) fixed, then (i) \(\partial \Pi^*/\partial c > 0\); (ii) \(\partial S^*/\partial c < 0\); and (iii) \(\partial W^*/\partial c < 0\).

**Lemma 3.4.** Given (A2) and holding \(l_1^*\) and \(l_2^* = 1 - l_1^*\) fixed, then (i) \(\partial S^*/\partial \beta < 0\) and (ii) \(\partial W^*/\partial \beta < 0\). However, (iii) \(\partial \Pi^*/\partial \beta\) is indeterminate.

**Lemma 3.5.** Given (A2) and holding \(l_1^*\) and \(l_2^* = 1 - l_1^*\) fixed, then (i) \(\partial \Pi^*/\partial k > 0\); (ii) \(\partial S^*/\partial k < 0\); and (iii) \(\partial W^*/\partial k = 0\).

Most of the results of lemmas 3.3-3.5 are very intuitive and will be discussed in the context of the remaining propositions. Only part (iii) of lemma 3.4 seems surprising and
needs immediate discussion. To see why the direct effect of learning on the total profit to both audit firms is ambiguous, recall that the total two-period profit that audit firm 1 earns from supplying services to a client \( z \in [0, 1/2] \), given the equilibrium audit fee schedules and audit firms’ specializations, is

\[
\Pi_1^*(l_1^*) + \Pi_2^*(l_2^*) = \begin{cases} 
  c(2(1 - l_1^* - z) - (1 + \beta)|l_1^* - z|) + k(1 - 2l_1^*) & \text{if } 0 \leq z < z_1, \\
  (1 + \beta)c(1 - 2z) & \text{if } z_1 \leq z \leq 1/2,
\end{cases}
\]

since \( l_2^* = 1 - l_1^* \) in equilibrium. Observe that while a decrease in \( \beta \) increases the two-period profit that audit firm 1 earns from client \( z \in [0, z_1) \), it decreases those from client \( z \in [z_1, 1/2] \). Moreover, the number of clients in the market segment \([z_1, 1/2]\) increases as \( \beta \) decreases.\(^{26}\) Therefore, the direct effect of \( \beta \) on the total profit to audit firm 1 depends on the equilibrium specializations of the two audit firms (which, together with \( \beta \), determine \( z_1 \)). As audit firms are \textit{ex ante} identical, a similar argument applies to audit firm 2. Hence, in a high learning, high switching cost audit market, the direct effect of \( \beta \) on the total profit to both audit firms is ambiguous.\(^{27}\)

Now, the implications of changes in auditing costs, learning rate and switching costs on the audit firms’ specializations and profits, the aggregate surplus to clients and the social welfare are readily presented in the following propositions.

**Proposition 3.5.** Given \((A2)\), in the SPNE, an increase in the auditing cost parameter, \( c \), increases the total profit to both audit firms, decreases the aggregate surplus to clients, but has an ambiguous effect on the social welfare.

Proposition 3.5 provides some comparative statics properties of an increase in the auditing costs. Increasing \( c \) directly increases audit firms’ profit mark-ups when specializations of audit firms are fixed and, hence, it is not surprising that the total profit to both audit firms increases as \( c \) increases. This increase is further accentuated by the fact that increasing \( c \)

\(^{26}\)Recall that \( z_1 \equiv l_1^* + \frac{\beta - k}{(1 + \beta)c}(1 - 2l_1^*) \). Therefore, decreasing \( \beta \) also decreases \( z_1 \).

\(^{27}\)However, given specific values of parameters, the direct effect of \( \beta \) on the total profit to both audit firms is easily determined.
results in a decrease in competition between audit firms as stated in proposition 3.4. As a result, increasing $c$ relaxes competition and allows audit firms to build up higher profits. It is also clear that increasing $c$ directly decreases the net value of an audit to clients when specializations of audit firms are fixed and, hence, the aggregate surplus to clients decreases as $c$ increases. This decrease is further aggravated by the fact that audit fees are also higher if audit firms choose more distinct specializations as a response of an increase in $c$. As a consequence, clients are worse off when there is an increase in auditing costs. However, the effect of an increase in $c$ on the social welfare is ambiguous. On the one hand, increasing $c$ directly decreases the total surplus available to be shared by the clients and the audit firms when specializations of audit firms are fixed. However, the full effect of an increase in $c$ on the social welfare is confounded by the fact that an increase in $c$ improves also social welfare since it drives the two audit firms’ specializations more distinct from each other but closer to the welfare-maximization specializations. In fact, social welfare may be higher if audit firms are more efficiently specialized as $c$ increases.

**Proposition 3.6.** Given $(A^2)$, an increase in the learning parameter, $\beta$, decreases the aggregate surplus to clients, but has an ambiguous effect on the total profit to both audit firms and the social welfare.

To understand the intuition behind proposition 3.6, notice that while an increase in learning allows a given audit firm to reduce its total auditing costs to its clients in the second period, which in turn induces it to enlarge its market share in the first period, it simultaneously induces its competitor to do the same thing. Hence, in equilibrium, audit firms compete more fiercely and choose specializations that are closer to the middle of the market. It is then not surprising that clients are better off as they pay lower audit fees as $\beta$ decreases. Whether audit firms are better off after an increase in the learning depends on whether the change has improved or reduced their profitability under the more competitive
environment. Similarly, even though decreasing $\beta$ directly increases social welfare, when specializations of audit firms are fixed, it also drives audit firms' specializations farther away from the welfare-maximization specializations. Hence, the full effect of a decrease in $\beta$ on social welfare is ambiguous.

**Proposition 3.7.** Given (A2), an increase in the switching cost parameter, $k$, decreases social welfare, but has an ambiguous effect on the total profit to both audit firms and the aggregate surplus to clients.

As stated before, an increase in $k$ has no direct impact on social welfare since clients do not switch audit firms in equilibrium. It only affects the amount being transferred from the clients to their incumbent audit firms. However, an increase in $k$ induces the two audit firms to choose specializations that are closer to the middle of the market in equilibrium, which in turn increases the total auditing costs in the audit market. As a result, social welfare decreases. However, increasing $k$ has an ambiguous effect on the total profit to both audit firms and aggregate surplus to clients. On the one hand, it is not surprising that increasing $k$ directly increases the total profit to both audit firms and decreases the aggregate surplus to clients, when specializations of audit firms are fixed. This simply reflects the fact that more rents are being extracted from the clients by their incumbent audit firms. However, this effect on the total profit to both audit firms or the aggregate surplus to clients is confounded by the fact that increasing $k$ also results in an increase of competition between the two audit firms. Thus, audit firms may be worse off if the increased competition drives their audit fees so low that their ultimate profits will be reduced even if $k$ is increased. By the same token, clients may be better off as a result of an increase in $k$ if they would have to pay lower audit fees owing to a more competitive audit market. This result is consistent with that of Gigler and Penno (1995), even though the economic mechanism is quite different.
3.5 Concluding Remarks

This chapter develops a simple two-period spatial duopoly model to analyze the effects of audit firms’ learning and clients’ costs of switching audit firms on auditing competition. In the model, audit firms make strategic specialization and pricing decisions. Through specialization, an audit firm achieves a comparative cost advantage over its rivals for all clients whose characteristics are closer to its area of specialization. This comparative cost advantage is further fortified by the presence of learning and switching costs. As a result, each audit firm obtains some market power and is able to price discriminate across clients by offering ‘specialization-and-relationship-specific’ audit fee schedules. The analysis demonstrates how equilibrium audit fee schedules and audit firm’s specialization depend on the auditing cost, the learning rate, and the switching costs. In this respect, the analysis may shed light on the potential conflict between the regulations that affect switching costs (e.g., Securities Release #34-9344, ASR-165, ASR-194 and ASR-247) the audit firms in the audit industry would like to adopt and those the regulators and/or the clients might want to impose. In particular, the analysis demonstrates that the economic forces considered in the model are such that lower switching costs can result in efficiency gains. Therefore, if the objective of the regulators (particularly the SEC) is to maximize social welfare (or equivalently, minimize total auditing costs), then they might want to consider regulations that induce lower switching costs. These regulations may raise audit firms’ profits at the expense of clients’ aggregate surplus, but they improve overall efficiency.
Figure 1

The second-period equilibrium audit fee schedule
The first-period equilibrium audit fee schedule

Figure 2
The first-period equilibrium audit fee schedule

Audit firm 1’s profit

Audit firm 1’s loss
Chapter 4

‘Low-balling’ and Efficiency

The practice of ‘low-balling’ has been cited by both the Securities and Exchange Commission (SEC) and the Commission on Auditors’ Responsibilities (Cohen Commission) as a factor which impairs auditor independence. Specifically, low-balling could impair auditor independence since it could provide clients with a credible threat of terminating incumbent audit firms should they refuse accounting concessions. However, based on a multi-period contestable market model of auditing, DeAngelo (1981a) argues that low-balling does not itself impair auditor independence. Instead, she claims that it is the existence of ‘relationship-specific economic interests’ (or ‘quasi-rents’ as defined by DeAngelo (1981a)) between clients and their incumbent audit firms and the competition among audit firms that may lower the ‘optimal’ amount of auditor independence and lead to low-balling. Thus, she concludes that there is no causal relationship between low-balling and impaired auditor independence. This

1A typical definition for ‘low-balling’ is provided by DeAngelo (1981a), in which DeAngelo defines low-balling as setting audit fees below total current costs on an initial audit engagement. Auditor independence, on the other hand, can be measured in different ways. For example, DeAngelo (1981a) measures the level of auditor independence as the conditional probability that an audit firm will truthfully report a misstatement. On the other hand, Magee and Tseng (1990) take auditor independence to mean an auditor’s decisions are consistent with his or her beliefs about a reporting issue.

2For example, the Cohen Commission Report (1978) contends that accepting an audit engagement with the expectation of offsetting early losses with future fees gives the auditor an interest in the financial success of the client and might influence the auditor’s independence in carrying out the examination. The SEC is so wary of this pricing practice that it requires disclosure of any audit fee that is significantly less than what would cover expected direct costs.

3Since initial fee reductions are sunk in future periods, they have no effect on auditor independence.
point is further explored by Magee and Tseng (1990) who provide five necessary conditions under which a relationship-specific economic interest that leads to low-balling may also lead to a compromise of auditor independence.\footnote{These conditions are: (1) auditors must disagree among themselves on a client's reporting issue; (2) at the time of initial engagement, auditors do not know their own positions on the reporting issue; (3) when the reporting issue arises, the client must not know the incumbent auditor's position on the reporting issue; (4) the reporting issue must affect the client for more than one reporting period; and (5) the client must benefit from the preferred reporting strategy even after an auditor switch (Magee and Tseng (1990), p. 317).} However, they argue that those conditions are usually not fulfilled. While the issue of auditor independence is admittedly interesting and important, this chapter focuses on the economic relation between low-balling and efficiency in the audit market.\footnote{Among other things, Magee and Tseng (1990) point out that when audit firms possess all the bargaining power and there is no disagreement among audit firms on the proper interpretation of generally accepted accounting principles (GAAP), clients have nothing to gain by threatening termination of incumbent audit firms and there is no impairment of auditor independence. Given that the assumptions adopted in our model are consistent with those of Magee and Tseng, one might argue that the same conclusion could be reached here.} That is, the purpose of this chapter is to examine the impact of banning audit firms from the practice of low-balling on social welfare, an important issue that has not been fully considered by the academic researchers or the regulators concerned with low-balling by audit firms.

The conclusions of this chapter depends on a comparison of the equilibrium outcomes derived in chapter 3 with those derived in an otherwise-equivalent economy where low-balling is not allowed, i.e., everything is the same as in the basic model in chapter 3 except that audit firms are required to price at or above their marginal auditing costs. It is demonstrated that although a policy of banning low-balling always reduces competition, it improves social efficiency in some cases. The key factor that drives this result is the adverse effect of competition on the total auditing costs in the audit market. In a world without regulations on audit pricing policy, audit firms cannot credibly commit to refuse to price below its marginal auditing cost in competing for initial audit engagements. As a result, audit firms seeking market dominance would like to strategically utilize their service specialization to advance their competitive position in the audit market. As mentioned in the earlier chapters, through specialization with respect to client characteristics, an audit firm can achieve a cost advan-
tage over its rivals for all clients whose characteristics are closer to its area of specialization. More specifically, given the specialization of its rival, an audit firm has an incentive to choose a specialization that is closer to the characteristics of the average client. This action will increase the audit firm's profit by enhancing the audit firm's competitive power. However, while it is individually rational for each audit firm to choose a specialization that is closer to the characteristics of the average client, all audit firms in the audit industry taken together are made worse off by the increased competition. In fact, competition is excessive even from a social welfare perspective if the equilibrium specializations of audit firms are so 'close' that the total auditing costs in the audit market become higher. Therefore, social welfare increases as the policy of banning low-balling allows audit firms to partially collude their audit pricing policies and induces them to specialize in a more efficient way.

The rest of the chapter is organized as follows. Section 4.1 compares our predictions on low-balling with those of the existing literature. Section 4.2 analyzes the impact of banning audit firms from the practice of low-balling. Section 4.3 concludes the chapter.

4.1 Low-balling

The existence of low-balling in the market for auditing services is suggested by the extant theoretical literature on audit pricing which is reviewed in chapter 1. This section compares the primary similarities and differences between the predictions on low-balling of our model and those of the existing literature. However, since not all the predictions on low-balling provided from the existing literature are readily comparable to ours, in the sequel, we only confine our comparison with the predictions provided from DeAngelo (1981a), Kanodia and Mukherji (1994) and Magee and Tseng (1990).

From the analysis in the chapter 3, it is easy to see that the unique first- and second-period SPNE market audit fee schedules offered by the supplying audit firm to each client
\( z \in [0, 1] \) are given by\(^6\)

\[
\begin{align*}
    f_1^*(l_1^*, 1 - l_1^*) &= \begin{cases} 
        c(1 - l_1^* - z) & \text{if } 0 \leq z < z^1, \\
        \beta c(1 - l_1^* - z) - k(1 - 2l_1^*) & \text{if } z^1 \leq z < 1/2, \\
        \beta c(z - l_1^*) - k(1 - 2l_1^*) & \text{if } 1/2 \leq z < z^2, \\
        c(z - l_1^*) & \text{if } z^2 \leq z < 1,
    \end{cases} \\
    f_2^*(l_1^*, 1 - l_1^*) &= \begin{cases} 
        c(1 - l_1^* - z) + k(1 - 2l_1^*) & \text{if } 0 \leq z \leq 1/2, \\
        c(z - l_1^*) + k(1 - 2l_1^*) & \text{if } 1/2 < z \leq 1,
    \end{cases}
\end{align*}
\]

where \( z^1 \equiv l_1^* + \frac{\beta c - k}{(1 + \beta)c}(1 - 2l_1^*) \) and \( z^2 \equiv 1 - l_1^* - \frac{\beta c - k}{(1 + \beta)c}(1 - 2l_1^*) \).

Following DeAngelo (1981a), the low-ball magnitude is defined as the difference between the first-period total auditing cost and the audit fee. Thus the low-ball magnitude in our model is given by

\[
\text{low-ball} \equiv \max \left\{ 0, \min \left\{ |c(1 - l_1^* - z)|, |c(1 - l_1^* - z)| - f_1^* \right\} \right\} \text{ for all } z \in [0, 1]
\]

\[
\begin{align*}
    = \begin{cases} 
        0 & \text{if } 0 \leq z < z^1, \\
        c[z - l_1^* - \beta(1 - l_1^* - z)] + k(1 - 2l_1^*) & \text{if } z^1 \leq z < 1/2, \\
        c[1 - l_1^* - z - \beta(z - l_1^*)] + k(1 - 2l_1^*) & \text{if } 1/2 \leq z < z^2, \\
        0 & \text{if } z^2 \leq z \leq 1.
    \end{cases}
\end{align*}
\]

In DeAngelo’s (1981a) model, the predicted low-ball magnitude is the sum of the incumbent audit firm’s learning and the client’s switching costs. On the other hand, the low-ball magnitude found in Kanodia and Mukherji (1994) is strictly less than that amount. Our result predicts that low-balling only occurs in the market segment \([z^1, z^2]\) where competition is fierce. Moreover, observe that for any client \( z \in [z^1, 1/2] \) (a similar argument applies to any client \( z \in (1/2, z^2) \)), the low-ball magnitude is given by

\[
c[z - l_1^* - \beta(1 - l_1^* - z)] + k(1 - 2l_1^*) \leq c(1 - \beta)(z - l_1^*) + k(1 - 2l_1^*),
\]

where the equality holds only when \( z = 1/2 \). Thus, the predicted low-ball magnitude in our model is less than the sum of the incumbent audit firm’s learning and switching costs almost everywhere. This result is consistent with that of Kanodia and Mukherji, even though our analysis is based on a different economic mechanism.

\(^6\)Recall that \( l_2^* = 1 - l_1^* \) in the unique symmetric SPNE choice of specializations.
The extant literature also uses 'price-cut' to measure the magnitude of the initial fee reductions. Following Magee and Tseng (1990), price-cut is defined as the difference between the second- and first-period audit fees.\(^7\) Thus the price-cut magnitude in our model is given by

\[
\text{price-cut} = \max \{0, f_2^{*} - f_1^{*}\} = \begin{cases} 
  k(1 - 2l_t^*) & \text{if } 0 \leq z < z^1, \\
  c(1 - \beta)(1 - l_t^* - z) + 2k(1 - 2l_t^*) & \text{if } z^1 \leq z \leq l/2, \\
  c(1 - \beta)(z - l_t^*) + 2k(1 - 2l_t^*) & \text{if } l/2 < z \leq z^2, \\
  k(1 - 2l_t^*) & \text{if } z^2 < z \leq 1.
\end{cases}
\]

Magee and Tseng show that the first-period price-cut should at most equal to the client's switching costs (i.e., price-cut = \(k(1 - 2l_t^*)\) in our model). They argue that "the first-period price-cut observed by Simon and Francis (1988) should be correlated with the client's costs of switching to new auditor, not with the auditor's learning cost (p. 320)." Thus, if client switching costs are less than the first-period price-cut, additional explanations for price-cutting are required. Our result shows that Magee and Tseng' suggestion is valid only in the market segments where low-balling does not occur, i.e., if \(z \in [0, z^1) \cup (z^2, 1]\). When low-balling occurs, i.e., if \(z \in [z^1, z^2]\), the first-period price-cut is always higher than the client's switching costs. The reason for this observation is as follows. For any client \(z\) in the market segments \([0, z^1)\) and \((z^2, 1]\), the second lowest-cost audit firm realizes that even if it were the incumbent audit firm for client \(z\), it cannot offer an audit fee to him that is as attractive as that of the lowest-cost audit firm in the second period. In other words, the second lowest-cost audit firm knows for sure that it will lose client \(z\) to the lowest-cost audit firm in the second period. Thus, it has no incentive to offer an audit fee that is lower than its current auditing cost to client \(z\) in the first period. Anticipating this behaviour of the second lowest-cost audit firm, the lowest-cost audit firm will not charge less than the auditing cost of the second lowest-cost audit firm to client \(z\) in the first period. As such, the price-cut

\(^7\)The term 'pricing-cutting' is empirically motivated. Empirical researchers define it as the different between the first-year audit fee and either prior auditor's fee or an estimated fee based on a cross-sectional model. See Francis and Simon (1987) for details.
to client $z$ in the first period just reflects the maximum amount of economic rent that the lowest-cost audit firm can extract from him in the second period owing to the existence of the client's switching costs, i.e., $\text{price-cut} = k(1 - 2t^*_1)$. On the other hand, for any client $z$ in the interval $[z^1, z^2]$, the existence of learning and switching costs in the second-period will increase the second-period profit of the second lowest-cost audit firm if it can attract the client to patronize it in the first-period. Thus the second lowest-cost audit firm, anticipating an economic rent earned as an incumbent in the second period, is willing to cut its fee in the first period to the extent that its second-period economic rent is exactly turned over to the client. This in turn drives the first-period audit fee of the lowest-cost audit firm downwards in order to meet the competition and results in a price-cut that is greater than the client's switching costs.

Notice that our result has a clear empirical implication for price-cutting. It suggests that empirical researchers should always observe price-cutting. However, empirical evidence on price-cutting is mixed. On the one hand, Baber, Brooks and Ricks (1987), Ettredge and Greenberg (1990), Francis and Simon (1987) and Simon and Francis (1988), find evidence of price-cutting in first-year audits. On the other hand, Francis (1984), Palmrose (1986) and Simunic (1980) find no significant evidence of price-cutting. We do not have an explanation for this mixed empirical evidence, and neither does the existing literature. Rather, we admit that our understanding of the dynamics of audit pricing is far from complete. We would expect that more important questions can be addressed by studying a $T$-period model with learning and switching costs, where $T > 2$. For an example, one may want to know whether price-cutting will persist beyond the initial period. Empirical evidence on this issue is also inconclusive. Simon and Francis (1988) find that there is price-cutting on initial audit engagements and the lower audit fee persists into the second and third years following an auditor change. On the other hand, Baber, Brooks, and Ricks (1987) provide evidence of price-cutting on initial audit engagements in the public sector. Their results do not indicate that price-cutting persists beyond the initial engagement year. We propose to study this
issue on our future work.

4.2 The Welfare Implications of Low-Balling

We now shift our focus to the welfare implications of low-balling. To this end, we first derive the equilibrium outcomes in an economy where low-balling is not allowed, i.e., audit firms are required to price at or above their marginal auditing costs. Then, the equilibrium outcomes derived will be compared with those derived in chapter 3, where no restrictions are imposed on audit pricing policy.

4.2.1 Equilibrium Outcomes without Low-Balling

Suppose that audit firms are required to price at or above their marginal auditing costs. Through competition and specification of the lowest bound on the audit price, it is straightforward to derive the equilibrium audit fee schedules as follows:

\[ f_{11}(l_1, l_2) = f_{21}(l_1, l_2) = \max \{ l_1 - z, l_2 - z \}, \]
\[ f_{12}(l_1, l_2) = f_{22}(l_1, l_2) = \max \{ l_1 - z, l_2 - z \} + k(l_2 - l_1). \]

That is, in equilibrium an audit firm will charge a client whose characteristics are in the proximity of its own specialization at the marginal auditing cost of its competitor in the first period, and at the marginal auditing cost of its competitor plus the client’s switching costs in the second period.

Given these audit fee schedules and the specialization of its rival, audit firm 1 chooses its specialization to maximize its total two-period profit, i.e., audit firm 1 solves the following profit-maximization problem given \( l_2 \): (P4.1)

\[
\max_{l_1} \int_0^{\frac{l_1 + l_2}{2}} c(l_2 - z - |l_1 - z|) \, dz + \int_0^{\frac{l_1 + l_2}{2}} [c(l_2 - z - \beta|l_1 - z|) + k(l_2 - l_1)] \, dz,
\]

since \( q_{1t}^* = 0 \) for all \( z \in ((l_1 + l_2)/2, 1] \) and \( t = 1, 2 \). By the same token, one can define...
audit firm 2’s profit-maximization problem.

The next proposition shows the unique SPNE choice of audit firm specializations when audit firms are not allowed to low-ball.

**Proposition 4.1.** *Given (A1), the unique symmetric SPNE choice of audit firm specializations under the duopoly auditing structure without low-balling, \((l_1^N, l_2^N)\), is given by

\[
l_1^N = 1 - l_2^N = \frac{c(3 + \beta)}{2(5c + 3\beta c + 2k)}.
\]

**Corollary 4.1.** *Suppose (A1) holds and audit firms are not allowed to low-ball. Then (i) as \(\beta\) approaches one in the limit, competition in the audit market is insufficient, i.e., \(\lim_{\beta \to 1} l_1^* < l_1^N\); (ii) as \(k\) approaches zero in the limit, competition in the audit market is excessive, i.e., \(\lim_{k \to 0} l_1^* > l_1^W\); (iii) as \(\beta\) approaches one and \(k\) approaches zero in the limit, the SPNE audit firm specializations are the same as the welfare-maximization specializations.*

The properties of the equilibrium audit firms’ specializations under the duopoly auditing structure without low-balling are the same as those with low-balling. The intuition described in proposition 3.2 and corollary 3.1 applies here.

Accordingly, given the SPNE choice of audit firms specializations under the duopoly structure without low-balling, \((l_1^N, l_2^N)\), where \(l_2^N = 1 - l_1^N\), the total two-period profit to both audit firms is given by

\[
\Pi^N(l_1^N) = \Pi(l_1^N, 1 - l_1^N) = \frac{5c - \beta c + 4k - 4l_1^N(c - \beta c + 2k) - 8cl_1^{N2}(1 + \beta)}{4},
\]

and the corresponding aggregate surplus to clients and social welfare, respectively, are given by

\[
S^N(l_1^N) = S(l_1^N, 1 - l_1^N) = 2b - \frac{3c + 2k - 4l_1^N(c + k)}{2},
\]

\[
W^N(l_1^N) = W(l_1^N, 1 - l_1^N) = 2b - \frac{c(1 + \beta)(1 - 4l_1^N + 8l_1^{N2})}{4}.
\]
4.2.2 Welfare Comparison

Now, the equilibrium outcomes derived in the previous section can be compared with the ones derived in the duopoly auditing structure without restrictions on audit pricing policy. The major results of this chapter are presented as follows:

Proposition 4.2. Suppose (A1) holds and audit firms are not allowed to low-ball. Then, compared to the outcome under the duopoly auditing structure with low-balling, audit firms’ specializations are more distinct from each other (i.e., \( l_1^N < l_1^* \)), total profit to both audit firms is higher, and aggregate surplus to clients is lower. The effect on social welfare is ambiguous.

It is easy to ascertain that the audit fees paid by clients are higher in both the first and second periods and the total profit to both audit firms is higher in an audit market without low-balling. The reason is that, without regulations on audit pricing policy, audit firms cannot credibly commit, in the first-period pricing stage of the game, to refuse to price below marginal auditing cost. Regulations on audit pricing policy provide a substitute for this precommitment. As a consequence, the banning of low-balling effectively relaxes price competition and allows the two audit firms to achieve a partial collusion in the audit market. However, it is of interest that the society as a whole may be either better off or worse off when low-balling is not allowed. To see this, subtracting (3.9) by (4.3), it is easy to show that

\[
W^* - W^N = c(1 + \beta)(l_1^* - l_1^N)[1 - 2(l_1^* + l_1^N)].
\]

Since \( l_1^* > l_1^N \) by the first part of proposition 4.2, then

\[
W^N > W^* \quad \text{if, and only if,} \quad l_1^* + l_1^N > \frac{1}{2}.
\]

This implies that there are two possible cases in which social welfare is higher without low-balling: (i) \( l_1^* > l_1^N > l_1^W \); or (ii) \( l_1^* > l_1^W > l_1^N \), with \( l_1^* - l_1^W > l_1^W - l_1^N \). In both cases, \( l_1^* > l_1^W \),
i.e., competition in the audit market is excessive without any regulations on audit pricing policy. Then, the above analysis and tedious calculation lead to the following proposition.

Proposition 4.3.. Suppose audit firms are not allowed to low-ball. Then, compared to the outcome under the duopoly auditing structure with low-balling, social welfare is higher if, and only if, (i) $\beta \leq 0.5$, or (ii) $0.5 < \beta < 0.778225$ and $c(1 - \beta)/2 < k < \beta c/2$.

Proposition 4.3 states that under some parameter values, social welfare increases as a result of the policy of banning low-balling. This result is driven by the fact that the policy of banning low-balling effectively allows audit firms to partially collude their pricing policies. In particular, audit firms are less concerned about their competitive advantage in the market segment $[l_1, l_2]$ where competition is the most fierce. Consequently, audit firms are induced to choose specializations that are more distinct from each other but are closer to the welfare-maximization specializations. Thus, increased efficiency justifies decreased competition. Nevertheless, in any case, total profit to both audit firms increases at the expense of the aggregate surplus to clients (actually, all clients are worse off as audit fees increase).

4.3. Concluding Remarks

This chapter compares the predictions on low-balling and price-cutting of our model with those of the existing literature. Our work is distinguished from the existing literature on the basis of its focus on an imperfect spatial audit market. While our analysis agrees with the existing literature that the practice of low-balling is a natural consequence of the competition among audit firms, we are able to identify that low-balling occurs only in a certain market segment where audit firms compete quite fiercely. On the other hand, if price-cut is defined as the difference between the second- and first-period audit fees, our analysis suggests that it should be observed as long as there are economic rents accruing to the incumbent audit firms. This chapter also examines the welfare consequences of banning the practice of low-balling.
Our analysis suggests that while a policy of banning low-balling always reduces competition, it improves social efficiency in some cases. Thus, the analysis in this chapter provides a legitimate reason for a regulator to consider the banning the practice of low-balling.
Chapter 5

Conclusion

This dissertation develops variants of the well-known Hotelling’s location model to examine the nature of competition in the audit market where audit firms make strategic specialization and pricing decisions. In this chapter, the major conclusions of this dissertation are enumerated in their order of presentation. Suggestions for future research are also provided. As in most formal models, our models are also a stylized representation of real world phenomena, employing a simple and restrictive framework to obtain tractability. The appropriateness of our assumptions hinges on the empirical validity of our predictions. Perhaps, future research will shed light on this issue.

Chapter 2 presents a multi-period oligopoly spatial model of auditing competition that captures all salient economic features of an audit market that involves a large number of audit clients with different characteristics relevant to audit production and relatively few audit suppliers who differ in their area of specialization with respect to client characteristics. In the model, audit firms strategically choose their area of service specialization and compete in audit fees. It is demonstrated that, through specialization, an audit firm achieves a cost advantage over its rival for all clients whose characteristics are closer to its area of specialization. Thus, each audit firm obtains some market power through specialization and is able to price discriminate across clients by offering ‘specialization-specific’ audit fee schedules.
As a result, given a configuration of audit firm specializations, the unique subgame perfect Nash equilibrium audit fee schedule is such that each audit firm charges the minimum of the marginal auditing costs of its rivals on services to clients in the vicinity of its specialization. This structure of the equilibrium audit price schedule is robust to arbitrary client distributions, audit cost functions, multidimensional client-characteristics space, and \( n > 2 \) audit firms. Given the equilibrium audit fee schedules, clients purchase audit services from the least-cost supplier. This implies that the cost effectiveness of audit firms determines their ultimate market shares. The unique pricing equilibrium is also shown to induce an allocation of clients’ surplus and audit firms’ profits that lies in the core of the economy. That is, at the final allocation, no group of clients can move to another audit firm for a mutually advantageous auditor-client re-alignment. When making their specialization decisions, audit firms anticipate the pricing and specialization decisions of their rivals. The competitive forces in the market induce audit firms to achieve constrained efficient utilization of specialized resources. We establish the existence of a subgame perfect Nash equilibrium choice of audit firm specializations and find that such a specialization equilibrium is such that each audit firm chooses a specialization that maximizes the expected social welfare, given the unique subgame perfect Nash equilibrium audit fee schedule and its rivals’ specializations. We also demonstrate that audit firms will not choose the same specialization in equilibrium. This is because the audit firms’ profits are driven down to zero by intense price competition if they choose the same specialization. Instead, in order to earn rents as ‘local monopolists’, audit firms search for ‘niche’ markets such as industry specialization. Thus, the model provides a theoretical link between audit firm specializations and the observed market segmentation in which clients with similar characteristics buy from the same audit firm which has cost efficiency advantage in serving them.

To obtain stronger results, chapter 3 focuses on a simple two-period spatial duopoly model of auditing competition in which there is (i) learning by the incumbent audit firms, and (ii) clients incur transactions costs if they switch audit firms. As in chapter 2, audit firms achieve
a comparative cost advantage through specialization. However, in the chapter 3 model, this comparative cost advantage is further fortified by the presence of learning and switching costs which create client-specific economic interests. Thus, audit firms optimally price discriminate among clients by offering them 'specialization-and-relationship-specific' audit fee schedules.

The analysis demonstrates that the practice of low-balling in initial audit engagements is a natural consequence of the competition among audit firms which are seeking to achieve market dominance. However, low-balling occurs only in a strict subset of the market in which audit firms compete quite fiercely. We also examine how equilibrium audit fee schedules, audit firms’ specializations’ and profits, clients’ surplus, and social welfare depend on the auditing costs, the learning rate, and the switching costs. Some of our results are shown to carry interesting policy implications. For example, the analysis enables us to understand why there may be a conflict between the regulations (e.g., Securities Release #34-9344, ASR-165, ASR-194 and ASR-247) the audit firms in the audit industry would like to adopt and those the regulators and/or the clients might want to impose. In this respect, our results suggest that if the objective of the regulators (particularly the SEC) is to maximize social welfare (or equivalently, minimize total auditing costs), then the regulators should impose regulations that induce lower switching costs. This policy may raise audit firms’ profits at the expense of clients’ aggregate surplus, but it improves overall efficiency.

The issue of low-balling has received considerable attention. In the past, the interest in low-balling has stemmed from the hypothesized link to a loss of auditor independence. While the issue of auditor independence is admittedly interesting and important, our focus in this dissertation is on the economic relation between low-balling and audit market efficiency, which has not been fully considered by the academics and regulators concerned with low-balling by audit firms. Applying the model developed in chapter 3, chapter 4 examines the welfare consequences of banning audit firms from the practice of low-balling. It is demonstrated that if low-balling is not allowed, while audit firms are better off and clients are worse off, the effect on social efficiency is ambiguous. Our contribution lies in the fact
that we are able to identify those conditions under which a policy of banning low-balling can improve social efficiency. Thus, the analysis in this chapter provides a legitimate reason for a regulator to consider the banning the practice of low-balling.

By and large, this dissertation shows that inefficiency in the audit market arises from the fact that the market equilibrium audit firm specializations are in general different from the corresponding welfare-maximization ones in an unregulated audit market. This suggests a rationale for regulating the audit market. However, it is natural for the regulatory decision on the audit market (particularly on audit pricing policy) to be made without considering how the anticipation of that decision might have affected audit firm specialization (which is already in place at the time the decision is taken). The analysis in this dissertation suggests that this 'case by case' approach may lead regulators to ignore important externalities associated with their decisions. Instead, it is emphasized that regulations on the audit market should be formulated with an understanding of their likely effect on audit firm specialization.

In conclusion, this dissertation represents an attempt to apply spatial economic analysis to understand the nature of competition in the audit market where audit firms make strategic specialization and pricing decisions. On the contrary, the extant theoretical research on auditing competition considers the audit market as a concentrated and ex-ante perfectly competitive market, and has primarily focused on the pricing behaviour of audit firms but ignored the issue of audit firm specialization. However, the analysis of spatial auditing competition in this dissertation is undoubtedly far from complete. The current models are restrictive in that they do not permit an explicit examination of other important issues such as auditor independence and the roles of different audit quality and asymmetric information regarding auditing costs in audit markets. Without an explicit consideration of the effect of auditor independence on the welfare of the financial statement users (other than the audit clients), it would be reckless to make any strong statements about the effect of a regulatory decision.\footnote{In the dissertation, we simply assume that the regulators' objective function is to minimize total auditing} In this respect, the purpose of this dissertation is less ambitious. It
provides structured models that highlight some of the basic trade-offs between the welfare of audit firms and clients that arise when regulations are changed. A formal exploration that considers auditor independence and the welfare of the end users of audited financial statements is left for future work.

Another obvious deficiency in the models is the assumption that clients perceive audit quality as constant across audit firms notwithstanding suggestions in the literature that clients correlate audit quality with 'brand name' of the audit firm (see DeAngelo (1981b)). Allowing audit quality to vary in the spatial framework would considerably complicate the analysis, but is of course a promising topic for future research.

Finally, the results of our models predict clients never switch audit firms in equilibrium. This is because incumbent audit firms always set their second-period audit fees to prevent entry so that clients will not have any incentive to change audit firms in equilibrium.\(^2\) Our result is consistent with those of DeAngelo (1981) and Magee and Tseng (1990), where deterministic auditing costs and no disagreement among audit firms on the client's reporting issue are assumed. On the other hand, clients do change their audit suppliers when there are auditor-client disagreements, as in Dye (1991) and Teoh (1992); when cost-minimization auditor-client matches change over time, as in Gigler and Penno (1995); or when auditor switches are the clients' rational response to limit the value of incumbency of the audit firms owing to their superior knowledge of the auditing costs at the time of the switch, as in Coate and Loeb (1994) and Kanodia and Mukherji (1994). Other than Gigler and Penno (1995), where auditor switching is an artifact of their assumption that efficient auditor-client matches change over time, the above approaches emphasize that some sort of information asymmetry is necessary for auditor switching to be sustainable as an equilibrium phenomenon. For our future work, we propose to incorporate information asymmetry regarding auditing costs or, equivalently, maximize social welfare (which, by definition, equals to the sum of the total profit to both audit firms and the aggregate surplus to clients) without providing any economic justification for why they would behave this way.

\(^2\)See Saloner (1987) and Milgrom and Roberts (1982) for theoretical discussions of predatory pricing.
into our models. This extension will check the robustness of the results obtained in this dissertation.
References


APPENDIX A

The Potential Benefits of External Auditing: An Information Economic Analysis

The primary purpose of this appendix is to understand how voluntary auditing services can increase audit purchasers' welfare. Prior research has used agency theory and information economics to suggest explanations for the production of audited information. All in all, the existing models of auditing assume that there is a welfare loss caused by information asymmetry between an insider (firm in our model), who has private information, and an outsider (lender in our model), who has not. This information asymmetry in the market gives rise to a demand for auditing as a means of information transfer. Our model goes one step backwards and assumes that the firm does not have superior information over the lender. However, the firm can choose to become privately informed if it wishes. We argue that, given the auditor's liability system, the economic value of an audit lies in the auditor's credibility in providing unbiased information compared to other means of providing the same information. That is, while in the existing models of auditing that include asymmetric

---

1 A non-mandated auditing framework is assumed because we want to show a demand for auditing not driven by exogenous regulation. Rather, we would like to treat the mandated auditing framework as a special case of the non-mandated framework. Moreover, Wallace (1980) points out that claims that auditing is prevalent due to regulation are inconsistent with the existence of audits prior to regulation.

2 For example, see Baiman, Evans and Noel (1987), Datar, Feltham and Hughes, Evans (1980), Melumad and Thoman (1990), and Shibano (1993), to name just a few.
information, the role of auditing is to attest and to verify the firm’s private information, we assume that the auditor has a stronger role of providing additional information that has economic value.\(^3\)

In our model, the owner of a firm (the audit client) has monopoly access to different mutually exclusive investment projects which are classified by risk and return. The firm has no resources and must turn to a competitive debt market for funding. We assume that the current financial condition of the firm not only affects the probability of project success, but also influences the firm’s attitude towards project risk. More specifically, we assume that while the high risk project is optimal for a firm with a good financial condition, it is not optimal for a firm with a bad financial condition. However, the firm does not have private information about its own financial condition. Moreover, we assume that without any information, the firm may shift to the high risk project once a loan contract is signed. Thus, a rational lender will make his loan decision based on his assessment of the firm’s current financial condition and his knowledge about the firm’s opportunistic behaviour. We show that given the uncertainty about the current condition of the firm and the fact that the firm may have an incentive to substitute a riskier project, the lender’s assessment of the firm’s ability to repay may be so low that the firm will underinvest. That is, the project choice as well as the investment amount will be suboptimal. As such, the firm may attempt to discover its current financial condition through private information production. However, this means of private information production may lack credibility to the lender. Thus, the firm may have an incentive to hire an external independent auditor to attest to the accuracy of its financial statements, such that the more appropriate project will be undertaken and more favourable loan terms will be accepted by the lender. In executing her attest duties, the auditor acquires information about the firm’s true financial condition. Based on the information, the auditor either agrees to attest to the firm’s proposed report and issues an clean opinion, or disagrees with it and issues a qualified opinion. The audit is thus envisioned

\(^3\)Titman and Trueman (1986) make the same assumption.
as a means for independently producing additional information upon which loan contract between the firm and the lender is based.⁴

Like the other players in the game, the auditor is also modelled as a rational economic agent who is subject to moral hazard.⁵ The moral hazard stems from the fact that neither the firm nor the lender can costlessly observe the auditor's action after the auditor is appointed. As a result, the auditor has an incentive to shirk if the audit is costly to perform. Since contingent fees for auditors are illegal, the court system is used to discipline the auditor.

We assume that if the auditor is sued, then she is held liable if she fails to exercise due diligence or care as an auditor. However, what is due care is not always obvious, i.e., there is a vague negligence standard. Generally Accepted Auditing Standards (GAAS) provide some useful guidance. Currently in the U.S. legal system, auditors usually defend against charges of having a negligent audit by demonstrating that they have complied with GAAS. However, courts do not always go by GAAS. As stated in Palmrose (1987): "... adherence to GAAP/GAAS does not provide absolute assurance that would dismiss auditors from any liability for material omissions or misstatement (p.91)." To illustrate this argument, we assume that the probability of conviction when an audit is conducted in accordance with GAAS is greater than zero. That is, we maintain the assumption that the courts are fallible. On the other hand, to maintain that adherence to GAAS is a reasonable defense for the auditor, we assume that the probability of conviction when an audit is conducted in accordance with GAAS is smaller than that when it is not. Moreover, we posit that the avoidance of litigation arising from substandard audits is the primary force motivating an auditor to adhere to auditing standards. There is no incentive for the auditor to exceed the prescribed standards as long as they are met. Nevertheless, the actual quality of an audit

⁴Although auditing does not provide information about the firm's project choice, it affects the firm's project choice because additional information about the firm's current financial condition might provide indirect information about the firm's project choice.

⁵As pointed out by Antle (1982), if one seeks to understand the behaviour of the firm and the lender by modelling them as expected utility maximizers, the same treatment has to be made to the auditor since the auditor's incentives are also endogenous.
is not publicly observable when it is conducted, or when the audited report is issued. At best, the actual quality of an audit can only be perceived by the firm and the lender. More formally, we model the perceived audit quality as linked to the auditor's attachable wealth, and assume that there always exist some auditors who are wealthy enough so that they will always comply with the prescribed auditing standards when they are hired. Hence, the court system and the auditing standards affect the interaction between the auditor and the firm. The interaction, in turn, determines the information content of the firm's audited report. The firm’s ability to repay is then determined by both the lender's rational expectations about the information content of the audited report and the firm's investment decision.

Since the debt market is competitive, the lender earns zero expected profits in equilibrium and, therefore, is indifferent about the existence of an audit. The firm, on the other hand, is better off with an audit since it improves capital allocation and facilitates investment decision. We demonstrate that if the marginal benefit to the firm of an increase in the quality of auditing standards is higher than the marginal cost, the gross benefit of an audit increases as the quality of auditing standards increases.

The organization of the rest of this appendix is as follows. Section A.1 describes and analyzes a simple borrower-lender model without auditing. Section A.2 introduces an auditor who is hired to investigate the borrower’s financial condition and report the findings. Section A.3 briefly concludes the appendix.

A.1 The Basic Borrower-Lender Model without Auditing

A.1.1 The Model

Consider a universal risk-neutral economy in which the owner of a firm has monopoly access to $M$ mutually exclusive investment projects, indexed by $x \in X$. The firm has no resources and must solicit at most a single loan from a competitive debt market in order to
fund any project. Without loss of generality, the net riskless interest rate in the economy is set to be zero. There are $N \geq M$ types of observationally identical firms, indexed by $y \in Y$. One may interpret $Y$ as a possible set of the firm's current financial condition which reflects the past performance of the firm and affects the probability of project success. Thus a lender makes a loan decision based on his assessment of the firm's current financial condition. The better the firm's current financial condition, the lower the default risk is. Because of this, the firm may want to submit a financial report regarding its current financial condition to a potential lender. However, even though the distribution of firm types is common knowledge to the firm and the lender, the firm does not have superior information about its own type. Moreover, it is assumed that there is no mechanism to penalize misstatements, such that the firm would always overstate its financial condition in order to get a better loan term. In this setting, a rational lender would then assess the firm's financial condition as if there is no useful information provided from the financial report. Later in this appendix, an auditor who has access to an audit technology that can reveal the firm's type will be introduced. In such a case, the firm must get the certification from an external auditor to show that its financial report is a fair representation.

The sequence of events is as follows. There are three dates, 0, 1 and 2. At date 0, the firm requests and obtains a single loan from a lender in a competitive debt market. That lender responds by accepting or rejecting the loan request. If the offer is rejected, the firm then goes to a new lender with another loan proposal, and so on, until an offer is finally accepted. After an agreement has been reached, the firm is prohibited from seeking additional loans.

At date 1, the firm privately chooses its optimal investment project given the loan amount obtained at date 0. For an investment amount $I$, project $x$ yields at date 2 a random terminal

---

7 It implies that the financial report would merely reflect a priori information in this case.
8 For example, a debt covenant that restricts future debt issues is included in the loan contract.
9 For simplicity, the agency problem between the owners and the managers of the firm is assumed away.
cash flow of $R_x(I) > 0$ if successful and zero if not.\textsuperscript{10} The support of the terminal cash flow is assumed to be independent of the firm’s type. This ensures that firm types cannot be \textit{ex post} distinguished by merely observing the firm’s performance. Moreover, the return function $R_x(I)$ satisfies the following usual regularity conditions: it is strictly increasing and strictly concave with $R_x(0) = 0$, $\lim_{I \to \infty} R_x(I) < \infty$ and $\lim_{I \to 0^+} \partial R_x/\partial I = \infty$ for all $x \in X$.\textsuperscript{11}

It is assumed that the realized project cash flows, while observable, are not contractible.\textsuperscript{12} Because of this, we limit the extent to which the loan contract can be based on the observation of the project’s realized cash flow, and simply assume that a standard loan contract is optimal in this setting.\textsuperscript{13} Specifically, a standard loan contract offered by the firm, $(r, I)$, is a pair that specifies a loan amount, $I \geq 0$, provided by the lender if he accepts the contract, and a nominal repayment amount, $r \geq I$, that the firm must repay to the lender after the project’s cash flow is realized (if it has sufficient funds).

For simplicity, we focus on the case where there are only two types of firms and two mutually exclusive investment projects. The firm is either ‘good’ ($G$) or ‘bad’ ($B$), indexed by $y \in \{G, B\}$. The common prior probability that the firm is a good type is $\phi \in (0, 1)$. The two investment projects are classified by risk and return and are indexed by $x \in \{L, H\}$. We refer to project $L$ as the low risk project and project $H$ as the high risk project. For an

\textsuperscript{10}The assumption that there is a single state in which the lender is paid is made for the sake of simplicity. Relaxing this assumption will not affect the results qualitatively.

\textsuperscript{11}These are the necessary and sufficient conditions for the existence of an interior solution.

\textsuperscript{12}This assumption rules out forcing contracts based on punishing the firm if the final project payoff indicates it chose the ‘wrong’ project. A justification for this assumption is that a single outcome is used here for simplicity. More generally, one could have multiple outcomes with constant support (i.e., investment amount, $I$, and project choice, $x$, only shift the probability of project success). When the project is successful, the first realized outcome is sufficiently large to repay the debt when the loan contract terminates. However, a significant time delay may exist between the loan repayment and the realizations of the remaining outcomes. Clearly, this justification is far from pleasing, but we believe this simplification does not destroy the essence of the results.

\textsuperscript{13}A formal way to establish the optimality of the standard loan contract is to assume that the realized project cash flow can be observed by the lender only by spending some costs in monitoring at date 2. The firm then must pay the lender at date 2 whenever it is solvent, or otherwise the lender monitors the firm and seizes the entire proceeds of the project. Then, given the structure of our model and assuming that the lender can commit to follow his state-contingent monitoring strategy, it is well known that the optimal incentive compatibility contract must be a standard loan contract. See, for example, Diamond (1984) and Gale and Hellwig (1985) for details.
investment amount $I$, the payoff of the high risk project is $R_H(I)$ with probability $p_G$ ($p_B$) and 0 with probability $1 - p_G$ ($1 - p_B$) if this project is managed by a good (bad) type firm. On the other hand, the low risk project yields a payoff of $R_L(I)$ with probability $\tilde{p}_L$ and 0 with probability $1 - \tilde{p}_L$, regardless of the type of firm.\textsuperscript{14} We assume that $R_L(I) = \sigma R_H(I)$ for all $I > 0$, where $0 < p_B < \sigma\tilde{p}_L < p_G < 1$. The restriction on the scaling factor, $\sigma$, ensures that project $H$ ($L$) generates a higher expected return than project $L$ ($H$) does if it is managed by a good (bad) type firm. This assumption captures the notion that high risk is usually associated with a greater number of opportunities whose exploitation would place a good type firm at an advantage relative to a bad type firm. We also assume that all projects will generate a positive return irrespective of the firm’s type, i.e., $\tilde{p}_L R_x(I) > I$ and $p_y R_x(I) > I$ for all $x \in \{L, H\}$, $y \in \{G, B\}$ and $I > 0$. Let $\bar{p}_H \equiv \phi p_G + (1 - \phi)p_B$ be the firm’s expected solvency probability given its project choice is $H$. We assume that $\sigma\tilde{p}_L > \bar{p}_H$, which implies that the low risk project generates a higher expected return than the high risk project does if the firm’s type is unknown.\textsuperscript{15} A sufficient condition for this condition to hold is that $\phi$ is sufficiently small. It is easy to see that this in turn requires

$$\phi < \frac{\sigma\tilde{p}_L - p_B}{p_G - p_B}.$$  \hspace{1cm} (A.1)

The equilibrium concept that we employ is Selten’s (1975) subgame perfect Nash Equilibrium (SPNE) and our attention is restricted to pure strategies only. Formally, a subgame perfect Nash Equilibrium (hereafter called equilibrium) requires that, at each node of the game, the equilibrium strategy of each player maximizes his expected terminal payoff given the strategies of the other players. A strategy for the firm is a combination of the loan amount and the project choice. A strategy for the lender is a loan contract that he is willing to accept from the firm. Hence, an equilibrium is an allocation, $[(r, I), x]$, which is a pair

\textsuperscript{14}We could let firm type influence the success probability for the low risk project as well as the high risk project without changing the qualitative nature of the results, provided that a good type firm has a higher probability of success in the high risk project than in the low risk project.

\textsuperscript{15}This condition gives rise to the classical risk incentive problem in which a bad type firm has an incentive to choose the high risk project even though the low risk project will yield a high expected return if the firm’s project choice is unobservable to the lender. See, for example, Green (1984).
consisting of a loan contract, \((r, I)\), and a project, \(x\), chosen by the firm.

A.1.2 Characterizing the Equilibrium

The equilibrium under the setting without auditing is characterized by solving the following principal-agent problem: (PA.1)

\[
\max_{I > 0, \ r > I, \ x^*} \quad \bar{p}_x [R_x(I) - r]
\]

s.t.
\[
\bar{p}_x r - I \geq 0,
\]
\[
x^* \in \arg \max_{x \in \{L, H\}} \bar{p}_x [R_x(I) - r].
\]

In words, an optimal loan contract is the one in which the expected terminal payoff of the firm, \((A.2)\), is maximized subject to some constraints.\(^{16}\) Constraint \((A.3)\) guarantees that the lender at least breaks even and constraint \((A.4)\) is the firm’s incentive-compatibility constraint.

To provide a benchmark, let us first derive the solution under the case where the firm’s project choice is publicly observable and contractible. In this case, the equilibrium is characterized by solving (PA.1) without imposing the incentive compatibility constraint \((A.4)\). Substituting the participation constraint \((A.3)\) into the objective function \((A.2)\), the principal-agent problem becomes

\[
\max_{x \in \{L, H\}, \ I > 0} \bar{p}_x R_x(I) - I.
\]

By the assumption on \(R_H\), the first-order condition is both necessary and sufficient for a global maximum. If the high risk project is optimal, then the corresponding optimal investment level, \(I_H\), is characterized by the following equation:

\[
\frac{\partial R_H(I_H)}{\partial I} = \frac{1}{\bar{p}_H}.
\]

\(^{16}\)Since the firm’s type is unknown to both the firm and the lender, all types of firms will offer the same loan contract to the lender.
Similarly, if the low risk project is optimal, the corresponding optimal investment level, $I_L$, is characterized by the following equation:

$$\frac{\partial R_H(I_L)}{\partial I} = \frac{1}{\sigma \bar{p}_L}.$$  \hfill (A.6)

Clearly, $\frac{\partial R_H(I_H)}{\partial I} > \frac{\partial R_H(I_L)}{\partial I}$ since $\sigma > \bar{p}_H / \bar{p}_L$. Then, $I_L > I_H$ follows immediately from the fact that $R_H(I)$ is strictly concave. Hence, the investment amount is higher if the low risk project is optimal. Indeed, the low risk project is optimal since

$$\sigma \bar{p}_L R_H(I_L) - I_L \geq \sigma \bar{p}_L R_H(I_H) - I_H > \bar{p}_H R_H(I_H) - I_H,$$

where the first inequality follows from the fact that $I_L$ is the maximum solution of $\sigma \bar{p}_L R_H(I) - I$, and the second inequality follows from the assumption that $\sigma > \bar{p}_H / \bar{p}_L$. Therefore, since the low risk project generates a higher expected return than the high risk project does, the equilibrium in the benchmark case is the allocation $[(I_L / \bar{p}_L, I_L), L]$, where $I_L$ solves (A.6).

Define $U$ as the sum of the firm’s expected terminal payoff and the lenders’ expected profits. Since the lenders earn zero expected profits, the firm’s expected terminal payoff is the same as $U$. In the benchmark case, $U_L$ is then given by

$$U_L = \sigma \bar{p}_L R_H(I_L) - I_L > 0.$$  \hfill (A.7)

Now we go back to the original setting in which the firm’s project choice is not publicly observable. The solution in the benchmark case may no longer be incentive compatible (i.e., risk incentive problem exists). It is because the firm would choose the high risk project, $H$, when it receives the loan contract $(I_L / \bar{p}_L, I_L)$, i.e., the incentive compatibility constraint (A.4) is violated. To ensure that the loan contract $(I_L / \bar{p}_L, I_L)$ is not incentive compatible, we require

$$\bar{p}_L \left[ \sigma R_H(I_L) - \frac{I_L}{\bar{p}_L} \right] < \bar{p}_H \left[ R_H(I_L) - \frac{I_L}{\bar{p}_L} \right],$$

which can be simplified as

$$\sigma < \frac{\bar{p}_H}{\bar{p}_L} + \left( 1 - \frac{\bar{p}_H}{\bar{p}_L} \right) \frac{I_L}{\bar{p}_L R_H(I_L)}.$$  \hfill (A.8)

97
It is easy to see that, because \( \sigma \bar{p}_L R_H(I_L) > I_L \), (A.8) implies \( \sigma < 1 \). Since it is not an interesting issue if the risk incentive problem does not exist, we assume (A.8) from now on.\(^{17}\)

We show in the next proposition that underinvestment is needed to resolve the risk incentive problem.

**Proposition A.1.** If there exists an \( I < I_L \) satisfying

\[
\sigma \bar{p}_L R_H(I) - I \geq \bar{p}_H R_H(I_H) - I_H,
\]

then the equilibrium under the setting with no auditing is the allocation \([(\hat{I}^*/\bar{p}_L, \hat{I}^*), L] \), where \( \hat{I}^* \) solves

\[
(\sigma \bar{p}_L - \bar{p}_H) R_H(I) = \left( 1 - \frac{\bar{p}_H}{\bar{p}_L} \right) I. \tag{A.10}
\]

Otherwise, the equilibrium is the allocation \([(I_H/\bar{p}_H, I_H), H] \), where \( I_H \) solves (A.5).

Proposition A.1 states that, with the risk incentive problem, underinvestment is required to motivate the firm to undertake the low risk project. The underlying intuition for why underinvestment can resolve the risk incentive problem is as follows. Notice that since \((1 - \sigma) R_H(I) \) is strictly increasing in the investment \( I \), the high risk project can be made less attractive in the solvency state by reducing the loan amount. In addition, because the probability of solvency is also lower for the high risk project, the low risk project may become optimal if the loan amount is sufficiently small, i.e., \( \hat{I}^* < I_L \). However, if condition (A.9) does not hold, there is no feasible loan contract that can induce the firm to choose the low risk project. In this case, undertaking the high risk project is the only credible outcome and the firm receives \( \bar{p}_H R_H(I_H) - I_H \).

In order to study the equilibrium outcome in which the low risk loan contract specified in proposition A.1 is indeed optimal, we assume condition (A.9) from now on. The firm's\(^{17}\)Note that condition (A.8) is not inconsistent with the earlier assumption that \( \sigma \bar{p}_L > \bar{p}_H \) since \( (1 - \frac{\bar{p}_H}{\bar{p}_L}) \frac{I_L}{p_L R_H(I_L)} > 0 \).
expected terminal payoff, \( \hat{U}^* \), is then given by

\[
\hat{U}^* = \sigma \bar{p}_L R_H(\hat{I}^*) - \hat{I}^*.
\]  

(A.11)

The following corollary provides some interesting comparative static properties of the equilibrium outcome.

**Corollary A.1:** Suppose the loan contract \((\hat{I}^*/\bar{p}_L, \hat{I}^*)\) is optimal, then the equilibrium loan amount and the firm’s expected terminal payoff increase with an increase in the expected solvency probability of the low risk project, \(\bar{p}_L\), or a decrease in the expected solvency probability of the high risk project, \(\bar{p}_H\).

The intuition for corollary A.1 is that as the expected solvency probability of the low risk project increases or that of the high risk project decreases, the low risk project becomes more attractive compared to the high risk project. Hence, the risk incentive problem is reduced and less underinvestment is required to resolve the problem. Since investment is productive, larger loan size implies larger expected terminal payoff for the firm.

The following numerical example illustrates the equilibrium outcome.

**Numerical Example:** Consider an example where \(R_H(I) = \sqrt{I}\) and \(R_L(I) = 0.8\sqrt{I}\). Suppose that \(\phi = 0.1\), \(p_G = 0.7\), \(p_B = 0.405\), and \(\bar{p}_L = 0.6\), then \(\bar{p}_H = 0.1 \times 0.7 + 0.9 \times 0.405 = 0.4345\), \(\bar{p}_H/\bar{p}_L = 0.72417\), and \(I_H = 0.047198\), \(I_L = 0.0576\) solve (A.5) and (A.6) respectively. The corresponding expected terminal payoffs of the firm are \(U_H = 0.047198 \times 0.21725 - 0.047198 = 0.047198\) and \(U_L = 0.8 \times 0.6 \times 0.24 - 0.0576 = 0.0576\) respectively. Observe that \(I_L/\bar{p}_L R_H(I_B) = 0.0576 \div (0.6 \times 0.24) = 0.4\), condition (A.8) is satisfied as \(0.8 < 0.72417 + 0.27583 \times 0.4 = 0.8345\). Thus the benchmark loan contract \((I_L/\bar{p}_L, I_L)\) cannot be attained. Then we find that \(\hat{I}^* = 0.02721 < I_H\) which solves (A.10). The corresponding expected terminal payoff of the firm is \(\hat{U}^* = 0.8 \times 0.6 \times 0.16496 - 0.02721 = 0.051968 > U_H\). Therefore, the loan contract \((\hat{I}^*/\bar{p}_L, \hat{I}^*)\) is indeed optimal.
A.2 The Setting with Auditing

Now we are ready to examine whether the firm can be made better off by the availability of external auditing services. Auditing services include: (1) performing an audit to the financial report; (2) expressing an opinion on the fairness of the report; and (3) supplying an audited report to the client. Since audits are voluntary in this model, the firm hires an auditor only if the firm’s expected terminal payoff increases by an amount greater than the cost of the audit. Our focus in this appendix is to understand the potential benefits of external auditing. Thus the discussion of the audit cost is intentionally suppressed.

It is worth mentioning that given the firm’s limited liability to pay, the function of auditors in this model is not only to mitigate the inefficiency caused by uncertainty regarding firm type, but also to provide ‘insurance’ to lenders. The first function arises from the fact that the audit provides information about the likelihood that the firm will be successful, thereby economizing on investment in a bad type firm. The second function, on the other hand, comes from the fact that the loan contract offered by the firm with a clean audited report includes a claim on the auditor’s wealth if the firm subsequently fails and the audit is found to have been conducted negligently. The firm can pay damages up to its return of investment, but even this amount is contingent on solvency. Auditors, on the other hand, are assumed to have ‘deep pockets’ sufficient to take care of the damages. However, the auditor is not hired to share risk since all players in the model are risk neutral. Instead, the purpose of the auditor’s liability is to discipline the auditor’s behaviour.

Formally, there are four stages in our incomplete information sequential game with auditing: the auditing stage, the borrowing stage, the investment stage and the litigation stage. The sequence of events proceeds as follows:

At the beginning of the period:

- Nature determines the firm’s type.
• The Auditing Stage:

- The firm chooses whether to hire an auditor to audit its financial report.
- Suppose the auditor is hired. Given the auditor’s liability rule and the prevailing auditing standards, the auditor completes the investigation, and issues an audited report.

• The Borrowing Stage:

- Based on the audited report, the firm proposes a loan contract to a lender who decides whether to approve the loan contract or not. If approved, the firm borrows.

• The Investment Stage:

- The firm chooses one of the two investment projects.

At the end of the period:

• Either the firm realizes revenue \( R_x(I) \) and pays the lender or the firm goes bankrupt.

• The Litigation Stage:

- The lender sues the auditor if the firm received a clean audited report but went bankrupt.
- The court conducts its investigation and assigns damage awards.

We now fully describe our game with auditing.

In the auditing stage, the firm chooses whether to hire an external auditor to investigate its financial condition.\(^{18}\) The payment to the auditor is the fee \( f \). Conforming to rules...

\(^{18}\)The firm does not need to hire an external auditor if it just wants to learn about its own type; presumably it is cheaper to ask its own accounting manager to do the job. Besides the agency problem between the owners and the accounting manager, the firm would also like to avoid the adverse consequence owing to the well-known signaling problem as the one described in Akerlof (1970) by making its financial report credible.
governing auditor independence that prohibit contingent fees, the auditor's fee cannot depend on her report. If the users of financial statements suspect a misstatement, the only remedy is to file a lawsuit.\textsuperscript{19} Although the auditor's fee is usually privately known by the firm and the auditor, the firm's action of hiring or not hiring an auditor is public information. For simplicity, we violate our previous assumption a bit by assuming that the firm has sufficient resources to pay the required audit fee even though it does not have its own resources to finance the project.\textsuperscript{20}

The auditor is a utility-maximizing, risk-neutral agent who strategically decides her fee, $f$, and quality of audit service, $q$.$^\text{21}$ Without loss of generality, $q$ is normalized such that $0 < q < 1$ and represents the probability that the auditor successfully detects a discrepancy between the firm's claim and its true type. Each auditor is assumed to have the same audit technology described as follows. First, recall that since there is no penalty for the firm's misstatement of its type, the firm will always report itself as a good type. However, the firm's claim is now subject to verification by the auditor. We assume that if the firm has reported a false type, the auditor's finding will be identical to the firm's message with probability $(1 - q)$ and will contradict the message with probability $q$; but if the firm makes a correct claim, there is no error to discover, and the auditor's finding will conform to the firm's message.\textsuperscript{22} To further simplify the analysis, we assume that the auditor will honestly report her finding in the auditor's opinion. If her finding agrees with the firm's reported

\textsuperscript{19}Note that a fixed audit fee plus a contingent liability is equivalent to a contingent fee scheme.

\textsuperscript{20}Alternatively, one might assume that the firm has no resources at all so that it must borrow the required audit fee in addition to the necessary funds for the investment project. Our assumption that the firm has money to pay the required audit fee is made to simplify the analysis. Relaxing this assumption will not affect the results qualitatively.

\textsuperscript{21}While admittedly unrealistic, the assumption that the auditor is risk neutral may be less so for auditors in large audit firms that are diversified across a large number of clients.

\textsuperscript{22}That is, we limit the auditor to making type II errors only. This assumption is consistent with the real-world auditing procedure whereby the firm's financial statements are modified only when the auditor discovers a discrepancy between these numbers and her findings.
state, the opinion is said to be clean. Otherwise, it is a qualified opinion. Although the auditor’s opinion is publicly observable, the quality of an audit is not.

After the audit is completed, the firm submits it to a potential lender with its loan application. More specifically, in the borrowing stage, the firm with an audited report requests a loan by offering a loan contract \((r, I)\) to a potential lender. We assume that if an auditor is hired, then the audited report must be included with the contract offer no matter what it contains. The firm knows the content of the audited report before it chooses the contract to be offered. The lender, when he decides whether to accept a proposed loan contract from the firm, must infer the information content of the audited report. Clearly, the latter is implied by the auditor’s quality choice. Even though the auditor’s quality choice is not publicly observable, the lender can make such an inference because he knows the elements of the auditor’s decision problem. The lender then uses the error inherent in the auditor’s optimal audit quality choice to revise his beliefs about the distribution of firm types upon observing the firm’s audited financial report. In equilibrium, the lender’s conjecture must be fulfilled.

Suppose the auditor’s individually rational audit quality choice is \(q\). Then the posterior probability of good and bad type firms will equal \(\phi + (1 - \phi)(1 - q)\) and \((1 - \phi)q\), respectively. By our audit technology assumption, if the audited report is a qualified report, the firm is a bad type for sure. That is, no mistake will be made by the auditor. Furthermore, suppose the following conditions hold:

\[
\sigma < \frac{p_B}{\bar{p}_L} + \left(1 - \frac{p_B}{\bar{p}_L}\right) \frac{I_L}{\bar{p}_L R_H(I_L)}, \tag{A.12}
\]

where \(I_L\) solves (A.6); and there exists an \(I < I_L\) satisfying

\[
\sigma \bar{p}_L R_H(I) - I \geq p_B R_H(\tilde{I}_H) - \tilde{I}_H, \tag{A.13}
\]

where \(\tilde{I}_H\) solves (A.5) with \(\bar{p}_H\) replaced by \(p_B\).\(^{23}\) Then the optimal loan contract for the bad type firm will be the low risk loan contract, \((I_B/p_B, I_B)\), where \(I_B\) solves (A.10) with \(p_B\).\(^{23}\)

\(^{23}\)The interpretation of (A.12) and (A.13) are similar to those of (A.8) and (A.9).
\(p_H\) replaced by \(p_B\). Notice that because there is no uncertainty regarding firm types once the firm is identified as a bad type firm, the optimal low risk loan contract for the bad type firm is independent of the audit quality \(q\). On the other hand, suppose that the auditor issues a clean audited report to the firm. Then using Bayes’ rule, the posterior conditional probability that the firm is a good type is given by

\[
\psi(q) = \frac{\phi}{\phi + (1 - \phi)(1 - q)} \in (\phi, 1).
\]

It is not difficult to verify that

\[
\frac{\partial \psi(q)}{\partial q} = \frac{\phi(1 - \phi)}{[\phi + (1 - \phi)(1 - q)]^2} > 0,
\]

and

\[
\frac{\partial^2 \psi(q)}{\partial q^2} = \frac{2\phi(1 - \phi)^2}{[\phi + (1 - \phi)(1 - q)]^3} > 0.
\]

That is, upon observing a clean audited report, the posterior conditional probability that the firm is a good type firm is strictly increasing and strictly convex in audit quality. Therefore, the uncertainty regarding firm type can be controlled by increasing audit quality.

The firm stays in the borrowing stage until it obtains a loan, then it enters into the investment stage by choosing its optimal project given its loan amount. Neither the auditor nor the lender can observe the firm’s project choice.

We now know that if the firm gets a qualified opinion, the optimal loan contract will be the low risk loan contract, \((I_B/\bar{p}_L, I_B)\). Accordingly, the ex-post terminal payoff (excluding audit fee) for the bad type firm is given by \(\sigma \bar{p}_L R_L(I_B) - I_B\). In the sequel, we will focus on the derivation of the optimal loan contract for the firm if it obtains a clean audited report from the auditor.

Suppose that the firm with a clean audited report obtains a loan contract \((r, I)\) and undertakes project \(x\). With audit quality \(q\) and common posterior assessment of good type firm \(\psi(q)\), the firm’s conditional probability of solvency if the low risk project is undertaken is given by \(\psi(q)\bar{p}_L + [1 - \psi(q)]\bar{p}_L = \bar{p}_L\), which is the same as if there is no additional information.
provided by the audited report. On the other hand, if the high risk project is undertaken, the firm’s conditional probability of solvency is given by

$$\pi_H(q) = \psi(q)p_G + [1 - \psi(q)]p_B$$

$$= \frac{\phi p_G + (1 - \phi)(1 - q)p_B}{\phi + (1 - \phi)(1 - q)} \in (\bar{p}_H, p_G).$$

It is easy to see that

$$\frac{\partial \pi_H(q)}{\partial q} = (p_G - p_B) \frac{\partial \psi(q)}{\partial q} > 0.$$ 

The firm’s conditional probability of solvency is strictly increasing in audit quality. In this case, the firm’s ex-post expected profit (excluding audit fee) is $$\pi_H(q)[R_H(I) - r]$$. Notice that given a clean audited report is issued, audit quality will have an effect on the firm’s conditional probability of solvency only when the high risk project is undertaken.

For the loan contract ($r, I$), the lender’s expected loan repayment received from the firm with a clean audited report is given by $$\pi_H(q)r$$. However, in making a loan decision, the lender will also take into account the potential damage award receivable from the auditor. The availability and amount of the damage awards hinges on the auditor’s liability rule in place.

It is noteworthy that the auditor is a rational economic agent. If there is no liability, then given that contingent audit fees are illegal and audit quality is costly to the auditor, the auditor would provide the least amount of quality once she is hired. As such, the auditor will issue a clean audited report regardless of the true type of the firm. Of course, all users of the audited report will recognize the auditor’s self-interested behavior and react accordingly. Thus the audited report will have no information content in the sense that the posterior assessment of a good type firm will be the same as the prior. That is, we are back to the setting where there is no auditor. Thus, we introduce a court system in which users of an audited report can sue the auditor if they believe that there is a material misstatement made in the audited report. The threat of third-party suits produces equilibria in which the audit report is informative.
Formally, in the litigation stage, the lender decides, based on the auditor’s report and the solvency position of the firm, whether to sue the auditor.\textsuperscript{24} For simplicity, we impose the following limitation on the types of suits that can be tried. The court only hears cases in which a party can claim to have been damaged by the auditor’s report.\textsuperscript{25} Hence, a lender can sue only if the firm was bankrupt after the auditor issued a clean audited report.

The auditor’s liability rule adopted by the court is assumed to be a vague negligence standard under the joint and several liability regime, which seems to be the most descriptive scenario of the current situation. Under the vague negligence standard, the auditor is held liable if she fails to exercise due care in her duties as an auditor. However, what is due care is not clearly defined. Normally, GAAS provide some useful guidance, but courts do not always go by GAAS. Thus, there always exists some uncertainty in determining whether due care is met. More specifically, let $q^*$ be the audit quality that is defined by GAAS.\textsuperscript{26} In a lawsuit, the court looks for discrepancies between the auditor’s report and the firm’s true type; we define such a discrepancy as an audit failure. We assume that, if an audit failure exists, based on the evidence provided by the auditor, the court discovers it with probability $\nu^* > 0$ if the auditor chooses audit quality $q \geq q^*$ and is sued. On the other hand, if the auditor chooses audit quality $q < q^*$ and is sued, the court discovers an audit failure with probability $\nu^0 > \nu^*$. Thus, our model captures the notion that adherence to GAAS can be a reasonable defense, and yet assuming that $\nu^*$ is greater than zero, we maintain that the courts are not infallible. As such, a reasonably diligent auditor who has conducted an audit in accordance with GAAS still faces some probability of having an error in the report and not being able to provide sufficient evidence of due care to clear herself. Moreover, we assume in our model

\textsuperscript{24}We assume that the lender can sue the auditor but not the firm. It is because the auditor, who has ‘deep pockets’, is the only potential defendant that is still solvent in the model. In the other words, this assumption is adopted to permit us to avoid questions of whether the firm can be sued when it is bankrupt and how damage payments are to be divided between a firm and its auditor.

\textsuperscript{25}This is consistent with Schultz and Pany (1980) whereby they present an overview of the public accountant’s civil liability and address four elements of proof which are central to the determination of this liability, namely, (i) the plaintiff must incur financial damages, (ii) there must be a material omission or misstatement in the financial statements, (iii) there must be reliance on the statements by the plaintiff, and (iv) the accountant’s conduct must be deficient in some respect for determination of liability.

\textsuperscript{26}The standards are usually set by the accounting profession and are taken to be exogenously determined.
that the prospect of facing litigation arising from substandard audits is the primary force that motivates an auditor to adhere to auditing standards. As such, once the prescribed standards are met, the only cost to the auditor is her cost of quality. Hence, the auditor will never exceed the quality level defined by GAAS. It is also worthwhile mentioning that the liability losses fall when the auditor complies with the prescribed auditing standards for two reasons: there is less chance of an audit failure and the auditor is less likely to be found negligent even if an audit failure occurs.

If the audit report is accurate (i.e., when there are no discrepancies between the auditor's report and the firm's true type), there is no audit failure to uncover, and the court's ruling agrees with the report. On the other hand, if the court discovers an audit failure and the auditor is judged negligent, the lender will get full compensation. Let $D(I) \geq 0$ be the damage award paying to the lender. The damage award in principle should be directly linked to the actual loss incurred by the lender, which in turn is a function of the actual loan size $I$. Since $I$ is endogenously determined, the penalty award, $D(I)$, is also endogenously determined. For simplicity, we assume that $D(I) = I$. When the firm goes bankrupt in an audit failure, the auditor, who is held jointly and severally liable, must also pay damages on behalf of the bankrupt firm, even if the auditor is assessed as being responsible for only a small fraction of the total liability.

Suits also involve legal costs which have to be paid by both the auditor and the lender. Under the American Rule of litigation cost allocation, each side will bear its own cost. The winning party cannot recover its litigation cost from the losing party. Let the legal cost for

\[\text{\footnotesize 27In general, the goal of the penalty has two parts. It aims at fairly compensating the victims and adequately deterring the auditor from any 'wrong-doing'. The endogenity of the penalty award in our model renders any declared liability standard fair, and accordingly, narrows our focus to the deterrence aspect.}\]

\[\text{\footnotesize 28The qualitative results do not change as long as the damage award } D(I) \text{ is assumed to be an increasing function of the actual loan size } I. \text{ We argue that such a direct link between auditor's liability and the loss incurred by the lender should bring out a more effective audit incentive structure because the auditor's liability will be more sensitive to her audit quality in this case.}\]

\[\text{\footnotesize 29Note that many statutes on the state and federal levels in the United States also provide for the shifting of fees under particular circumstances. Some of these rules resemble the English system whereby the loser is typically forced to bear the winner's legal expenses. The English Rule of litigation cost allocation is not considered in this study.}\]
the auditor be \( \kappa \in (0, I) \).\(^{30}\) On the other hand, we are not going to model an explicit legal

cost for the lender but simply assume that the lender always sues the auditor whenever the

firm goes bankrupt.\(^{31}\) This assumption seems close to the current litigation environment

wherein trial lawyers always seem to be available to work on a contingency basis. Therefore,
the lender has little to lose to sue the auditor. Of course, the lawyers must expect to recover
their costs. We assume that trial lawyers are risk-neutral and incur a cost of \( \iota \) for litigating.

If the lender wins the suit, the lawyers keep a portion of the damages collected. Let the
fraction of the damages paid to trial lawyers as contingent fees be \( 1 - \alpha \), \( \alpha \in (0, 1) \). We
assume that the lawyers' legal costs are always smaller than the expected contingent fees
such that they are always willing to take the case. Hence, in this framework, the lender's
strategy about whether to sue the auditor is taken as fixed. As such, we can analyze the
auditor's optimal response to a potential legal liability.\(^{32}\)

Notice that the effective auditor's liability for damage, \( L(I, w) \), also hinges on the au-
ditor's attachable wealth, \( w \), i.e., \( L(I, w) = \min \{ I, w \} \).\(^{33}\) For simplicity, we assume that
the liability rule/standard combination \( \{[L(\cdot), \nu^*], q^* \} \) is typical as defined by Dye (1993),
such that there always exists some auditors with attachable wealth, \( w \), who would adopt a
standard quality audit if she is hired, i.e., \( q = q^* \). Moreover, we assume that the firm would
only hire an auditor who is very likely to comply with the prescribed auditing standards
\( q^* \). Since big audit firms are more likely to comply with the audit standards than small
audit firms because, other things being equal, larger audit firms have more wealth to lose,
we assume that the firm will go to a big audit firm for audit services.\(^{34}\)

\(^{30}\)If \( \kappa \geq I \), the auditor is better off settling the suit out-of-court. The possibility of settlement is not
considered in this study.

\(^{31}\)This is similar to the limited litigation case considered by Melumad and Thoman (1990).

\(^{32}\)It is worthwhile mentioning that an 'effective audit' equilibrium can exist only if the auditor is motivated
to be diligent because of her fear of being sued, and the lender is motivated to sue the auditor when the firm

goes bankrupt because of the potential to recoup his losses and legal costs. If the lender is not motivated
to sue, the auditor is not motivated to be diligent. In our model, the lender is motivated to sue the auditor
because adherence to GAAS does not provide absolute assurance that would dismiss the auditor from any
liability and the lender has little to lose to sue the auditor. This is the key to why our approach works.

\(^{33}\)The auditor's attachable wealth is a net amount after considering the audit fee, the cost of performing
the audit and the legal costs.

\(^{34}\)This consistent with the common perception that audit quality is positively correlated with audit firm
To further simplify the analysis, we assume that, given audit standard $q^*$ and a clean audited report, the optimal project for the firm (and also the lender) is the high risk project. That is, we require that, for all $I > 0$,

$$\{ \pi_H(q^*) p_G + [1 - \pi_H(q^*)] p_B \} R_H(I) > \sigma \bar{p}_L R_H(I),$$

which is equivalent to

$$q^* > \frac{\sigma \bar{p}_L - p_B - \phi (p_G - p_B)}{(1 - \phi) (\sigma \bar{p}_L - p_B)} \in (0, 1). \quad (A.14)$$

In the event that the audit is determined to have been conducted negligently, the lender’s expected damage recoverable from the auditor is given by

$$\Omega_H(q^*, I) = [1 - \psi(q^*)] (1 - p_B) \nu^* \alpha I$$

$$= \frac{(1 - \phi)(1 - q^*)(1 - p_B) \nu^* \alpha I}{\phi + (1 - \phi)(1 - q^*)} > 0.$$

Observe that

$$\frac{\partial \Omega_H}{\partial I} = \frac{(1 - \phi)(1 - q^*)(1 - p_B) \nu^* \alpha I}{\phi + (1 - \phi)(1 - q^*)} > 0$$

for all $q^* \in (0, 1)$ and

$$\frac{\partial \Omega_H}{\partial q^*} = -\frac{\phi (1 - \phi)(1 - p_B) \nu^* \alpha I}{[\phi + (1 - \phi)(1 - q^*)]^2} < 0,$$

for all $I > 0$. That is, the lender’s expected damage recoverable from the auditor is strictly increasing in the firm’s investment level but strictly decreasing in prescribed audit quality.

The optimal high risk loan contract for the ‘good type’ firm is then characterized by solving the following principal-agent problem:\textsuperscript{35} (PA.3)

$$\max_{I > 0, r > I} \quad \pi_H(q^*) [R_H(I) - r]$$

s.t. \quad $\pi_H(q^*) r + \Omega_H(q^*, I) - I \geq 0.$ \quad (A.15) (A.16)

\textsuperscript{35}The term ‘good type’ is in quotes because the firm may not actually be good, but rather deemed as good under the given audit technology.
Substituting the participation constraint (A.16) into the objective function (A.15), the principal-agent problem (PA.3) becomes

$$\max_{I>0} \pi_H(q^*) R_H(I) + \Omega_H(q^*, I) - I.$$ 

By the assumption on $R_H$, the first-order condition is both necessary and sufficient for a global maximum. The optimal investment level, $I_H$, is then characterized by the following equation:

$$\frac{\partial R_H(I_H)}{\partial I} = \frac{1 - \frac{\partial \Omega_H}{\partial I}}{\pi_H(q^*)} = \frac{\phi + (1 - \phi)(1 - q^*)(1 - (1 - p_B)\nu^*\alpha)}{\phi p_G + (1 - \phi)(1 - q^*)p_B}.$$  (A.17)

Hence, the optimal high risk loan contract for the 'good type' firm is given by

$$\left[ I_H(q^*) - \Omega_H(q^*, I_H(q^*)) \right] \cdot \pi_H(q^*), \quad I_H(q^*)$$

where $I_H(q^*)$ solves (A.17). The expected terminal payoff of the firm (excluding audit fee) under the setting with auditing is then given by

$$U(q^*) = [\phi + (1 - \phi)(1 - q^*)][\pi_H(q^*) R_H(I_H(q^*)) + \Omega_H(q^*, I_H(q^*)) - I_H(q^*)]$$

$$+ (1 - \phi)q^*[\sigma p_L R_H(I_B) - I_B].$$  (A.18)

The gross benefit of an audit with audit quality $q^*$ can then be established by comparing (A.18) with (A.11). As uncertainty regarding the firm's type is reduced by the audit, we would expect such service to benefit the firm. Moreover, as the prescribed audit quality $q^*$ increases, one in general would expect the gross benefit of the audit increases since capital resources will be more efficiently allocated. The next proposition states the necessary and sufficient condition under which the gross benefit of the audit is strictly positive and increasing in the quality prescribed by the prevailing auditing standards.

**Proposition A.2.** Given (A.14), the gross benefit of an audit, $b(q^*) \equiv U(q^*) - \bar{U}^*$, is strictly positive. Furthermore, $b(q^*)$ is increasing in the prescribed audit quality, $q^*$, if, and
only if, \(36\)

\[
(1 - p_B)\nu^* \alpha I_H(q^*) < \left[ \sigma \tilde{p}_L R_H(I_B) - I_B \right] - \left[ p_B R_H(I_H(q^*)) - I_H(q^*) \right].
\]  

(A.19)

Proposition A.2 demonstrates that if condition (A.19) holds, the gross benefit of an audit, \(b(q^*)\), which is strictly positive given (A.14), is strictly increasing in the prescribed quality, \(q^*\). Condition (A.19) has an intuitive economic interpretation. The right-hand-side of (A.19) is the marginal benefit to the firm of an increase in the prescribed audit quality. The marginal benefit stems from the result of a better investment decision of the bad type firm. On the other hand, the left-hand-side of (A.19) is the marginal cost to the firm of an increase in the prescribed audit quality. The marginal cost arises from the fact that, if the bad type firm is identified, the lender will lose the expected damage award from the auditor. Since the debt market is perfectly competitive, this implies that the lender will require a higher loan repayment from the firm. Thus, condition (A.19) simply says that the marginal benefit to the firm of an increase in the prescribed audit quality is higher than the marginal cost. Of course, given that audits are voluntary in this model, the firm is willing to hire an auditor if, and only if, the gross benefit of an audit is greater than the required audit fee, i.e., \(b(q^*) \geq f\).

Before ending this appendix, we use our previous numerical example to demonstrate the value of an audit with 95% audit assurance, i.e., if a clean audited report is issued, the posterior conditional probability that the firm is a good type is equal to 0.95.

Numerical Example Continued: Suppose that the standard audit quality is \(q^* = 0.99415\). Then \(\psi(q) = 0.1 \div [0.1 + (1 - 0.1) \times (1 - 0.99415)] = 0.95\). That is, an audit conducted under the prevailing auditing standards provides 95% audit assurance. Suppose further that \(\nu^* = 0.1\) and \(\alpha = 0.5\). It is easy to calculate that \(\tau_H(q^*) = 0.95 \times 0.7 + \ldots \)

\(^{36}\)Using (A.13), we have \(\sigma \tilde{p}_L R_H(I_B) - I_B \geq p_B R_H(I_H) - I_H > p_B R_H(I_H(q^*)) - I_H(q^*)\), where the second inequality follows from the fact that \(I_H\) maximizes \(p_B R_H(I) - I\). Then the right-hand-side of (A.19) is strictly positive.
$0.05 \times 0.405 = 0.68525$, $I_B = 0.05325$, $I_H(q^*) = 0.11774$, $\Omega_H(q^*, I_H(q^*)) = 0.0001752$ and 

$U(q^*) = [0.1 + (1 - 0.1)(1 - 0.99415)][0.68525 \times 0.34313 - (0.11774 - 0.0001752)] + (1 - 0.1) \times 0.99415 \times [0.8 \times 0.6 \times 0.23076 - 0.05325] = 0.063836$. Hence, the value of a standard audit service is given by $b(q^*) \equiv U(q^*) - \bar{U} = 0.063836 - 0.051968 = 0.011868$.

### A.3 Concluding Remarks

This appendix presents a model in which audit services purchased by the firm provide information to both the firm and the potential lender about the firm’s current financial condition. The firm’s current financial condition affects the firm’s investment decision and risk incentive, which in turn determine the firm’s ability to repay and the lender’s willingness to contract. We do not assume that the audit is mandated; rather, the firm has an incentive to hire an external auditor to attest its financial statements. We show that without the information provided from an audit, the firm will underinvest and the socially desirable project will be foregone. Reducing such an inefficiency crucially depends on the optimal use of the information on hand in the initial situation and additional information which can be used to further reduce the residual inefficiency. We then show that the presence of an auditor mitigates the inefficiency caused by imperfect information and the resulting risk incentive problem between the lender and the firm. Thus, our model provides a theoretical link between auditing and the efficiency of the capital market.
APPENDIX B

Proofs of Propositions

Proof of Proposition 2.1.

Suppose audit firm i’s auditing cost to serve client z, m(||i − z||), is lower than m(||l − z||) for all j ≠ i. Then audit firm i is said to have a cost advantage to serve client z. Let \( \bar{f}^p_i \) be the lowest profit-maximizing audit fee quoted to client z when audit firm i cannot charge more than \( \min_{j \neq i} m(||l_j − z||) \). Then, it is clear that \( \bar{f}^p_i \) must take its greatest value, i.e.,

\[
\bar{f}^p_i = \min_{j \neq i} m(||l_j − z||).
\]

Thus, for a given client z, Bertrand competition drives audit fees down to the level of the second lowest-cost to that client, which then allows the lowest-cost audit firm to serve the client and charge an audit fee of that amount. If the second lowest-cost audit firm charged over its cost, the lowest-cost audit firm would charge at this amount and get the client. This would induce the second lowest-cost audit firm to cut its audit fee.

On the other hand, suppose audit firm i has no cost advantage to serve client z, i.e.,

\[
m(||l_i − z||) > \min_{j \neq i} m(||l_j − z||).
\]

Then for any audit fee \( f_i^p > m(||l_i − z||) \), audit firm i has no demand since it will always be undercut by at least one of its rivals. Similarly, audit firm i makes no profit when \( m(||l_i − z||) = \min_{j \neq i} m(||l_j − z||) \). In either case, audit firm i earns zero profits and pricing at \( m(||l_i − z||) \) is optimal.  \( \square \)
Proof of Proposition 2.2.

Since the first term in the expression of \( W(l_1, l_2, \ldots, l_n) \) is independent of \((l_1, l_2, \ldots, l_n)\), then for any \((l^*_1, l^*_2, \ldots, l^*_n)\) that maximizes \( W(l_1, l_2, \ldots, l_n) \), it also minimizes \( C(l_1, l_2, \ldots, l_n) \).
\( \square \)

Proof of Proposition 2.3.

Since by assumption \( m(||l_i - z||) \) is continuous, \( W(l_1, l_2, \ldots, l_n) \) is also continuous because minimization and integration preserve continuity. By the Weierstrass theorem, \( W(l_1, l_2, \ldots, l_n) \) has a maximum on the compact set \( Z \).
\( \square \)

Proof of Proposition 2.4.

For all \( l_i \in Z, i = 1, 2, \ldots, n, \)

\[
\Pi_i(l_i, l_{-i}, f_i^{z*}, f_{-i}^{z*}) = \frac{\delta(1 - \delta^T)}{1 - \delta} \int_{Z_i} [\min_{j \neq i} m(||l_j - z||) - m(||l_i - z||)] h(z) \, dz
\]

\[
= \frac{\delta(1 - \delta^T)}{1 - \delta} \left[ \int_{Z} [\min_{j \neq i} m(||l_j - z||)] h(z) \, dz - \int_{Z} [\min_{j \neq i} m(||l_i - z||)] h(z) \, dz \right]
\]

\[
= \frac{\delta(1 - \delta^T)}{1 - \delta} \int_{Z} [\min_{j \neq i} m(||l_j - z||)] h(z) \, dz - C(l_i, l_{-i})
\]

\[
= \frac{\delta(1 - \delta^T)}{1 - \delta} \int_{Z} [\min_{j \neq i} m(||l_j - z||)] h(z) \, dz + W(l_i, l_{-i}).
\]

If \((l^*_i, l^*_{-i})\) are equilibrium specializations, then \( \Pi_i(l^*_i, l^*_{-i}, f_i^{z*}, f_{-i}^{z*}) \geq \Pi_i(l_i, l_{-i}, f_i^{z*}, f_{-i}^{z*}) \), which is equivalent to

\[
\frac{\delta(1 - \delta^T)}{1 - \delta} \int_{Z} [\min_{j \neq i} m(||l_j^* - z||) - b^z] h(z) \, dz + W(l_i^*, l_{-i}^*)
\]

\[
\geq \frac{\delta(1 - \delta^T)}{1 - \delta} \int_{Z} [\min_{j \neq i} m(||l_j^* - z||) - b^z] h(z) \, dz + W(l_i, l_{-i}).
\]

The condition specified in the proposition follows immediately. Finally, the existence of a specialization configuration which satisfies that condition has been proved by proposition 2.3. This completes the proof. \( \square \)
Proof of Proposition 2.5.

The proof is by induction. First, suppose that it is an equilibrium for all audit firms to choose the same specialization, say $l_0^*$. Then the unique audit fee equilibrium schedule is given by the Bertrand solution, i.e., $f_1^{z*} = f_2^{z*} = \ldots = f_n^{z*} = m(||l_0^* - z||)$. Consequently, the corresponding profits are necessarily zero. On the other hand, consider any configuration with one audit firm, say audit firm $\tilde{i}$, who deviates from the proposed specialization equilibrium by choosing a distinct specialization, $l_\tilde{i} \neq l_0$. The market region exclusively served by that audit firm is denoted by $Z_\tilde{i}$ which is nonempty. Then, the equilibrium audit fee charged by audit firm $\tilde{i}$ in its own market region $Z_\tilde{i}$ will be $m(||l_0^* - z||) > m(||l_\tilde{i}^* - z||)$, where the inequality follows from the definition of $Z_\tilde{i}$. Thus, choosing a specialization away from $l_0^*$ must increase the profit of audit firm $\tilde{i}$. Therefore, the proposed ‘$n$-firms pooling equilibrium’ will be broken. The same logic can then be apply to break any proposed ‘$n-1$-firms pooling equilibrium’, with $n \geq 3$. This completes the proof. □

Proof of Proposition 3.1.

Assume an interior solution exists. Then, since $W(l_1, l_2)$ is strictly concave in $(l_1, l_2)$ (or, equivalently, $C(l_1, l_2)$ is strictly convex in $(l_1, l_2)$), the first-order conditions are necessary and sufficient for a global maximum (minimum). Solving the first-order conditions yields the result. □

Proof of Proposition 3.2.

Observe that

$$
\Pi_1(l_1, l_2) = \int_0^{l_1} c(l_2 - z) \, dz + \int_{l_1}^{l_1+\frac{\theta_2-\theta_1}{(1+\beta)}} c(l_1 + l_2 - 2z) \, dz \\
+ \int_{l_1+\frac{\theta_2-\theta_1}{(1+\beta)}}^{l_1+\frac{\theta_2-\theta_1}{(1+\beta)}} \{c[\beta(l_2 - z) - (z - l_1)] - k(l_2 - l_1)\} \, dz \\
+ \int_0^{l_1} \{c[l_2 - z - \beta(l_1 - z)] + k(l_2 - l_1)\} \, dz
$$

115
\[ + \int_{l_1}^{l_1 + l_2} \{e[l_2 - z - \beta(z - l_1)] + k(l_2 - l_1) \} \, dz. \]

Then, by Leibnitz's rule,
\[
\frac{\partial \Pi_1(l_1, l_2)}{\partial l_1} = \frac{2}{2(1 + \beta)^2c} \left[ (3 + \beta)(l_1 + l_2)(e^2 + \beta^2 e^2 + 2ck - 2\beta ck + 2k^2) \right. \\
-2l_1(4e^2 + 4\beta e^2 + 6\beta^2 e^2 + 2\beta^3 e^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2) \bigg],
\]

and
\[
\frac{\partial^2 \Pi_1(l_1, l_2)}{\partial l_1^2} = \frac{-(5e^2 + 7\beta e^2 + 9\beta^2 e^2 + 3\beta^3 e^2 + 8ck + 6k^2 + 2\beta k^2)}{2(1 + \beta)^2c} < 0.
\]

By the same token, one can get \( \frac{\partial \Pi_2(l_1, l_2)}{\partial l_2}, \frac{\partial^2 \Pi_2(l_1, l_2)}{\partial l_2^2} \) and easily verify that
\[
\frac{\partial^2 \Pi_2(l_1, l_2)}{\partial l_2^2} < 0.
\]

Hence, the first-order conditions are necessary and sufficient for a global maximum. For the symmetric specialization equilibrium, \( l^*_1 = 1 - l^*_2 \). Thus, solving the equation \( \frac{\partial \Pi_1(l^*_1, l^*_2)}{\partial l_1} = 0 \) yields \( l^*_1 \) as reported in the proposition.

Finally, it remains to show the global stability of the given symmetric specialization equilibrium. Uniqueness then follows immediately. This can be done by proving that both audit firms' reaction functions are upward sloping and audit firm 1's reaction function is everywhere steeper than that of audit firm 2 on the \( l_1-l_2 \) plane, so that they intersect at most once. \( \Box \)

**Proof of Corollary 3.1.**

(i) It is straightforward to show that
\[
\lim_{\beta \to 1} l^*_1 - l^*_1 = \frac{(2k - c)k}{4(4c^2 + ck + 2k^2)} < 0,
\]
for all \( c > 0 \) and \( 0 < k < c/2 \).
(ii) Similarly, one can get
\[
\lim_{k \to 0} l_1^* - l_1^W = \frac{1 + 5\beta + 2\beta^2}{4(2 + 2\beta + 3\beta^2 + \beta^3)} > 0,
\]
for all 0 < \beta < 1.

(iii) Finally, it is easy to show that
\[
\lim_{\beta \to 1, k \to 0} l_1^* = \frac{1}{4}. \quad \square
\]

Proof of Lemma 3.1.

Denote \( D = 4c^2 + 4\beta c^2 + 6\beta^2 c + 2\beta^3 c^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2 > 0. \)

(i) Partially differentiating \( l_1^* \) with respect to \( \beta \) yields
\[
\frac{\partial l_1^*}{\partial \beta} = -\frac{c(1 + \beta)}{2D^2} \left[ 8c^3 + 27c^2 k + \beta c^2 k + 32ck^2 + 20\beta ck^2 + 4\beta^2 ck^2 \\
+ 10k^3 + 2\beta k^3 + (4\beta c^3 + 9\beta c^2 k)(1 - \beta) + 3\beta c^2 k(1 - \beta^2) \right] \\
< 0,
\]
for all \( c > 0, 0 < k < \beta c/2 \) and 0 < \( \beta \) < 1.

(ii) Partially differentiating \( l_1^* \) with respect to \( c \) yields
\[
\frac{\partial l_1^*}{\partial c} = -\frac{k(3 + \beta)(1 + \beta)^2}{2D^2} \left( c^2 - 3\beta^2 c^2 + 4ck + 4\beta ck + 2k^2 \right) \\
< (>) 0,
\]
if \( c^2 - 3\beta^2 c^2 + 4ck + 4\beta ck + 2k^2 > (<) 0. \)

(iii) Similarly, partially differentiating \( l_1^* \) with respect to \( k \) yields
\[
\frac{\partial l_1^*}{\partial k} = \frac{c(3 + \beta)(1 + \beta)^2}{2D^2} \left( c^2 - 3\beta^2 c^2 + 4ck + 4\beta ck + 2k^2 \right) \\
> (<?) 0,
\]
if \( c^2 - 3\beta^2 c^2 + 4ck + 4\beta ck + 2k^2 > (<) 0. \) \quad \square
Proof of Lemma 3.2.

Denote $D = 4c^2 + 4\beta c^2 + 6\beta^2 c + 2\beta^3 c^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2 > 0$.

(i) Partially differentiating $\Pi^*$ with respect to $l_1^*$ yields

$$\frac{\partial \Pi^*}{\partial l_1^*} = \frac{2}{c(1+\beta)^2} \left[ c^2 - 3\beta c^2 + \beta^2 c^2 + \beta^3 c^2 + 5ck - 6\beta ck - 3\beta^2 ck + 6k^2 + 2\beta k^2 
- l_1^*(5c^2 + \beta c^2 + 7\beta^2 c^2 + 3\beta^3 c^2 + 12ck - 8\beta ck - 4\beta^2 ck + 12k^2 + 4\beta k^2) \right]
= -\frac{c}{D} \left[ (7 - \beta^4)c^2 + 10\beta c^2 + 14\beta^2 c^2 + 2\beta^3 c^2 + 12ck + 4\beta ck + 12\beta^2 ck
+ 4\beta^3 ck + 8k^2 \right] < 0,$$

for all $c > 0$, $0 < k < \beta c/2$ and $0 < \beta < 1$.

(ii) Partially differentiating $S^*$ with respect to $l_1^*$ yields

$$\frac{\partial S^*}{\partial l_1^*} = -\frac{1}{c(1+\beta)^2} \left[ 9\beta c^2 - c^2 + \beta^2 c^2 - \beta^3 c^2 - 10ck + 12\beta ck + 6\beta^2 ck - 12k^2 - 4\beta k^2
+ 2l_1^*(3c^2 - 5\beta c^2 + \beta^2 c^2 + \beta^3 c^2 + 12ck - 8\beta ck - 4\beta^2 ck + 6k^2 + 4\beta k^2) \right]
= \frac{c}{D} \left[ (5 - 4\beta^4)c^2 + 10\beta c^2 + 16\beta^2 c^2 + 2\beta^3 c^2 + 7ck + (5c - 8k)\beta k
+ 21\beta^2 ck + 7\beta^3 ck + 2k^2(1 - \beta^2) \right] > 0,$$

for all $c > 0$, $0 < k < \beta c/2$ and $0 < \beta < 1$.

(iii) Similarly, partially differentiating $W^*$ with respect to $l_1^*$ yields

$$\frac{\partial W^*}{\partial l_1^*} = c(1+\beta)(1 - 4l_1^*)
= -\frac{c(1+\beta)}{D} \left( 2c^2 - 2\beta c^2 + 5ck - 6\beta ck - 3\beta^2 ck + 6k^2 + 2\beta k^2 \right)
< (>) 0,$$

if $2c^2 - 2\beta c^2 + 5ck - 6\beta ck - 3\beta^2 ck + 6k^2 + 2\beta k^2 c > (<) 0$. □
Proof of Proposition 3.3.

Denote $D = 4c^2 + 4\beta c^2 + 6\beta^2 c + 2\beta^3 c^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2 > 0$.

The proposition then follows from the fact that

$$l_1' - \frac{1}{4} = \frac{2c^2 - 2\beta c^2 + 5ck - 6\beta ck - 3\beta^2 ck + 6k^2 + 2\beta k^2}{4D} > 0,$$

which is implied by (A2). □

Proof of Proposition 3.4.

The proposition follows directly from applying the results of lemma 3.1 and imposing (A2). □

Proof of Lemma 3.3.

Denote $D = 4c^2 + 4\beta c^2 + 6\beta^2 c + 2\beta^3 c^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2 > 0$.

(i) Partially differentiating $\Pi^*$ with respect to $c$ yields

$$\frac{\partial \Pi^*}{\partial c} = \frac{1}{2c^2(1 + \beta)^2}\left[c^2 + 7\beta c^2 + \beta^2 c^2 - \beta^3 c^2 + 6k^2 + 2\beta k^2 + 4l_1^*(c^2 - 3\beta c^2 + \beta^2 c^2\right.$

$$+ \beta^3 c^2 - 6k^2 - 2\beta k^2) - 2l_1^*(5c^2 + \beta c^2 + 7\beta^2 c^2 + 3\beta^3 c^2 - 12k^2 - 4\beta k^2)\right]\right.\]

$$= \frac{1}{4D^2}\left[(35 - 3\beta^6)c^4 + (99 - 37\beta^6)\beta c^4 + 203\beta^2 c^4 + 219\beta^3 c^4 + 165\beta^4 c^4 + 53\beta^5 c^4\right.$

$$+ 16(7 - \beta^4)c^3 k + 160\beta c^3 k + 224\beta^2 c^3 k + 32\beta^3 c^3 k + 2(91 - 3\beta^4)c^2 k^2\right.$

$$+ 6(33 - \beta^4)\beta c^2 k^2 + 336\beta^2 c^2 k^2 + 128\beta^3 c^2 k^2 + 144ck^3 + 96\beta ck^3 + 160\beta^2 ck^3\right.$

$$+ 96\beta^2 ck^3 + 16\beta^3 ck^3 + 48k^4 + 16\beta k^4\left.\right] > 0,$$

for all $c > 0$, $\min\{0.171573c, \beta c/2\} < k < \beta c/2$ and $0 < \beta < 0.885618$.

(ii) Partially differentiating $S^*$ with respect to $c$ yields

$$\frac{\partial S^*}{\partial c} = \frac{1}{4c^2(1 + \beta)^2}\left[3c^2 + 17\beta c^2 + 5\beta^2 c^2 - \beta^3 c^2 + 12k^2 + 4\beta k^2 + 4l_1^*(c^2 - 9\beta c^2\right.$

$$+ 2\beta k^2) - 2l_1^*(5c^2 + \beta c^2 + 7\beta^2 c^2 + 3\beta^3 c^2 - 12k^2 - 4\beta k^2)\right]\right.\]

$$= \frac{1}{4D^2}\left[(35 - 3\beta^6)c^4 + (99 - 37\beta^6)\beta c^4 + 203\beta^2 c^4 + 219\beta^3 c^4 + 165\beta^4 c^4 + 53\beta^5 c^4\right.$

$$+ 16(7 - \beta^4)c^3 k + 160\beta c^3 k + 224\beta^2 c^3 k + 32\beta^3 c^3 k + 2(91 - 3\beta^4)c^2 k^2\right.$

$$+ 6(33 - \beta^4)\beta c^2 k^2 + 336\beta^2 c^2 k^2 + 128\beta^3 c^2 k^2 + 144ck^3 + 96\beta ck^3 + 160\beta^2 ck^3\right.$

$$+ 96\beta^2 ck^3 + 16\beta^3 ck^3 + 48k^4 + 16\beta k^4\left.\right] > 0,$$
\[-\beta^2 c^2 + \beta^3 c^2 - 12 k^2 - 4 \beta k^3) - 4 l_1^2 (3 c^2 - 5 \beta c^2 + \beta^2 c^2 + \beta^3 c^2 - 12 k^2 - 4 \beta k^2)\]

\[= -\frac{1}{4 D^2} \left[ (45 - \beta^7) c^4 + 121 \beta c^4 + 249 \beta^2 c^4 + 285 \beta^3 c^4 + 223 \beta^4 c^4 + 91 \beta^5 c^4 + 11 \beta^6 c^4 + 2(75 - 11 \beta^4) \beta^3 c^3 k + 2(129 - \beta^4) \beta^2 c^2 k + 72 \beta c^2 k + 255 c^2 k^2 + 251 \beta^2 c^2 k^2 + 354 \beta^3 c^2 k^2 + 210 \beta^4 c^2 k^2 + 47 \beta^5 c^2 k^2 + 3 \beta^6 c^2 k^2 + 216 c k^3 + 144 \beta c k^3 + 96 \beta^2 c k^3 + 48 \beta^3 c k^3 + 8 \beta^4 c k^3 + 84 k^4 + 76 \beta k^4 + 28 \beta^2 k^4 + 4 \beta^3 k^4 \right] < 0,

for all \(c > 0\), min \(\{0.171573 c, \beta c/2\} < k < \beta c/2\) and \(0 < \beta < 0.885618\).

(iii) Similarly, partially differentiating \(W^*\) with respect to \(c\) yields

\[\frac{\partial W^*}{\partial c} = -\frac{(1 + \beta)(1 - 4 l_1^* + 8 l_1^2)}{4}\]

\[= -\frac{1}{4 D^2} \left[ 10 c^4 + 22 \beta c^4 + 46 \beta^2 c^4 + 66 \beta^3 c^4 + 58 \beta^4 c^4 + 38 \beta^5 c^4 + 14 \beta^6 c^4 + 2 \beta^7 c^4 + 2(19 - 3 \beta^4) c^3 k + 12(3 - \beta^4) \beta c^3 k + 2(17 - \beta^4) \beta^2 c^3 k + 40 \beta^3 c^3 k + 73 c^2 k^2 + 53 \beta c^2 k^2 + 18 \beta^2 c^2 k^2 + 82 \beta^3 c^2 k^2 + (53 c - 8 k) \beta^4 c k^2 + 9 \beta^5 c k^2 + 8(9 - \beta^2) c k^3 + 48(1 - \beta^2) \beta c k^3 + 36 k^4 + 60 \beta k^4 + 28 \beta^2 k^4 + 4 \beta^3 k^4 \right] < 0,

for all \(c > 0\), min \(\{0.171573 c, \beta c/2\} < k < \beta c/2\) and \(0 < \beta < 0.885618\).

\[\Box\]

**Proof of Lemma 3.4.**

Denote \(D = 4 c^2 + 4 \beta c^2 + 6 \beta^2 c + 2 \beta^3 c + 7 c k - 2 \beta c k - \beta^2 c k + 6 k^2 + 2 \beta k^2 > 0\).

(i) Partially differentiating \(S^*\) with respect to \(\beta\) yields

\[\frac{\partial S^*}{\partial \beta} = \frac{(\beta c - c - 2 k) (2 l_1^* - 1)^2 (11 c + 4 \beta c + \beta^2 c + 10 k + 2 \beta k)}{4 e (1 + \beta)^3}\]

\[= \frac{c (1 + \beta) (\beta c - c - 2 k) (c + \beta c + k)^2 (11 c + 4 \beta c + \beta^2 c + 10 k + 2 \beta k)}{4 D^2}\]

\[< 0,

for all \(c > 0\), \(b > (2 + \beta)/2\), min \(\{0.171573 c, \beta c/2\} < k < \beta c/2\) and \(0 < \beta < 0.885618\).
(ii) Partially differentiating $W^*$ with respect to $\beta$ yields

$$
\frac{\partial W^*}{\partial \beta} = - \frac{c(1 - 4l_1^* + 8l_1^{*2})}{4} \\
= - \frac{c}{4D^2} \left[ 10c^4 + 12\beta c^4 + 34\beta^2 c^4 + 32\beta^3 c^4 + 26\beta^4 c^4 + 12\beta^5 c^4 + 2\beta^6 c^4 \\
+ 2(19 - \beta)c^3 k + 2(18 - 5\beta^2)\beta^2 c^3 k + 2(2 - \beta^2)\beta^3 c^3 k^2 + (73 - 20\beta)c^2 k^2 \\
+ 38\beta^2 c^2 k^2 + 4(11\beta c - 10k)\beta^2 ck^2 + (9\beta c - 8k)\beta^3 ck^2 + 24(3 - \beta)ck^3 + 36k^4 \\
+ 24\beta^4 k^4 + 4\beta^2 k^4 \right] \\
< 0,
$$

for all $c > 0$, $\min \{0.171573c, \beta c/2\} < k < \beta c/2$ and $0 < \beta < 0.885618$.

(iii) Similarly, partially differentiating $\Pi^*$ with respect to $\beta$ yields

$$
\frac{\partial \Pi^*}{\partial \beta} = \frac{2}{c(1 + \beta)^3} \left[ (-4l_1^*)(5c^2 - 5\beta c^2 - 3\beta^2 c^2 - 3\beta^3 c^2 + 16ck + 10k^2 + 2\beta k^2) \\
+ 2l_1^{*2}(9c^2 - 13\beta c^2 - 9\beta^2 c^2 - 3\beta^3 c^2 + 32ck + 20k^2 + 4\beta k^2) \right] \\
= \frac{c}{4D^2} \left( c^4 + 14\beta c^4 - 25\beta^2 c^4 - 52\beta^3 c^4 - 45\beta^4 c^4 - 18\beta^5 c^4 - 3\beta^6 c^4 + 16c^3 k \\
+ 128\beta c^3 k + 48\beta^2 c^3 k + 22c^2 k^2 + 216\beta c^2 k^2 + 88\beta^2 c^2 k^2 - 16\beta^3 c^2 k^2 \\
- 6\beta^4 c^2 k^2 + 144\beta c^2 k^3 + 96\beta^2 c^2 k^3 + 16\beta^3 c^2 k^3 - 16c^4 \right). 
$$

To see the ambiguity of $\partial \Pi^*/\partial \beta$, notice that for all $c > 0$, $\min \{0.171573c, \beta c/2\} < k < \beta c/2$ and $0 < \beta < 0.885618$,

$$
\lim_{\beta \to 0, k \to 0} \frac{\partial \Pi^*}{\partial \beta} = 0.015625c > 0,
$$

$$
\lim_{\beta \to 0.885618, k \to 0.171573c} \frac{\partial \Pi^*}{\partial \beta} = -0.05101415c < 0.
$$

Proof of Lemma 3.5.

Denote $D = 4c^2 + 4\beta c^2 + 6\beta^2 c + 2\beta^3 c^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2 > 0$. 

121
(i) Partially differentiating \( \Pi^* \) with respect to \( k \) yields

\[
\frac{\partial \Pi^*}{\partial k} = \frac{2(2l_1^* - 1)}{c(1 + \beta)^2} \left[ c - 2\beta c - \beta^2 c + 3k + \beta k - l_1^*(3c - 2\beta c - \beta^2 c + 6k + 2\beta k) \right]
= \frac{c^2(1 + \beta)^2(c + \beta c + k)}{D^2} \left[ c + (3\beta c - 2k)(1 - \beta^2) + (9\beta c - 8k)\beta \right]
\]

> 0,

for all \( c > 0 \), \( \min \{0.171573c, \beta c/2\} < k < \beta c/2 \) and \( 0 < \beta < 0.885618 \).

(ii) Partially differentiating \( S^* \) with respect to \( k \) yields

\[
\frac{\partial S^*}{\partial k} = 2(1 - 2l_1^*)[c - 2\beta c - \beta^2 c + 3k + \beta k - l_1^*(3c - 2\beta c - \beta^2 c + 6k + 2\beta k)]
= \frac{c^2(1 + \beta)^2(c + \beta c + k)}{D^2} \left[ c + (3\beta c - 2k)(1 - \beta^2) + (9\beta c - 8k)\beta \right]
\]
\[
< 0,
\]

for all \( c > 0 \), \( \min \{0.171573c, \beta c/2\} < k < \beta c/2 \) and \( 0 < \beta < 0.885618 \).

(iii) Similarly, partially differentiating \( W^* \) with respect to \( k \) yields

\[
\frac{\partial W^*}{\partial k} = 0. \quad \square
\]

Proof of Proposition 3.5.

Differentiating \( \Pi^* \) with respect to \( c \) yields

\[
\frac{d\Pi^*}{dc} = \frac{\partial \Pi^*}{\partial l_1^*} \times \frac{\partial l_1^*}{dc} + \frac{\partial \Pi^*}{dc}
\]

> 0,

since by lemmas 3.1, 3.2 and 3.3, \( \partial l_1^*/\partial c < 0 \), \( \frac{\partial \Pi^*}{\partial l_1^*} < 0 \) and \( \partial \Pi^*/\partial c > 0 \), respectively.

Differentiating \( S^* \) with respect to \( c \) yields

\[
\frac{dS^*}{dc} = \frac{\partial S^*}{\partial l_1^*} \times \frac{\partial l_1^*}{dc} + \frac{\partial S^*}{dc}
\]

< 0,
since by lemmas 3.1, 3.2 and 3.3, \( \partial l_1^*/\partial c < 0, \frac{\partial S^*}{\partial l_1^*} > 0 \) and \( \partial S^*/\partial c < 0 \), respectively.

Differentiating \( W^* \) with respect to \( c \) yields
\[
\frac{dW^*}{dc} = \frac{\partial W^*}{\partial l_1^*} \frac{\partial l_1^*}{\partial c} + \frac{\partial W^*}{\partial c}.
\]
Since by lemmas 3.1, 3.2 and 3.3, \( \partial l_1^*/\partial c < 0, \frac{\partial W^*}{\partial l_1^*} < 0 \) and \( \partial W^*/\partial c < 0 \), respectively, then it follows that the sign of \( dW^*/dc \) is indeterminate. \( \Box \)

**Proof of Proposition 3.6.**

Differentiating \( S^* \) with respect to \( \beta \) yields
\[
\frac{dS^*}{d\beta} = \frac{\partial S^*}{\partial l_1^*} \frac{\partial l_1^*}{\partial \beta} + \frac{\partial S^*}{\partial \beta} < 0,
\]
since by lemmas 3.1, 3.2 and 3.4, \( \partial l_1^*/\partial \beta < 0, \frac{\partial S^*}{\partial l_1^*} > 0 \) and \( \partial S^*/\partial \beta < 0 \), respectively.

Differentiating \( \Pi^* \) with respect to \( \beta \) yields
\[
\frac{d\Pi^*}{d\beta} = \frac{\partial \Pi^*}{\partial l_1^*} \frac{\partial l_1^*}{\partial \beta} + \frac{\partial \Pi^*}{\partial \beta}.
\]
Since by lemmas 3.1, 3.2 and 3.4, \( \partial l_1^*/\partial \beta < 0, \frac{\partial \Pi^*}{\partial l_1^*} < 0 \) and \( \partial \Pi^*/\partial \beta \) is indeterminate, respectively, then it follows that the sign of \( d\Pi^*/d\beta \) is also indeterminate.

Differentiating \( W^* \) with respect to \( \beta \) yields
\[
\frac{dW^*}{d\beta} = \frac{\partial W^*}{\partial l_1^*} \frac{\partial l_1^*}{\partial \beta} + \frac{\partial W^*}{\partial \beta}.
\]
Since by lemmas 3.1, 3.2 and 3.4, \( \partial l_1^*/\partial \beta < 0, \frac{\partial W^*}{\partial l_1^*} < 0 \) and \( \partial W^*/\partial \beta < 0 \), respectively, then it follows that the sign of \( dW^*/d\beta \) is indeterminate. \( \Box \)

**Proof of Proposition 3.7.**

Differentiating \( W^* \) with respect to \( k \) yields
\[
\frac{dW^*}{dk} = \frac{\partial W^*}{\partial l_1^*} \frac{\partial l_1^*}{\partial k} + \frac{\partial W^*}{\partial k} < 0,
\]

123
since by lemmas 3.1, 3.2 and 3.5, \( \partial I^*_1 / \partial k > 0, \) \( \partial W^* / \partial l_1 > 0 \) and \( \partial W^* / \partial k = 0 \), respectively.

Differentiating \( \Pi^* \) with respect to \( k \) yields

\[
\frac{d\Pi^*}{dk} = \frac{\partial \Pi^*}{\partial l_1^*} \times \frac{\partial l_1^*}{\partial k} + \frac{\partial \Pi^*}{\partial k}.
\]

Since by lemmas 3.1, 3.2 and 3.5, \( \partial I^*_1 / \partial k > 0, \) \( \partial W^* / \partial k < 0 \) and \( \partial W^* / \partial k = 0 \), respectively, then it follows that the sign of \( d\Pi^*/dk \) is indeterminate.

Differentiating \( S^* \) with respect to \( k \) yields

\[
\frac{dS^*}{dk} = \frac{\partial S^*}{\partial l_1^*} \times \frac{\partial l_1^*}{\partial k} + \frac{\partial S^*}{\partial k}.
\]

Since by lemmas 3.1, 3.2 and 3.5, \( \partial I^*_1 / \partial k > 0, \) \( \partial S^*/\partial l_1^* > 0 \) and \( \partial S^*/\partial k < 0 \), respectively, then it follows that the sign of \( dS^*/dk \) is indeterminate. \( \square \)

**Proof of Proposition 4.1.**

Analogous to the proof of proposition 3.2. \( \square \)

**Proof of Corollary 4.1.**

Analogous to the proof of corollary 3.1. \( \square \)

**Proof of Proposition 4.2.**

Denote \( D = 4c^2 + 4\beta c^2 + 6\beta^2 c + 2\beta^3 c^2 + 7ck - 2\beta ck - \beta^2 ck + 6k^2 + 2\beta k^2 > 0 \). Define

\[
\Delta I^N_1 \equiv I^*_1 - I^N_1, \ \Delta \Pi^N \equiv \Pi(I^*_1, l_2^*) - \Pi(I^N_1, l_2^N), \ \Delta S^N \equiv S(I^*_1, l_2^*) - S(I^N_1, l_2^N) \text{ and } \Delta W^N \equiv W(I^*_1, l_2^*) - W(I^N_1, l_2^N). \]

Then

\[
\Delta I^N_1 = \frac{(3 + \beta)(c + \beta c + k)(c - \beta c + 2k)^2}{2(5c + 3\beta c + 2k)D^2} > 0,
\]

\[
\Delta \Pi^N = -\frac{1}{8(5c + 3\beta c + 2k)^2D^2} \left[ 1589c^7 + 4507\beta c^7 + 8440\beta^2 c^7 + 11128\beta^3 c^7 + 10526\beta^4 c^7 + 78424\beta^5 c^7 + 4024\beta^6 c^7 + 1048\beta^7 c^7 + 61\beta^8 c^7 + 13\beta^9 c^7 + 9120\beta^6 c^7 + 18320\beta^8 c^7 \right]
\]

124
\[ +24128\beta^2 c^6 k + 21536\beta^2 c^6 k + 1024\beta^4 c^6 k + 3792\beta^5 c^6 k + 2112\beta^6 c^6 k + 768\beta^7 c^6 k \\
+ 96\beta^8 c^6 k + 4(6199 - 80\beta^6)c^5 k^2 + 4(9576 - 11\beta^6)\beta c^5 k^2 + 39160\beta^2 c^5 k^2 + 31428\beta^3 c^5 k^2 \\
+ 12076\beta^4 c^5 k^2 + 1032\beta^5 c^5 k^2 + 39368\beta^4 c^5 k^3 + 47600\beta^4 c^5 k^3 + 33464\beta^2 c^4 k^3 \\
+ 22560\beta^3 c^5 k^3 + 9016\beta^4 c^5 k^3 + 1520\beta^5 c^4 k^3 + 72\beta^6 c^4 k^3 + 38952\beta^2 c^5 k^2 + 39176\beta^3 c^5 k^2 \\
+ 17584\beta^2 c^4 k^3 + 7088\beta^3 c^3 k^4 + 2152\beta^4 c^3 k^4 + 264\beta^5 c^3 k^4 + 23584 c^2 k^5 + 20928\beta c^2 k^5 \\
+ 6784\beta^2 c^2 k^5 + 1088\beta^3 c^2 k^5 + 96\beta^4 c^2 k^5 + 7968ck^6 + 6304\beta ck^6 + 1696\beta^2 ck^6 \\
+ 160\beta^3 ck^6 + 1152k^7 + 768\beta k^7 + 128\beta^2 k^7 \right] \\
< 0, \\
\\
\Delta S^N = \frac{1}{8(5c + 3\beta c + 2k)D^2} \left[(301 - 3\beta^8)c^6 + 720\beta^2 c^6 + 1296\beta^2 c^6 + 1456\beta^3 c^6 + 1206\beta^4 c^6 \\
+ 832\beta^5 c^6 + 304\beta^6 c^6 + 32\beta^7 c^6 + 1586c^5 k + 2402\beta c^5 k + 3114\beta^2 c^5 k \\
+ 1930\beta^3 c^5 k + 254\beta^4 c^5 k + 206\beta^5 c^5 k + 198\beta^6 c^5 k + 38\beta^7 c^5 k + 332(12 - \beta^5)c^4 k^2 \\
+ 4(1049 - 16\beta^5)\beta c^4 k^2 + 4784\beta^2 c^4 k^2 + 3112\beta^3 c^4 k^2 + 192\beta^4 c^4 k^2 \\
+ 5752c^3 k^3 + 4032\beta c^3 k^3 + 3240\beta^2 c^3 k^3 + 2424\beta^3 c^3 k^3 + 736\beta^4 c^3 k^3 + 72\beta^5 c^3 k^3 \\
+ 5048c^2 k^4 + 2624\beta c^2 k^4 + 736\beta^2 c^2 k^4 + 352\beta^3 c^2 k^4 + 72\beta^4 c^2 k^4 + 16(159 - \beta^3)ck^5 \\
+ 1328\beta ck^5 + 112\beta c^2 k^5 + 576k^6 + 384\beta k^6 + 64\beta^2 k^6 \right] \\
> 0, \\
\\
\text{for all } c > 0, 0 < k < \beta c/2 \text{ and } 0 < \beta < 1. \\
\\
\text{Similarly, it is easy to show that} \\
\\
\Delta W^N = -\frac{c^2(1 + \beta)(3 + \beta)(c + \beta c + k)(c - \beta c + 2k)^2}{2(5c + 3\beta c + 2k)D^2} \left(7c^2 - 2\beta c^2 - 2\beta^2 c^2 \\
- 2\beta^3 c^2 - \beta^4 c^2 + 14ck - 18\beta ck - 22\beta^2 ck - 6\beta^3 ck + 16k^2 + 8\beta k^3 \right). \\
\\
\text{To see the ambiguous effect of low-balling on the social welfare, notice that for all } c > 0,
0 < k < βc/2 and 0 < β < 1,

\[
\lim_{\beta \to 1} \Delta W^N = \frac{2c^2k^2(3c - 3k)(2c + k)}{(4c + k)^2(4c^2 + ck + 2k^2)^2} > 0,
\]

\[
\lim_{k \to 0} \Delta W^N = \frac{(β - 1)^3(1 + β)^2(3 + β)(7 + 5β + 3β^2 + β^3)}{8(5 + 3β)^2(2 + 2β + 3β^2 + β^3)^2} < 0. \quad \Box
\]

**Proof of Proposition A.1.**

Suppose that the low risk project is optimal (i.e. \( \hat{x} = L \)). Substituting the lender's break-even constraint (A.3) into the objective function (A.2) and the incentive compatibility constraint (A.4) yields the Lagrangian for (PA.1):

\[
\sigma \bar{p}_L R_H(I) - I + \zeta \left[ (\sigma \bar{p}_L - \bar{p}_H) R_H(I) - \left( 1 - \frac{\bar{p}_H}{\bar{p}_L} \right) I \right],
\]

where \( \zeta \geq 0 \) is the Lagrange multiplier for (A.4). By the assumption on \( R_H \), the first-order conditions are necessary and sufficient for a global maximum. The first-order condition with respect to \( I \) yields

\[
\zeta = \frac{\sigma \bar{p}_L \frac{\partial R_H(I^*)}{\partial I} - 1}{\left( \bar{p}_H \frac{\partial R_H(I^*)}{\partial I} - \bar{p}_H \right) - \left( \sigma \bar{p}_L \frac{\partial R_H(I^*)}{\partial I} - 1 \right)}.
\]

It must be true that \( \zeta > 0 \). If not, \( \zeta = 0 \) would imply that

\[
\sigma \bar{p}_L \frac{\partial R_H(I^*)}{\partial I} - 1 = \sigma \bar{p}_L \left[ \frac{\partial R_H(I^*)}{\partial I} - \frac{\partial R_H(I_L)}{\partial I} \right] = 0,
\]

where the first equality follows from (A.6). Note that this in turn requires \( I^* = I_L \). But then it is not difficult to verify that (A.4) does not hold given (A.8). Thus, \( \zeta > 0 \) and (A.4) is binding. Solving (A.3) and (A.4) yields (A.10).

Now, we show that \( \hat{I}^* < I_L \). Suppose not, \( \hat{I}^* > I_L \) would imply from the above analysis that the numerator of the expression of \( \zeta \) is negative as \( R_H(I) \) is strictly concave. On the other hand, using (A.6) the denominator of the expression of \( \zeta \) can be written as

\[
\bar{p}_H \left[ \frac{\partial R_H(I^*)}{\partial I} - \sigma \frac{\partial R_H(I_L)}{\partial I} \right] - \sigma \bar{p}_L \left[ \frac{\partial R_H(I^*)}{\partial I} - \frac{\partial R_H(I_L)}{\partial I} \right]
\]
where the last inequality follows from $\sigma < 1$, $\hat{I}^* > I_L$ and $R_H(I)$ is strictly concave. But then $\zeta$ is negative, which is a contradiction. Moreover, we have shown from the above analysis that $\hat{I}^* \neq I_L$. Hence, we can conclude that $\hat{I}^* < I_L$.

Now, suppose that the high risk project is optimal (i.e. $\hat{x}^* = H$), then there is no incentive problem. Hence, the optimal loan contract is $(I_H/p_H, I_H)$, where $I_H$ solves (A.5). Which project is indeed optimal then depends on condition (A.9) \( \square \)

**Proof of Corollary A.1:**

Total differentiating (A.11) with respect to $\hat{U}^*$ and $\hat{I}^*$ and rearranging terms yields

\[
\frac{d\hat{U}^*}{d\hat{I}^*} = \sigma \bar{p}_L \frac{\partial R_H(\hat{I}^*)}{\partial I} - 1 = \sigma \bar{p}_L \left( \frac{\partial R_H(\hat{I}^*)}{\partial I} - \frac{\partial R_H(I_L)}{\partial I} \right) > 0,
\]

where the second equality follows from (A.6) and the inequality follows from the fact that $\hat{I}^* < I_L$ and $R_H(I)$ is strictly concave. This implies that the firm's expected terminal payoff $\hat{U}^*$ and the investment amount $\hat{I}^*$ change in the same direction in equilibrium. Hence, if we can show that $\partial \hat{I}^*/\partial \bar{p}_L > 0$ and $\partial \hat{I}^*/\partial p_L < 0$, we are done.

Differentiating (A.10) with respect to $\bar{p}_L$ yields

\[
\frac{\partial \hat{I}^*}{\partial \bar{p}_L} = \frac{\sigma \bar{p}_L R_H(\hat{I}^*) - \frac{\hat{I}^*}{\bar{p}_L}}{\bar{p}_L \left( (\bar{p}_H \frac{\partial R_H(\hat{I}^*)}{\partial I} - \frac{\hat{I}^*}{\bar{p}_L}) - (\sigma \bar{p}_L \frac{\partial R_H(I_L)}{\partial I} - 1) \right)} \geq \frac{\frac{\partial \hat{I}^*}{\partial \bar{p}_L}}{\bar{p}_L \left( (\bar{p}_H \frac{\partial R_H(I_H)}{\partial I} - \hat{I}^*) - (\sigma \bar{p}_L \frac{\partial R_H(I_L)}{\partial I} - 1) \right)},
\]

where the first inequality follows from (A.9). Clearly, the numerator of the expression of $\partial \hat{I}^*/\partial \bar{p}_L$ is positive. By the expression of $\zeta$ and the proof in proposition A.1, the denominator is also positive. Hence, $\partial \hat{I}^*/\partial \bar{p}_L$ is positive.
Similarly, differentiating (A.10) with respect to $\bar{p}_H$ yields

$$\frac{\partial \hat{I}^*}{\partial \bar{p}_H} = -\frac{R_H(\hat{I}^*) - \frac{\hat{I}^*}{p_L}}{\bar{p}_H - \frac{\partial R_H(\hat{I}^*)}{\partial I} - \frac{\partial R_H(\hat{I}^*)}{\partial I} - 1} \left(\frac{\bar{p}_L(1 - \sigma)R_H(\hat{I}^*)}{(\bar{p}_L - \bar{p}_H)\left\{\bar{p}_H - \frac{\partial R_H(\hat{I}^*)}{\partial I} - \bar{p}_L - (\sigma \bar{p}_L - \frac{\partial R_H(\hat{I}^*)}{\partial I} - 1)\right\}\right) < 0,$$

where the second equality follows from (A.10). □

**Proof of Proposition A.2.**

Using the result of corollary A.1, we know that $\sigma \bar{p}_LR_H(I_B) - I_B > \sigma \bar{p}_LR_H(\hat{I}^*) - \hat{I}^*$ since $p_B < \bar{p}_H$. Then, we have

$$U(q^*) - \hat{U}^* > \left[\phi + (1 - \phi)(1 - q^*)\right]\left\{\sigma \bar{p}_LR_H(I_H(q^*)) + \Omega_H(q^*, I_H(q^*)) - I_H(q^*) \right\}$$

$$> 0,$$

where the second inequality follows from the fact that, given (A.14), the high risk project is optimal and the optimal investment amount is $I_H(q^*)$.

Furthermore, we have

$$\frac{\partial U(q^*)}{\partial q^*} = \left[\phi \pi_G + (1 - \phi)(1 - q^*)p_B\right] \frac{\partial R_H}{\partial I} \times \frac{\partial I_H}{\partial q^*} - (1 - \phi)p_B R_H(I_H(q^*))$$

$$- \left[\phi + (1 - \phi)(1 - q^*)\right] \left(\frac{\partial I_H}{\partial q^*} - \frac{\partial \Omega_H}{\partial q^*} - \frac{\partial \Omega_H}{\partial I} \times \frac{\partial I_H}{\partial q^*}\right)$$

$$+ (1 - \phi)[I_H(q^*) - \Omega_H(q^*, I_H(q^*))] + (1 - \phi)\left[\sigma \bar{p}_LR_H(I_B) - I_B\right]$$

$$= (1 - \phi)\left[\sigma \bar{p}_LR_H(I_B) - I_B\right] - \sigma \bar{p}_LR_H(I_H(q^*))$$

$$+ (1 - \phi)[I_H(q^*) - \Omega_H(q^*, I_H(q^*))] + \left[\phi + (1 - \phi)(1 - q^*)\right] \frac{\partial \Omega_H}{\partial q^*}$$

$$= (1 - \phi)\left[\sigma \bar{p}_LR_H(I_B) - I_B\right] - \sigma \bar{p}_LR_H(I_H(q^*)) + [1 - (1 - p_B)\nu^* \alpha]I_H(q^*),$$

where the second equality follows from (A.17) that

$$\frac{\partial R_H}{\partial I} = \frac{1 - \frac{\partial \Omega_H}{\partial I}}{\pi_H(q^*)} = \frac{[\phi + (1 - \phi)(1 - q^*)](1 - \frac{\partial \Omega_H}{\partial I})}{\phi \pi_G + (1 - \phi)(1 - q^*)p_B},$$

128
and the third equality follows from

\[
(1 - \phi)[I_H(q^*) - \Omega_H(q^*, I_H(q^*))] + \left[ \phi + (1 - \phi)(1 - q^*) \right] \frac{\partial \Omega_H}{\partial q^*} \\
= \frac{(1 - \phi)\left\{ \phi + (1 - \phi)(1 - q^*)[1 - (1 - p_B)\nu^*\alpha]\right\} I_H(q^*)}{\phi + (1 - \phi)(1 - q^*)} - \frac{\phi(1 - \phi)(1 - p_B)\nu^*\alpha I_H(q^*)}{\phi + (1 - \phi)(1 - q^*)} \\
= \frac{(1 - \phi)\left\{ \phi + (1 - \phi)(1 - q^*) - (1 - p_B)\nu^*\alpha[(1 - \phi)(1 - q^*) + \phi]\right\} I_H(q^*)}{\phi + (1 - \phi)(1 - q^*)} \\
= (1 - \phi)[1 - (1 - p_B)\nu^*\alpha]I_H(q^*).
\]

Hence, \(\partial U(q^*)/\partial q^* > 0\) if, and only if

\[
\sigma\bar{p}LR_H(I_B) - I_B > p_B R_H(I_H(q^*)) - [1 - (1 - p_B)\nu^*\alpha]I_H(q^*),
\]

which is equivalent to the condition specified in the proposition. This completes the proof.

\(\square\)